

Leibniz-Institut für Astrophysik Potsdam

Cosmological Radiative Transfer and the Ionisation of the Intergalactic Medium

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
in der Wissenschaftsdisziplin “Astrophysik”

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von

Adrian M. Partl

geboren am

16.08.1982 in Zürich, CH

Potsdam, im Mai 2011

To my parents and Bettina.

Abstract

Gravitational instability during the evolution of the Universe formed a large scale filamentary structure, known as the cosmic web. Baryons embedded in this cosmic web constitute the intergalactic medium (IGM) with hydrogen making up around 76% of its total mass. A large fraction of the baryons are kept in a highly ionised state by an ultraviolet (UV) background field. High resolution spectra of quasars (QSOs) reveal a vast amount of absorption features blue-ward of the QSO's Lyman- α emission line: the H I Lyman- α forest. The Ly α forest traces the filamentary cosmic web, where a small remaining fraction of neutral hydrogen (H I) in the filaments produces all the absorption. The Ly α forest thus provides a unique method to constrain the history of the cosmic web. However due to the high ionisation state of the IGM, the formation history of the cosmic web can only be inferred with a detailed knowledge of the UV background, requiring the usage of both observations and simulations.

In the first part of this thesis we wish to characterise the formation of the structure giving rise to the Ly α forest. We thus derive the Ly α absorber number density evolution and the differential column density distribution. The number density evolution of high column density absorbers reveal a yet unknown dip in the number density at around $z \sim 2.1$. We further show, that this depression in the absorber number density is directly connected to a dip in the differential column density distribution at $N_{\text{HI}} > 10^{14} \text{ cm}^{-2}$. This is most likely related to the high star-formation rate at $z \sim 2$.

A small number of absorbers in the Ly α forest stem from ionic metal lines, indicating that some parts of the gas causing the Ly α forest have been metal enriched. In this thesis we investigate the statistical properties of enriched hydrogen absorbers. Further, we constrain the volume averaged $N_{\text{HI}} - N_{\text{CIV}}$ relation, which shows a constant relation between the two constituents at high N_{HI} . However, we find that the $N_{\text{HI}} - N_{\text{CIV}}$ relation drops off steeply at $N_{\text{HI}} \sim 10^{15.2} \text{ cm}^{-2}$. This indicates that the IGM around galaxies is only metal enriched up to a characteristic radius. We argue that these findings are similar to results of simulated density-metallicity relations. These observations help to provide constraints on the coupling between galaxies and their environment.

In the second part of this thesis, we focus on solving the 3D radiative transfer equation numerically. In recent years, solving the 3D radiative transfer has become a new exciting field in numerical cosmology. We have developed a method of adapting the cosmological radiative transfer code `CRASH2` for distributed memory clusters. We show that the resulting parallel MPI application `pCRASH2` performs and scales well, enabling the simulation of complex and large problems with high resolution.

In the third part of this thesis, we model the QSO line of sight proximity effect using radiative transfer in a cosmological context. Due to the UV radiation emitted by QSOs, the hydrogen ionisation fraction increases in their vicinity. This proximity effect can be used to determine the UV background flux. With simulations, we confirm the assumption

Abstract

of geometrical dilution used in analytical formulations of the effect. Furthermore, we find a relation between the environmental density fluctuations and the proximity effect signal. The density fluctuations of the cosmic web are responsible for a large scatter in measurements of the proximity effect signal derived from Ly α forest spectra. The scatter is found to increase with decreasing redshift and decreases with increased QSO luminosity. Furthermore, we find that the distribution of normalised optical depths, resembles a log-normal distribution at large distances from the QSO. However the distribution becomes increasingly skewed when approaching the QSO. The proximity effect strength is found to be weakly correlated with the host halo mass, and tightly correlated with the mean density in the large scale environment of a QSO host. If a large scale overdensity is present inside a sphere of $10 h^{-1}$ Mpc comoving radius, we show that the proximity effect strength decreases, resulting in an overestimation of the UV background flux. We quantify the dependence of this bias on the QSO luminosity, the host halo mass, and the redshift. These results will help to correct this environmental bias.

Zusammenfassung

Während der kosmischen Entwicklung formten sich durch die Gravitationsinstabilität filamentartige großräumige Strukturen, welche als das kosmische Netz bezeichnet werden. Die in diesem kosmischen Netz enthaltenen Baryonen bilden das intergalaktische Medium (IGM), dessen Masse zu 76% aus Wasserstoff besteht. Ein großer Teil der Baryonen wird durch die ultraviolette (UV) Hintergrundstrahlung hoch ionisiert. Hochaufgelöste Quasarspektren zeigen eine Vielzahl von Absorptionslinien blauwärts der Lyman- α Emissionslinie. Dieser Ly α -Wald hat seine Ursache in dem im kosmischen Netz enthaltenen restlichen neutralen Wasserstoff (HI). Der Ly α -Wald bietet somit eine einzigartige Möglichkeit, die Entwicklungsgeschichte des kosmischen Netzes zu erforschen. Durch den hohen Ionisierungsgrad des intergalaktischen Mediums kann jedoch die Entwicklungsgeschichte des kosmischen Netzes nur mit detaillierter Kenntnis des UV-Hintergrundfeldes abgeleitet werden. Dies erfordert eine kombinierte Analyse der Beobachtungen mit Simulationen.

Mit Hilfe von hochaufgelösten Spektren des Ly α -Waldes werden im ersten Teil dieser Arbeit die Entwicklung der Ly α -Absorberanzahlsdichte und der differentiellen Säulendichteverteilung abgeleitet. Die Entwicklung der Absorberanzahlsdichte von Systemen mit hoher Säulendichte zeigt einen bisher unbekanntem Abfall in der Anzahlsdichte bei $z \sim 2.1$. Darüber hinaus wird gezeigt, dass dieser Abfall direkt mit einem Abfall der differentiellen Säulendichteverteilung bei Säulendichten $N_{\text{HI}} > 10^{14} \text{ cm}^{-2}$ zusammenhängt. Dies könnte auf die hohe Sternenstehungsrate bei $z \sim 2$ zurückzuführen sein.

Ein kleiner Teil der Absorptionslinien im Ly α -Wald stammt von hoch ionisierten Metalllinien, was darauf hinweist, dass der Ly α -Wald partiell mit Metallen angereichert ist. In dieser Arbeit werden verschiedene statistische Eigenschaften von solchen assoziierten Absorbern bestimmt. Es wird die volumengemittelte $N_{\text{HI}}-N_{\text{CIV}}$ -Relation untersucht. Das Verhältnis der beiden Ionisationsstufen ist bei hohen Säulendichten etwa konstant, fällt aber bei einer Säulendichte von $N_{\text{HI}} \sim 10^{15.2} \text{ cm}^{-2}$ steil ab. Dieses Verhalten deutet darauf hin, dass das IGM nur bis zu einem charakteristischen Radius rund um die Galaxien mit Metallen angereichert ist. Die beobachtete $N_{\text{HI}}-N_{\text{CIV}}$ -Relation gleicht den Resultaten von simulierten volumengemittelten Dichte-Metallizitätsrelationen und liefert somit Anhaltspunkte zur Physik der Kopplung von Galaxien an ihre Umgebung.

Im zweiten Teil dieser Arbeit wird der Fokus auf die Modellierung von Strahlungstransportprozessen gelegt. Simulationen von 3D Strahlungstransportprozessen haben sich in den letzten Jahren zu einem neuen und aufregenden Feld der numerischen Kosmologie entwickelt. In dieser Arbeit wird eine Methode entwickelt, welche es ermöglicht, den kosmologischen Monte-Carlo-Strahlungstransport Code **CRASH2** auf Grossrechnern mit verteiltem Arbeitsspeicher zu benutzen. Es wird gezeigt, dass die daraus resultierende MPI-Anwendung **pCRASH2** gute Leistung erzielt und sehr gut skaliert. Dies ermöglicht hoch aufgelöste Simulationen von großen und komplexen Problemen.

Zusammenfassung

Im dritten Teil wird der Quasar-Umgebungseffekt mit Hilfe von Strahlungstransport-Simulationen untersucht. Aufgrund der von Quasaren emittierten UV-Strahlung nimmt der Anteil an ionisiertem Wasserstoff in der Quasarumgebung zu. Dieser Umgebungseffekt wird für die Bestimmung des UV-Hintergrundflusses benutzt. Mit den Simulationsergebnissen wird die Annahme der geometrischen Verdünnung der ionisierenden Strahlung bestätigt, welche für analytische Formulierungen des Effekts verwendet wird. Es wird ein Einfluss der Dichtefluktuationen auf das Signal des Umgebungseffektes festgestellt, welches aus dem Ly α -Wald gewonnen wird. Die Dichtefluktuationen sind für eine große Streuung des Signals des Umgebungseffektes verantwortlich. Die Streuung nimmt mit abnehmender Rotverschiebung zu und darüber hinaus mit zunehmender Quasarleuchtkraft ab. Die Verteilung der normierten optischen Tiefen entspricht bei großem Abstand zum Quasar einer log-normalen Verteilung, welche mit kleiner werdendem Abstand zum Quasar zunehmend asymmetrisch wird. Die Stärke des Umgebungseffektes korreliert wenig mit der Masse des Halos. Jedoch besteht eine enge Korrelation der Stärke mit der mittleren großräumigen Dichte um den Quasar. Liegt eine großräumige Überdichte in einer Sphäre von mitbewegten $10 h^{-1}$ Mpc um den Quasar vor, zeigt sich eine Abnahme der Stärke des Umgebungseffektes. Dies kommt einer Überschätzung des UV-Hintergrundflusses gleich. Die Abhängigkeit dieses Bias von der Quasarleuchtkraft, der Masse des Wirtshalos und der Rotverschiebung wird quantifiziert.

Contents

Abstract	iii
Zusammenfassung	v
1. Introduction	1
1.1. Structure formation in the cosmological framework of Λ CDM	1
1.2. From first stars to a transparent Universe	4
1.3. The evolution of the ionised IGM after reionisation	7
1.4. Thesis outline	10
2. The forest CIV at $2 < z < 3.5$	13
2.1. Introduction	13
2.2. Data and Voigt profile fitting analysis	15
2.2.1. Voigt profile fitting analysis	17
2.2.2. System	20
2.3. The relations between H I and C IV column densities	23
2.4. Comparisons with the optical depth analysis	27
2.4.1. The $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation	28
2.4.2. Velocity offset between H I and C IV	29
2.4.3. The fraction of H I pixels associated with C IV	30
2.5. Simple shell model for the $N_{\text{HI}} - N_{\text{CIV}}$ relation	32
2.5.1. The observed rate of incidence of non-highly ionised C IV absorbers	32
2.5.2. Simple shell model	33
2.6. Summary	34
3. The IGM H I column density evolution	37
3.1. Introduction	38
3.2. Data and Voigt profile fitting	40
3.3. Comparison with previous studies using Ly α only	44
3.3.1. Absorber number density evolution dn/dz	44
3.3.2. The differential column density distribution function	48
3.4. Analysis using higher-order transition lines	49
3.4.1. The mean number density evolution	49
3.4.2. The differential column density distribution function	53
3.5. Characteristics of the metal-enriched forest	56
3.5.1. Method	57
3.5.2. Results	57

3.6. Conclusions	65
3.7. Data	67
4. Enabling parallel computing in CRASH	69
4.1. Introduction	69
4.2. CRASH2: Summary of the algorithm	71
4.3. Parallelisation strategy	72
4.3.1. Domain decomposition strategy	74
4.3.2. Parallel photon propagation	76
4.3.3. Parallel pseudo-random number generators	78
4.4. Performance	78
4.4.1. Test 0: Convergence test of a pure-H isothermal sphere	78
4.4.2. Comparing pCRASH2 with CRASH2	81
4.4.3. Scaling properties	89
4.4.4. Dependence of the solution on the number of cores	94
4.4.5. Thousand sources in a large cosmological density field	94
4.5. Summary	97
4.6. Generating the Peano-Hilbert curve	98
5. Cosmological radiative transfer for the LOS proximity effect	101
5.1. Introduction	101
5.2. Simulations	104
5.2.1. Initial realisation (IGM model)	104
5.2.2. Model and calibration of the intergalactic medium	105
5.3. Radiative transfer in cosmological simulations	108
5.3.1. Method	108
5.3.2. UV background field	109
5.3.3. QSO radiation	110
5.3.4. Simulation setup and convergence	111
5.4. Line-of-sight proximity effect	112
5.5. Radiative transfer effects	113
5.5.1. The overionisation profile	113
5.5.2. Shadowing by Lyman limit systems	114
5.5.3. Diffuse recombination radiation	115
5.6. Radiative transfer on the line-of-sight proximity effect	118
5.6.1. Strength parameter	118
5.6.2. Additional models	118
5.7. Mean proximity effect	119
5.7.1. Mean normalised optical depth	119
5.7.2. Performance of the SED correction term	120
5.7.3. Influence of the large-scale environment	121
5.7.4. Influence of the host environment	124
5.7.5. Redshift evolution	126
5.7.6. Luminosity dependence	127
5.8. Proximity effect strength distribution	127
5.9. Conclusions	129

5.10. Strong absorber contamination of the normalised optical depth	131
6. Large scale environmental bias on the LOS proximity effect.	135
6.1. Introduction	135
6.2. Simulation	137
6.2.1. Distribution of baryons in the IGM	137
6.2.2. Model and calibration of the intergalactic medium	138
6.3. Method	142
6.3.1. Line-of-sight proximity effect	142
6.3.2. Environments	145
6.3.3. Measuring the proximity effect in spectra	146
6.4. Results	148
6.4.1. Null hypothesis: Random locations	148
6.4.2. Dependence on the spectra signal to noise	149
6.4.3. Large scale environment	150
6.4.4. Proximity effect strength as function of halo mass	153
6.4.5. Influence of the large scale overdensity	155
6.5. Conclusion	158
7. Summary & Conclusions	161
A. CRASH: Cosmological RAdiative transfer Scheme for Hydrodynamics	167
A.1. Overview of radiative transfer schemes	167
A.1.1. The radiative transfer equation	169
A.2. How it all works - The CRASH RT scheme	172
A.2.1. Coloured photon packets	172
A.2.2. Package propagation	174
A.2.3. Solving for the chemistry	175
A.2.4. Cross-sections and rates	177
Bibliography	183

Very strange people, physicists - in my experience the ones who aren't dead are in some way very ill.

Mr. Standish "The Long Dark Tea-Time Of The Soul" by
Douglas Adams

1

Introduction

In this chapter we give a brief overview of the Universe's evolution. We focus on its ionisation history and describe how the change of the initially neutral Universe to a completely ionised one occurred. In this introduction only a broad overview is given. More detailed introductions to the various topics of this thesis are given individually in each chapter.

1.1. Structure formation in the cosmological framework of Λ CDM

According to the widely accepted Λ CDM cosmological model, the history of the Universe is thought to have started around 13.75 ± 0.11 Gyr ago. During this time, the fascinating story of our Universe unfolded up to the present day. The various chapters of this story represent the many metamorphoses the Universe lived through, such as the era of nucleosynthesis, the era of recombination, the dark ages, the era of reionisation, and the era of galaxy formation. Surprisingly, the main actor in large parts of this story is not the matter we deal with in everyday life. According to the Λ CDM model, the structure we find today in the Universe would not exist without Cold Dark Matter (CDM, from now on referred to as dark matter or DM) and dark energy (expressed through a cosmological constant Λ in the Einstein field equation, which describe the coupling between the geometry of spacetime with matter and energy).

The possible existence of some form of unobserved and thus dark matter was first proposed by Zwicky (1933) and Zwicky (1937) to account for the large velocity dispersion in the Coma galaxy cluster. The discrepancy between the mass needed to account for the high velocity dispersions observed in clusters and the mass derived from the observable luminous matter is not confined to the Coma cluster, but has been observed in various other clusters as well (e.g. Merritt 1987; Carlberg et al. 1997). However not only galaxy clusters show signs of unaccounted for mass. Also the rotation curves of galaxies show indications of missing mass. In the outskirts of galaxies the rotations curves do not follow the Keplerian law which would dictate a decrease in rotation velocity with larger galactocentric radii. Instead the rotation curves stay constant (Volders 1959; Rubin & Ford 1970; van Albada et al. 1985; de Blok & Bosma 2002), which indicates the presence of large amounts of unobserved mass at large radii. Further evidence for a dark matter component is found in observations of gravitational lenses (e.g. Clowe et al. 2004) and the observed galaxy cluster temperatures derived from X-ray emission (e.g. White et al. 1993; Evrard et al. 1996).

1. Introduction

However when the total energy content of the Universe is inferred today, one finds that dark matter together with the baryonic matter only make up for 27% of the total energy (Jarosik et al. 2011). The remaining 74% are assumed to stem from a cosmological constant Λ , which is also known as dark energy (e.g. Carroll 2001). This dark energy is responsible for an accelerated expansion of the Universe **at** $z \lesssim 1$, which is also observationally established through distance measurements of galaxies using supernovae as standard candles (Colgate 1979; Amanullah et al. 2010). The assumed existence of dark energy, dark matter, and the observed flatness of space are the three main constituents of the Λ CDM model.

The story of the Universe started with the big bang which left our Universe in a very hot and dense state. The young Universe expanded rapidly (e.g. Peebles 1993). While the Universe expanded, it also cooled down and with decreasing temperature electrons, protons, and neutrons formed. While this plasma further cooled, deuterium and helium nuclei assembled. However the temperature of the Universe was still high enough, so that the electrons were kept separated from the hydrogen, deuterium and helium nuclei by collisions. At this early stage, around 75% of the baryonic matter were hydrogen and 25% were helium with traces of deuterium and other slightly more massive atoms (Kolb & Turner 1990; Tytler et al. 2000). More massive atoms, e.g. such as carbon, nitrogen, and oxygen, had not yet been formed.

As the Universe cooled further to $T \approx 3000$ K, the collisions between matter and photons ceased, the photons decoupled from the plasma and the electrons recombined with the atom nuclei. This last scattering of the photons occurred at a redshift of $z \approx 1100$. The Universe was transparent from then on and photons from the last scattering surface were able to travel basically unobstructed through space. Today these photons are observed as the cosmic microwave background (CMB) (Penzias & Wilson 1965; Mather et al. 1994; Jarosik et al. 2011) and the best measurement has been obtained by the WMAP satellite¹. Higher resolution maps of the CMB will be obtained in the near future with results from the PLANCK mission².

Observations of the CMB show, that matter was distributed almost homogeneously at a redshift of $z \approx 1100$. Only small fluctuations in relation to the mean density of $\rho/\bar{\rho} \sim 10^{-5}$ were present in the early Universe (de Bernardis et al. 2000; Hanany et al. 2000). These fluctuations later evolved due to gravitational collapse into the structure we observe today. After the photons decoupled from the matter, the Universe entered what came to be known as the dark ages, since not a single star had formed yet. For the time between $z = 1100$ and $z \approx 30$ (e.g. Loeb & Barkana 2001) the perturbations in the matter density grew more and more and eventually collapsed to the first luminous objects in the Universe.

In Λ CDM, the evolution of the Universe after the last scattering is fully characterised by the primordial power spectrum of Gaussian density fluctuations, the mean matter density, and the initial temperature, density and molecular composition of the baryonic gas. The evolution of the Universe as such is further expressed through the expansion factor $a(t) = (1 + z)^{-1}$. The expansion factor itself is given by the Hubble parameter $H(t) = \dot{a}/a$. From the Einstein field equation it is possible to obtain the Friedman

¹<http://map.gsfc.nasa.gov/>

²<http://sci.esa.int/planck>

equation $H^2(a) = H_0^2 (\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_\Lambda)$ which describes the expansion history of a flat Universe (e.g. Peacock 1999), where the Hubble constant at $z = 0$ is according to WMAP 7 $H_0 = 70.4_{-1.4}^{+1.3}$ km/s/Mpc (Jarosik et al. 2011). The expansion of the Universe is thus described through a combination of energy density parameters $\Omega_i = \rho_i/\rho_c$, where ρ_i is the energy density of the various components in the Universe, such as for instance baryonic matter or dark energy. Furthermore ρ_c denotes the critical density which characterises the density that is required for the Universe to be spatially flat. The following energy density parameters are derived from the CMB according to the WMAP 7 results: the baryon density parameter $\Omega_b = 0.0456 \pm 0.0016$, the dark matter density parameter $\Omega_c = 0.227 \pm 0.014$, and the dark energy density parameter $\Omega_\Lambda = 0.728_{-0.016}^{+0.015}$.

Any structure in the dark matter distribution of the Universe shows three stages during its evolution, as has already been predicted by Zel'dovich (1970) and Shandarin & Zel'dovich (1989). A perturbation first collapses along its shortest axis into a sheet like structure, also known as Zel'dovich pancake. Then the structure further collapses along the second axis to form a filament, and then further collapses to form a compact clump of matter known as halo. All these different structures make up the so-called cosmic web. In order to follow this evolution not only in isolation, but in a more realistic environment with tidal forces acting between the different structures, numerical simulations are needed. Starting out from the initial conditions obtained from the CMB measurements, such simulations are able to trace the evolution of the Universe forward in time (e.g. Springel et al. 2005; Gottlöber et al. 2010) and reveal the formation of a fine web of filaments and sheets. The success story of the Λ CDM model is based in parts on the agreement between these numerical results and the wealth of observed properties of this cosmic web, such as the distribution of galaxies (e.g. Hawkins et al. 2003), which is successfully regained with simulations (e.g. Cole et al. 1998; Benson et al. 2001). Numerical simulations thus provide the field of cosmology and galaxy formation a laboratory to explore many different physical processes involved in the formation of galaxies and eventually stars.

The baryonic component from which the first stars and eventually galaxies formed, predominantly follows the evolution of the dark matter. A close relation between the two components exists as long as an equilibrium configuration between the baryonic gas pressure and the gravitational forces develops. However, as has been shown by Jeans (1928), a cloud of gas becomes gravitationally unstable if the free-fall time is less than the sound-crossing time. The internal gas pressure is then unable to act against gravitational collapse. A cloud with an overdensity $\delta = \rho/\bar{\rho}$, where ρ is the matter density and the barred quantity its cosmic mean, becomes gravitationally unstable if the cloud's matter content (dark matter plus baryons) is more massive than the Jeans mass

$$M_J \approx 6 \times 10^7 M_\odot \delta^{-1/2} \left(\frac{T}{10^3 \text{ K}} \right)^{3/2} \left(\frac{1+z}{10} \right)^{-3/2} \left(\frac{\mu}{1.22} \right)^{-3/2} \left(\frac{\Omega_m h^2}{0.13} \right)^{-1/2} \quad (1.1)$$

(e.g. Loeb 2006). Here T is the gas temperature, μ is the mean molecular mass in units of hydrogen mass and h is the Hubble parameter $h = H_0/100$ km/s/Mpc. Eq. 1.1 is only valid for an ideal gas and a spherical symmetric cloud. However Eq. 1.1 can also be used as a good approximation for non spherical clouds.

Whenever a cloud of gas reached the Jeans mass, it collapsed. The first bound objects formed through this process are thought to be the first objects to have emitted light

1. Introduction

after the last scattering and thus ended the dark ages (e.g. Loeb & Barkana 2001). The collapse is stopped as soon as the gas pressure is again large enough to prevent further collapse (Couchman & Rees 1986; Haiman & Loeb 1997; Ostriker & Gnedin 1996). These first small objects formed the building blocks of Λ CDM hierarchical structure formation, which started, depending on the model parameters, at around $z \approx 30$ (e.g. Wise & Abel 2008). Subsequently large objects formed in Λ CDM through accretion and merging of smaller objects (e.g. Peebles 1993; Peacock 1999; Mo et al. 2010).

1.2. From first stars to a transparent Universe

The first filaments and halos that formed during the dark ages are thought to have been the nurseries of the first stars. In this weblike structure, stars 100 to 1000 times heavier than the Sun have formed and eventually ended the dark ages by lighting up the Universe (e.g. Yoshida et al. 2003). From CMB polarisation measurements, the formation of the first stars must have happened between $z \sim 30$ and $z \sim 10$ (Page et al. 2007).

Stars are only able to form in cold and dense gas clouds, requiring a coolant to cool the gas in a cloud to around $T \sim 200$ K (Bromm et al. 2002; Abel et al. 2002). Since the gas that formed the first stars lacked any metals that could act as coolant, the only possible mechanism to remove excess energy from the gas was through molecular hydrogen. In order for this process to work, enough molecular hydrogen needed to be available primordially. Molecular hydrogen had to be sufficiently abundant such that the cooling time scale in the cloud became smaller than the free-fall time scale, a necessary condition for star formation (Rees & Ostriker 1977; Silk 1977). The cooling time for a $(3 - 4)\sigma$ peak in the primordial density field is smaller than the free-fall time for a gas with a molecular hydrogen fraction larger than 10^{-4} and a temperature larger than $T \sim 10^3$ K (e.g. Bromm & Larson 2004). Molecular hydrogen could only cool primordial gas with $T < 10^4$ K through radiative de-excitation of collisionally excited molecular hydrogen. The first stars that formed in the Universe through this process are known as Population III (Pop III) stars.

The radiation produced by these Pop III stars however dissociated molecular hydrogen in the Lyman-Werner band (e.g. Glover & Brand 2001), thus depleting gas that might have formed new stars at a later time from its coolant. This mechanism quenches the formation of stars after the first stars have appeared (Haiman et al. 1997). The first stars however already formed metals, which could then act as coolant for star formation if the metals are transported back into the surrounding medium (e.g. through supernovae). For this mechanism to work, the gas needed to be enriched up to a metallicity of around $Z \sim 10^{-4} Z_{\odot}$ (Schneider et al. 2002).

The rate with which the primordial gas was enriched depends largely on the masses of the Pop III stars and the way they ended their lifetime (e.g. Ciardi & Ferrara 2005). Stars with masses of $10 \lesssim M \lesssim 30 M_{\odot}$ perished in supernova explosions, which ejected gas into the stars' surrounding and enriched it with metals. Stars with masses of $30 \lesssim M \lesssim 140 M_{\odot}$ and $M \gtrsim 260 M_{\odot}$ directly collapsed to a black hole, swallowing most of the stellar material. These stars did not return any metals that formed in the process to its surroundings but may instead have been the progenitors of supermassive black holes in the centre of galaxies. Stars with masses of $140 \lesssim M \lesssim 260 M_{\odot}$ however

were good candidates for enriching the primordial medium. They are thought to have ended in pair-instability supernovae which completely disrupted the whole star, leaving no remnant. Thus all the star's gas is transported into its surrounding (e.g. Heger & Woosley 2002). Pair-instability supernovae provided an efficient metal enrichment process in the early Universe (e.g. Bromm et al. 2009). The metal enriched gas was then able to form stars again through metal cooling, gradually forming the building blocks of the first galaxies (Bromm & Larson 2004).

Ultra violet (UV) photons originating in these first objects started to ionise their surrounding inter galactic medium (IGM), which started the era of reionisation. The first galaxies and stars eventually carved out ever-growing H II regions out of the neutral IGM. The ionisation fronts of these H II regions propagated faster in regions with low density than in dense regions, causing underdense regions to become ionised faster. During this initial stage, the H II regions of these individual sources were separated by the remaining neutral hydrogen. The H II regions continued to grow as long as the source produced UV photons, until an equilibrium between photoionisation and recombination was reached (Strömgren 1939). At some stage the large individual H II regions started to overlap and merged to one large H II region (e.g. Gnedin 2000; Iliev et al. 2006; Trac & Cen 2007; McQuinn et al. 2007; Baek et al. 2009). Now the expansion of these compound H II regions were not driven by one source anymore, but by all the sources enclosed in the compound H II region. This sudden increase in the number of available UV photons accelerated the reionisation of the Universe, and the remaining neutral regions became quickly ionised. Only clumps which were dense enough to self-shield themselves from the UV radiation remained neutral. However even these neutral pockets were gradually ionised by the combined UV radiation of all the sources in the Universe (e.g. Barkana & Loeb 2001).

Reionisation thus started out in patches where enough dense material was available to form the first galaxies. The first phase of reionisation thus proceeded in an inside out fashion, and after the overlap of H II regions reionisation changed into an outside in process, gradually ionising the remaining clumps of neutral gas. Reionisation ended when most of the Universe's hydrogen became ionised and the UV radiation of all the sources combined into one UV background flux field. Simulations of reionisation predict the Universe to have become completely ionised at a redshift around $z \approx 6 - 8$ (e.g. Ciardi et al. 2000; Miralda-Escudé et al. 2000; Iliev et al. 2006; Zahn et al. 2007; Shin et al. 2008; Lee et al. 2008).

Unfortunately modelling the evolution of reionisation is susceptible to weakly constrained parameters. Present day models all depend strongly on the evolution of the UV emitting sources responsible for reionisation. One large uncertainty is the stellar initial mass function during the era of reionisation which might differ from what is observed today (Larson 2003; Schneider et al. 2006). The ionisation history is found to be strongly dependent on the initial mass function and the type of sources responsible for reionisation. Due to these uncertainties, reionisation might even have ended at redshifts larger than $z \gtrsim 7$ (Ciardi et al. 2003; Wyithe & Cen 2007).

Another very uncertain parameter which is important in modelling reionisation is the number of UV photons actually escaping the object that produced them. The precise value of this escape factor depends on the amount and distribution of dust in the galaxy and is still under heavy debate. The escape fraction is found to range from

1. Introduction

$f_{\text{esc}} \lesssim 0.1$ up to $f_{\text{esc}} \sim 0.8$ (e.g. Gnedin et al. 2008; Wise & Cen 2009; Razoumov & Sommer-Larsen 2010; Yajima et al. 2010). Next generation radio interferometers such as LOFAR³, MWA⁴, and SKA⁵ will be able to shed light on the era of reionisation and provide direct measurements of reionisation's topology. This will help to eliminate some of these uncertainties, completing the picture of early star and galaxy formation.

Observationally, the only constraints for the era of reionisation available today are obtained through CMB polarisation measurements and the spectra of high-redshift quasars (QSO). From the WMAP polarisation data it is possible to measure the Thomson scattering optical depth of CMB photons $\tau_e = 0.087 \pm 0.014$ (Jarosik et al. 2011). Assuming an instantaneous reionisation model (Page et al. 2007) the reionisation redshift can be derived from the Thomson scattering optical depth. The latest WMAP 7 results indicate a reionisation redshift of $z = 10.4 \pm 1.2$. One drawback of this method however is that reionisation is most likely not an instantaneous process, but evolves in patches with regions showing faster reionisation than others.

Another method of constraining the reionisation redshift is, as mentioned, to look at spectra of high-redshift QSOs and to see, whether the transmission spectra blueward of the QSO's Lyman- α emission line is completely absorbed by neutral hydrogen or not. Already a tiny amount of neutral hydrogen produces large optical depths for Lyman- α photons, which leads to the Gunn-Peterson optical depth

$$\tau_{\text{GP}} = \frac{\pi e^2}{m_e c} f_e \frac{\lambda_\alpha}{H(z)} n_{\text{HI}} \quad (1.2)$$

(Gunn & Peterson 1965), where m_e is the electron mass, c the speed of light, $f_e = 0.41641$ denotes the oscillator strength of the H I Ly α transition, λ_α is the wavelength of the Ly α transition, $H(z)$ is the Hubble parameter at redshift z , and n_{HI} denotes the number density of neutral hydrogen. For neutral hydrogen one thus obtains

$$\tau_{\text{GP,HI}}(z) = 3.99 \times 10^5 \left(\frac{1+z}{7} \right)^{3/2} x_{\text{HI}} (1 + \delta_{\text{H}}) \quad (1.3)$$

(Shull et al. 2010), where x_{HI} is the neutral hydrogen fraction and δ_{H} denotes the hydrogen density normalised to the cosmic mean hydrogen density. Assuming that the Ly α absorption seen in QSO spectra becomes saturated for optical depths $\tau \gg 10$, a region with a low hydrogen density of $1 + \delta_{\text{H}} = 0.1$ would already show saturated Ly α absorption at $z = 6$ with a small neutral fraction of $x_{\text{HI}} \approx 10^{-4}$. Hydrogen Ly α absorption is thus sensitive to the slightest traces of neutral hydrogen. Larger neutral fractions can be probed however with the Ly β and Ly γ transition. Due to their smaller oscillator strengths and wavelengths, the Ly β transition is about 6 times less sensitive to neutral hydrogen, and Ly γ almost 18 times less sensitive, hence providing better constraints on the evolution of neutral hydrogen at the end of reionisation.

From QSO spectra, Songaila (2004) observed an increase in the Ly α optical depth and with this an increase in neutral hydrogen at $z > 5$. They further find that the

³<http://www.lofar.org>

⁴<http://www.mwatelescope.org>

⁵<http://www.skatelescope.org>

transmission flux of their QSO sample approaches zero flux for $z > 5.5$. Spectra of SDSS quasars reveal large saturated regions blueward of the QSO's Ly α emission line at $z \approx 6$ (for a compilation of $z \sim 6$ QSO spectra, see Fan et al. (2006a)). For $z > 6$ the volume averaged hydrogen neutral fraction is found to rapidly increase to $x_{\text{H I}} = 10^{-3.5}$ or larger (Lidz et al. 2002; Cen & McDonald 2002; Fan et al. 2002, 2006b). All these observations provide indications that the era of reionisation ended at around $z \sim 6$

1.3. The evolution of the ionised IGM after reionisation

After reionisation ended and most of the matter in the Universe was ionised, the Universe was kept in this state by the integrated UV field produced by all the galaxies and QSOs (Madau 1995; Haardt & Madau 1996; Madau et al. 1999; Bianchi et al. 2001; Faucher-Giguère et al. 2009). This UV background (UVB) is a combination of photons originating in QSOs, which are the dominant contributors at $z \sim 3$, and star forming galaxies. Due to the high ionisation state of the IGM, the mean free paths for UV photons are extremely large and UV photons can pass through the IGM almost unobstructed (e.g. Miralda-Escudé et al. 2000). At $z = 6$ the UV photon mean free path in the IGM lies around proper $10 - 15 h^{-1}$ Mpc (Gnedin 2000; Meiksin & White 2004). At lower redshifts, the IGM become more and more ionised and the mean free path rises to about proper $60 - 90 h^{-1}$ Mpc (Meiksin & White 2004; Faucher-Giguère et al. 2008a).

The UV background was not uniform in space, but fluctuated in intensity depending on the environment (Maselli & Ferrara 2005; Furlanetto 2009; Mesinger & Furlanetto 2009). Just after reionisation ended, the fluctuations were strong and became less pronounced at lower redshifts. The UV background flux is higher in overdense locations, where many UV producing sources are clustered together. In underdense regions, where the density of UV sources is low, the UV background flux is lower than in the mean. The strength of these fluctuations is dependent on the mean free path. When the mean free path is short, the fluctuations are enhanced due to the absorption of the UV radiation stemming from the sources. If the mean free path is large, photons travel larger distances and smooth the fluctuations out. As a result the background becomes more homogeneous. Right after reionisation the UVB is found to have fluctuated up to a factor of three around the mean flux (Mesinger & Furlanetto 2009). Due to the ever increasing ionisation state of the IGM, fluctuations in the photoionisation rate are found to have decreased down to 30% for H I and 60% for He II at redshifts around $z \sim 3$ (Maselli & Ferrara 2005). The fluctuations in the He II photoionisation rate are due to the fact, that at $z = 3$ helium was not yet as highly (doubly) ionised as hydrogen. Helium was thus not yet completely reionised at the redshift of hydrogen reionisation but only became highly ionised at later times.

Not only was hydrogen ionised during reionisation, also the 25% helium were ionised to become singly ionised helium (He II). The He II regions were comparable in size to the H II regions, since the He I photoionisation rate was similar to the one of H I. However the photoionisation cross section for He II is much smaller than for He I. Therefore, only a small fraction of helium was doubly ionised at the end of hydrogen reionisation. In order to ionise He II efficiently, hard UV photons are needed. Pop III stars may not have produced enough hard UV photons to have effectively ionised He II (Venkatesan et al.

1. Introduction

2003). It is still debated though, whether helium was reionised by hard energetic photons stemming from QSOs, or whether enough soft photons with energies around the Lyman edge were available from star forming galaxies to complete He II reionisation (e.g. Bolton et al. 2009). However He III recombines 5 – 6 times faster than H II (e.g. Osterbrock & Ferland 2006), which means that a large amount of hard UV photons are needed to keep helium doubly ionised, favouring AGNs as the dominant source of He II reionisation. If QSOs were the main driver for He II reionisation, the maximum of the QSO space density at $z \sim 3$ (e.g. Warren et al. 1994; Wolf et al. 2003; Croom et al. 2009) implies that He II was efficiently ionised at around that redshift. Thus He II reionisation may have ended then. Observationally the ionisation history is accessible through He II absorption spectra. Similar to the hydrogen case, the He II optical depth can be determined through the amount of saturated absorption in the He II forest (e.g. Reimers et al. 1997; Shull et al. 2004). He II absorption however probes 50 – 100 times higher ion abundance fractions than H I. This means that only an He II fraction of $x_{\text{He II}} \sim 10^{-2}$ is needed at $z = 3$ for an absorber with a density of $1 + \delta = 0.1$ to saturate He II Ly α absorption (Shull et al. 2010). Observational and theoretical constraints indicate that He II reionisation must have indeed ended between $2.7 \lesssim z \lesssim 4$ (Miralda-Escudé et al. 2000; McQuinn et al. 2009; Furlanetto & Dixon 2010; Becker et al. 2011; Syphers et al. 2011).

The history of neutral hydrogen becomes observationally accessible through high redshift QSO spectra. High resolution spectra of high redshift QSOs show a wealth of absorption lines blueward of the QSO’s Ly α emission line. This forest of H I absorption lines is known as the Ly α forest (Rauch 1998) and is produced by neutral hydrogen located between the observer and the QSO. As the light emitted blueward of the QSO’s Ly α emission line travels towards the observer, it is redshifted. Whenever there is intervening neutral hydrogen and photons are redshifted towards the wavelength of the Ly α resonance line, photons are scattered out of the line of sight (e.g. Lynds 1971; Sargent et al. 1980). This thus gives rise to a Ly α absorption line at the corresponding redshift of the intervening H I gas. It has been shown, that the H I absorption in the Ly α forest traces the cosmic web and that the intervening absorption systems correspond to filaments and sheets (e.g. Petitjean et al. 1995; Bi & Davidsen 1997; Gnedin 1998b; Davé et al. 1999). This fact has been established by generating H I Ly α forest spectra from simulated matter density distributions and comparing statistical properties of simulated spectra, such as the absorber density evolution or the distribution function of the absorber line widths, with the observed ones (e.g. Miralda-Escudé et al. 1996; Hernquist et al. 1996). For generating Ly α spectra however, the ionisation state of the IGM needs to be known, which is determined by the UV background photoionisation rate (e.g. Hui et al. 1997). Good measurements of the UV background are thus needed in order to compare the simulations with observations (e.g. Bolton et al. 2005; Faucher-Giguère et al. 2008a; Dall’Aglio et al. 2008a; Calverley et al. 2010). Nevertheless, the Ly α forest is a powerful tool to probe and constrain the formation of structure in the Universe. For instance it is possible to derive constraints for the baryonic matter density $\Omega_b h^2$ (Rauch et al. 1997b) from the Ly α forest or the equation of state of dark energy (Kujat et al. 2002; Viel et al. 2003).

The shape of absorbers in the Ly α forest resemble Voigt profiles. The Voigt profile describes the line profile of an absorber which is Doppler and Lorentzian broadened (e.g. Unsöld 1955; Tepper-García 2006). The Voigt profile is characterised by the column

density N of the absorber and the b -parameter. The b -parameter is a measure of the velocity dispersion in the absorber and parameterises the broadening of the line. This leads to the observationally motivated categorisation of Ly α forest absorbers into three categories. H I absorbers with a column density of $12 < \log N_{\text{HI}} < 17.2 \text{ cm}^{-2}$ are considered to be Ly α forest absorbers. Systems with column densities of $17.2 < \log N_{\text{HI}} < 20.2 \text{ cm}^{-2}$ are optically thick to ionising radiation and are known as Lyman limit systems. For higher column density systems which are mostly neutral, Lorentzian damping wings appear in the absorption line profile and are thus generally referred to as damped Ly α systems (e.g. Rauch 1998). Damped Ly α systems have been shown to probe protogalactic clumps (Haehnelt et al. 1998) and provide constraints on the formation and evolution of galaxies (e.g. Wolfe et al. 2005).

Since Ly α forest lines are thought to trace the cosmic web, it is common to derive statistical properties from the observed Ly α forest and compare them with cosmological simulations of structure formation (e.g. Hui et al. 1997). Two distinct methods are used to analyse the Ly α forest: Voigt profile fitting of Ly α absorption lines (e.g. Kim et al. 2001) and pixel statistics (e.g. Schaye et al. 2003). In a Voigt profile analysis, the Ly α absorbers are decomposed into multiple Voigt profiles until the model spectrum matches the observed one. Due to line blending, the result is degenerate and many different combinations of Voigt profiles may lead to the same result. The Voigt profile fitting analysis however has the advantage that physical quantities of the absorber are measured, such as the redshift, column density and the velocity dispersion. For a Voigt profile fitting analysis, high resolution spectra are needed. The problem of degeneracy does not exist in the pixel statistics method. There, each pixel in the spectra is treated on its own which makes the method ideal for low resolution spectra. With a Voigt profile fitting analysis, one can, for instance, infer the absorber number density evolution (e.g. Kim et al. 2001) or, when metal lines are present, one can calculate metallicities in the IGM (e.g. Ellison et al. 2000; Songaila 2001). Pixel statistics method on the other hand provide for instance the evolution of the mean optical depth or the power spectrum (e.g. Gnedin & Hamilton 2002).

A small fraction of absorption lines in the Ly α forest is not produced by intervening H I but belongs to highly ionised metal ions such as triply ionised carbon C IV, triply ionised silicon Si IV, or five times ionised oxygen O VI (e.g. Norris et al. 1983; Lu 1991; Songaila & Cowie 1996; Adelberger et al. 2003). Roughly half of all H I absorber with a column density of $N_{\text{HI}} = 3 \times 10^{14} \text{ cm}^{-2}$ show associated C IV with a metallicity of $\sim 10^{-2} Z_{\odot}$ at $z = 3$ (Cowie et al. 1995; Tytler et al. 1995; Songaila & Cowie 1996).

How the IGM was metal enriched is still under debate. It may be that the IGM was already metal enriched early on in the history of the Universe by the first stars. Metals that formed in these first objects can be distributed over large cosmological volumes by supernovae (e.g. Madau et al. 2001) thus enriching large volumes of the IGM. An indication for this would be, if H I absorbers of all column densities show associated metal absorption lines. In this picture, different metal absorber probe different density regimes, depending on their ionisation state (Rauch et al. 1997a). Another process to enrich the IGM suggests that metals which formed in galaxies are transported out into the IGM. This can be achieved either with strong galactic winds (e.g. Aguirre et al. 2001b; Oppenheimer & Davé 2006) or through ejection of gas in galaxy mergers (Gnedin 1998a). Any enrichment process that would deposit metals only in the filamentary

1. Introduction

structure of the cosmic web and leave the remaining volume of the Universe pristine, should manifest itself in a drop of metallicity at the characteristic H I column densities of filaments, $\log N_{\text{HI}} \sim 14$ (Rauch 1998). Observationally such a drop has not yet been established. Thus, the exact nature and the evolution of the enrichment process can only be determined with good observational data giving strong constraints on the evolution of the metal absorbers. Recent observations of the C IV mass density showing an increase in the C IV mass density with decreasing redshift suggest, that current galactic feedback models are still missing a crucial component, in order to describe the observed evolution (D’Odorico et al. 2010). We will look into the metal enriched IGM in more detail in chapters 2 and 3.

1.4. Thesis outline

In this thesis I study two different properties of the Ly α forest using observations and simulations. Using observed spectra of high-redshift QSOs I characterise the C IV enriched Ly α forest. This helps to answer the question of how much of the IGM is metal enriched and what the properties of C IV enriched H I absorbers are. Furthermore, using radiative transfer simulations, the QSO line of sight proximity effect which is used to derive the UVB photoionisation rate is studied. It is assessed how these measurements are influenced by various physical effects, such as radiative transfer effects or the influence of large scale density inhomogeneities.

In chapter 2 the volume averaged H I and C IV column densities for C IV absorption systems is determined, using 17 high-redshift and high resolution QSO spectra. Observations of the C IV optical depth as a function of H I optical depth using pixel statistics indicate a linear relation between the two quantities (e.g. Schaye et al. 2003). I study whether this holds as well when relating the H I system column density with the C IV column density. Numerical simulations of the IGM metal enrichment however predict a drop in volume averaged metallicity at characteristic H I densities (e.g. Aguirre et al. 2001b; Springel & Hernquist 2003). They further predict that no strong correlation between the metallicity and the H I density exists for densities above the characteristic drop-off density. With the observations this behaviour is assessed, which helps to constrain future numerical simulations of IGM metal enrichment (e.g. Tescari et al. 2010) and may help to improve galaxy feedback models.

In chapter 3 measurements of the Ly α forest absorber number density evolution and the differential column density distribution are updated, again using the high-redshift and high resolution QSO spectra already used in chapter 2. Especially the depression in the differential column density distribution (see Petitjean et al. 1993) is studied. The sample size is large enough, to derive the combined number density evolution and study the evolution of the differential column density distribution in three redshift bins. From the combined sample a dip in the number density evolution of high column density systems is seen at $z \sim 2$. This dip is also connected to the depression in the differential column density distribution above $\log N_{\text{HI}} \sim 14$. Further, the absorber sample is divided into C IV enriched and unenriched absorbers, and the number density evolution and the column density distribution function are derived for each sample individually. It is found, that the number of highly ionised C IV systems (strong C IV systems which are associated

to low column density H I systems) increases with decreasing redshift. Further it is found, that C IV enriched H I systems show a different slope in the column density distribution, than the unenriched ones. Thus distinct characteristics of C IV enriched and unenriched systems are obtained.

In chapter 4 I develop a numerical scheme to parallelise the cosmological radiative transfer code **CRASH2** (Maselli et al. 2003, 2009). This enables the simulation of complex high-resolution radiative transfer problems, such as reionisation or the physical characteristics of the UV background flux. The **CRASH2** Monte-Carlo ray tracing scheme is adopted to run on distributed memory computer clusters (a complete description of the **CRASH** code is given in chapter A). The new parallel code, **pCRASH2**, is tested thoroughly and its performance is evaluated. The new parallel version is shown to perform well and is perfectly suited to study complex problems with large box sizes, which require photons produced by many sources to be followed with high precision.

In chapter 5 I model the QSO line of sight proximity effect using cosmological simulations in connection with detailed 3D radiative transfer simulations. The proximity effect is used to measure the UV background flux at intermediate redshifts of $2 < z < 6$ (e.g. Dall’Aglio et al. 2008a). However results obtained with the proximity effect method differ systematically from results obtained with the independent flux transmission statistics method (e.g. Bolton et al. 2005; Faucher-Giguère et al. 2008a). The major assumptions in the widely used analytical formulation of the proximity effect (Bajtlik et al. 1988; Liske & Williger 2001) are tested and confirmed with radiative transfer simulations. Further the influence of the local QSO host environment on the proximity effect signal is explored. For this results for a QSO contained in the most massive halo, in a random filament, and in a random void are obtained. Additionally the influence of the QSO’s luminosity on the proximity effect signal is studied. Indications are found, that the results are biased by large scale overdensities, which have been shown by Faucher-Giguère et al. (2008b) to influence the determination of the UV background photoionisation rate. This effect is discussed in more details in the chapter 6.

In chapter 6 I revisit the QSO line of sight proximity effect and study the influence of large scale density structures on the proximity effect signal. Using a high resolution cosmological simulation, the dependency of the proximity effect strength parameter on the host halo mass is explored in three halo mass bins. Further the dependency of the strength parameter on the mean density around the host in a sphere of $10 h^{-1}$ Mpc comoving radius is derived. A large scatter between the various halos is observed. However a dependence of the strength parameter on the halo mass cannot be confirmed. Nevertheless the proximity effect strength parameter is shown to be clearly correlated with the mean density around the host. Large scale overdensities are found to reduce the proximity effect’s strength, while large scale underdense regions increase its strength.

I summarise and conclude the results obtained in this thesis in chapter 7.

Triply ionised carbon associated with the Ly α forest at $2 < z < 3.5$ ¹

Numerical simulations of the chemical enrichment in the intergalactic medium (IGM) predicted the volume averaged relation between the gas density and its metallicity. In this chapter, we want to establish a similar relation from Ly α forest observations. In observations the gas density can be inferred from the column density of Ly α forest hydrogen absorbers $N_{\text{H I}}$. Further the metallicity of the IGM is traced by the column density of forest metal absorbers, such as triply-ionised carbon $N_{\text{C IV}}$. Using a sample of 17 high-redshift and high resolution QSO spectra taken from the ESO VLT UVES archive, a good coverage of the Ly α forest between $2 < z < 3.5$ is obtained. The Ly α forest has been analysed using a Voigt profile analysis. For each line of sight, all the available H I Lyman lines and metal lines were fitted with Voigt profiles and a robust estimate of the absorber's column densities is obtained. The volume averaged $N_{\text{H I}}$ and $N_{\text{C IV}}$ are determined for each C IV system, where a system is defined as a $\pm 250 \text{ km s}^{-1}$ velocity interval centred on the C IV flux minimum. From the Voigt profile analysis we obtain the $N_{\text{H I}} - N_{\text{C IV}}$ relation, which is similar to the numerically predicted density-metallicity relation. The $N_{\text{H I}} - N_{\text{C IV}}$ relation is characterised by a sharp drop in C IV column density at $N_{\text{H I, drop}} \sim 10^{15.2} \text{ cm}^{-2}$. Further, larger H I column densities $N_{\text{C IV}}$ only weakly relate to $N_{\text{C IV}}$. The shape of the observed $N_{\text{H I}} - N_{\text{C IV}}$ relation suggests that the forest may not be metal-enriched below a hydrogen column density of $N_{\text{H I}} \leq 10^{14} \text{ cm}^{-2}$. We present a simplistic shell model composed of an H I halo following a NFW profile and an embedded smaller C IV halo with an exponential density profile which is able to qualitatively explain the observed $N_{\text{H I}} - N_{\text{C IV}}$ relation. Our simple model suggests that C IV absorbers which are associated with the Ly α forest might have their origin in a metal enriched circum-galactic medium embedded in galactic halos.

2.1. Introduction

In spectra of high redshift QSOs, numerous narrow absorption lines are present blueward of the Ly α emission line. Most of these absorbers are produced by Ly α absorption of neutral hydrogen (H I) intersecting the QSO sight line. The so called Ly α forest arises from the remaining neutral hydrogen in the warm ($\sim 10^4 \text{ K}$) photoionised intergalactic medium and traces the underlying dark matter distribution (Cen et al. 1994; Petitjean et al. 1995; Bi & Davidsen 1997).

The discovery of triply ionised carbon C IV doublets ($\lambda\lambda 1548, 1550 \text{ \AA}$) which are associated with absorbers in the Ly α forest at $z \sim 3$ showed, that the Ly α forest is chemically

¹This part contains work I was involved in with T.-S. Kim, R.F. Carswell, and J.P. Mücke. I will present the parts I have contributed to a paper which will be submitted to MNRAS. Sections 2.2.1 (except Sect. 2.2.2) and Section 2.5.1 describe work carried out by T.-S. Kim.

2. The forest CIV at $2 < z < 3.5$

enriched up to $10^{-2.5} Z_{\odot}$ for H I absorbers with column densities of $N_{\text{HI}} \sim 10^{14.5} \text{ cm}^{-2}$ (Cowie et al. 1995; Tytler et al. 1995; Songaila 1998) or lower (Ellison et al. 2000; Schaye et al. 2003, 2007). However the forest is not only enriched with carbon, but with other ions, such as silicon and oxygen (Songaila & Cowie 1996; Davé et al. 1998; Schaye et al. 2000a). Since the IGM itself cannot form stars due to its high temperature and low density, all these metals may have been produced by stars in galaxies at $z > 6$ which were then transported into the IGM (Madau et al. 2001). The metals in the Ly α forest thus help to understand the formation of galaxies in the universe and the feedback between high-redshift galaxies and the neighbouring IGM (Davé et al. 1998; Aguirre et al. 2001b; Schaye et al. 2003; Oppenheimer & Davé 2006).

Various models describing the enrichment process of the IGM have been proposed. The metals might originate in Population III stars at $10 < z < 20$ (Ostriker & Gnedin 1996; Haiman & Loeb 1997) or they might be dynamically stripped from galaxies by mergers or tidal interaction (Gnedin & Ostriker 1997; Gnedin 1998a). Furthermore models of dust ejection by stellar radiation pressure (Aguirre et al. 2001a) or the accretion of pre-enriched gas from the very first stars in the universe have been proposed (Haehnelt et al. 1996b; Rauch et al. 1997a; Madau et al. 2001). Widely used scenarios for enriching the IGM in simulations are galactic winds blown by star forming galaxies. By observing local starburst galaxies, outflows operating on scales of a few kpc (Heckman et al. 2000; Strickland et al. 2004; Martin 2005, 2006) have been found. At higher redshifts, outflows with velocities of hundreds of km s^{-1} have been detected in Lyman break galaxies (Pettini et al. 1998; Shapley et al. 2003), supporting the model of IGM enrichment through galactic winds. Furthermore this model is supported by observations showing that most CIV absorbers are related to galaxies (Adelberger et al. 2005).

Although there is no consensus on the details of the outflow mechanism yet, numerical simulations agree in general on the evolution of the chemical enrichment, such as the volume averaged density-metallicity relation. Using simulations, the density-metallicity relation is predicted to drop off sharply at overdensities lower than a characteristic overdensity. Above this threshold, however, the IGM metallicity shows only a weak dependency on the IGM density (Aguirre et al. 2001b; Springel & Hernquist 2003; Oppenheimer & Davé 2006). The exact shape of the relation is strongly dependent on the outflow velocity of the galactic winds. With stronger outflows, larger distances from the galaxies can be enriched with metals, and the drop-off in the density-metallicity relation will shift towards lower densities. Also the time-scales on which winds operate affects the relation similar to the outflow speed.

On the other hand, a similar relation between the metallicity and the IGM density has been found by Gnedin & Ostriker (1997) and Gnedin (1998a). However unlike above, the driving mechanisms for the enrichment of the IGM are not outflows, but rather the transport of metals through mergers and tidal interactions.

Observationally, the density and the metallicity of the IGM cannot be directly measured. The only quantities that can be obtained from QSO spectra are the optical depth τ or the column density of H I or other ions. Given a cosmological model, the over-density can be related to the H I column density and optical depth τ_{HI} (Schaye 2001). However the metallicity is not as easily inferred. By assuming photoionisation equilibrium, it is possible to calculate the metallicity for a given UV background (Hui & Gnedin 1997; Davé et al. 1999; Schaye 2001). However, whether photoionisation equilibrium really

holds for these metal ions, is still debated (Haehnelt et al. 1996a, Kim & Carswell 2011 in preparation).

Up to today, the predicted density-metallicity relation has not been detected in observations yet. In previous studies using Voigt profile analysis of the Ly α absorbers, this is mainly due to the small number of data points available in previous measurements of the $N_{\text{HI}} - N_{\text{CIV}}$ relation (e.g. Simcoe et al. (2004)). In studies which make use of the optical depth analysis, where spectral features are correlated on a pixel by pixel basis, a linear relation between the C IV optical depth τ_{CIV} and τ_{HI} has been revealed up to the confusion limit (Ellison et al. 2000; Schaye et al. 2003). However the $\tau_{\text{CIV}} - \tau_{\text{HI}}$ relation does not follow the predicted density-metallicity relation.

In this chapter, we present observational variant of the theoretically predicted density-metallicity relation at redshifts $\langle z \rangle = 2.22$ and $\langle z \rangle = 2.76$. Using 17 high-resolution and high signal-to-noise (S/N) spectra obtained from the ESO VLT UVES archive, we derive the $N_{\text{HI}} - N_{\text{CIV}}$ relation, covering the Ly α forest at $2 < z < 3.5$.

In the next Sect. 2.2, we will discuss the observational data obtained from the ESO archive. In Sect. 2.2.1 we will summarise the methodology used for fitting Voigt-profiles to the data set. We will also describe the working definition of an absorption *system* used in this work. In Sect. 2.3 our results for the $N_{\text{HI}} - N_{\text{CIV}}$ relation will be discussed. In Sect. 2.4 we will compare our QSO sample with previous results and in Sect. 2.5 a simple model that qualitatively reproduces the data is presented. We will summarise the results in Sect. 2.6.

Throughout this chapter, we assume the cosmological parameters to be $\Omega_{\text{M}} = 0.24$, $\Omega_{\Lambda} = 0.73$, and $h = 0.71$ in accordance with the 5-year WMAP results (Hinshaw et al. 2009).

2.2. Data and Voigt profile fitting analysis

The same 17 VLT/UVES spectra of high-redshift QSOs taken from the UVES archive have been used as by Kim et al. (2007). The data reduction is described in Kim et al. (2004) and Kim et al. (2007). These high-resolution spectra are characterised by a resolution of $R \sim 45\,000$ and the individual properties of these 17 QSOs are given in Table 2.1. The instrument wavelength gaps and regions with strong telluric contamination are given in the 7th and 8th column of Table 2.1. Only those regions affected by telluric lines are given, that overlap with regions containing C IV absorbers. Further note, that the sample was selected in such a way that it does not contain any damped Ly α systems (DLAs, $N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$). However they do include sub-damped Ly α systems (sub-DLAs, $N_{\text{HI}} = 10^{19.0-20.3} \text{ cm}^{-2}$).

The proximity effect region was avoided by excluding a region of $4\,000 \text{ km s}^{-1}$ blueward of the QSO's Ly α emission line. With this wavelength cut, we obtain the highest possible redshift range for reliably finding C IV absorbers, as given in Table 2.1 3rd column. The lowest redshift where C IV is not blended with HI forest absorbers is given by the QSO's Ly α emission line. For smaller redshifts, our sample is most likely incomplete, due to line blending. We have restricted the lowest redshift for C IV absorbers to $z = 2$, since robust N_{HI} measurements of saturated HI absorption lines can only be obtained for $z \geq 2$. This is due to the fact, that up to $z = 2$, HI absorbers can be better constrained with higher

Table 2.1.: Analysed QSOs

QSO	z_{em}^{a}	$z_{\text{CIV}}^{\text{b}}$	$\lambda_{\text{CIV},1548}(\text{\AA})$	$z_{\text{Ly}\beta}^{\text{b}}$	$z_{\text{SiIV,min}}$	Gaps (\AA) ^c	Telluric regions (\AA) ^d	CIV region rms ^e
Q0055-269	3.655	2.662-3.390	5670-6797	2.130	3.041		6276-6315	0.020/0.031
PKS2126-158	3.279	2.360-3.205	5201-6510	2.396	2.718	5595-5681	6275-6327/6472-6504	0.006/0.008
Q0420-388	3.116	2.240-3.038	5016-6251	2.667	2.575	5596-5682		0.009/0.010
HE0940-1050	3.088	2.212-3.006	4973-6202	2.464	2.545	5762-5847	5873-5955	0.008/0.014
HE2347-4342	2.764	2.067-2.819	4748-5913	2.071	2.385	5744-5837		0.005/0.011
Q0002-422	2.767	2.000-2.705	4644-5736	2.022	2.256			0.006/0.009
PKS0329-255	2.704	2.000-2.651	4644-5653	2.090	2.219			0.016/0.025
Q0453-423	2.658	2.000-2.588	4644-5555	1.978	2.172			0.008/0.016
HE1347-2457	2.616	2.000-2.553	4644-5500	1.986	2.112			0.007/0.017
Q0329-385	2.434	2.000-2.377	4644-5228	1.998	1.973			0.018/0.030
HE2217-2818	2.413	2.000-2.365	4644-5210	1.977	1.964			0.007/0.010
Q0109-3518	2.405	2.000-2.348	4644-5183	1.977	1.949			0.008/0.010
HE1122-1648	2.404	2.000-2.358	4644-5199	1.975	1.965			0.005/0.014
J2233-606	2.250	2.000-2.201	4644-4955	2.000	1.845			0.025
PKS0237-23	2.223	2.000-2.179	4644-4922	1.974	1.806			0.010/0.014
PKS1448-232	2.219	2.000-2.175	4644-4916	1.975	1.789			0.010/0.020
Q0122-380	2.193	2.000-2.141	4644-4864	1.976	1.772			0.022

^a The redshift is measured from the observed Ly α emission line of the QSOs.

^b This column indicates the redshift above which the observed spectra cover both Ly α and Ly β regions.

^c Only the QSOs whose CIV region is affected by the instrumental setups are listed.

^d Only the QSOs whose CIV region is strongly affected by the telluric lines are listed. These regions are discarded in the present work. The regions with a low-level telluric contamination are not listed. They are included to search for stronger CIV systems.

^e Since the S/N ratio varies along each spectrum, the 1σ r.m.s. values of the normalised flux in the CIV continuum regions also vary. The smaller (larger) number corresponds roughly to the highest (smallest) S/N ratio. When the S/N ratio is roughly constant, only one number is listed.

order lines, such as Ly β and Ly γ . However at lower redshifts, the necessary wavelength ranges for these higher transition lines are not covered anymore.

2.2.1. Voigt profile fitting analysis

An extensive Voigt profile fitting analysis has been performed on the 17 QSO spectra. For each absorber, its redshift z , its column density N , and its Doppler parameter b has been determined using VPFIT². We will outline the basic procedure how the fitting analysis has been carried out. For a detailed discussion see Carswell et al. (2002) and Kim et al. (2007).

As a first step, the continuum of the QSO spectra was determined. Then, metal lines were searched for starting from longest wavelengths (high redshift) to shorter ones. Any identified metal line was then fitted with a Voigt profile, using same z and b values for equal species transition. Whenever a metal line was embedded in the H I forest, the H I absorbers were included in the fitting procedure. By fitting different transitions at the same time, corrections could be applied to the continuum determination, to obtain acceptable ion ratios. After all identifiable metal lines have been fitted, the remaining absorption features were fitted. The remaining systems were all assumed to represent H I absorbers. Any available higher-order Lyman series line was additionally considered when fitting the H I Ly α absorber. As with the metals, some small adjustments to the continuum fit were needed in order to retain the intensity ratios between the different transitions. After this first iteration and an improved determination of the continuum fit, the whole spectrum was refitted from scratch again using the new adjusted continuum. This iterative procedure was repeated several times until a satisfactory model for the observed spectra was obtained.

Robust H I column density measurement

Since we are interested in the column densities of H I absorbers that have associated C IV, robust measurements of both N_{HI} and N_{CIV} are needed. Unfortunately, most H I absorbers with associated C IV are saturated (C IV is usually not saturated). In these cases there is no unique method to determine the number of components from the Ly α transition alone. Assuming a single component for the saturated absorber only provides a lower limit on the column density. However when high-order Lyman series are included in the fitting procedure, better results can be obtained, since many saturated H I Ly α lines become unsaturated in higher orders due to their smaller oscillator strengths. It is even possible that a systems which is in Ly α apparently just one component, breaks up into several weaker components in the higher orders.

In Fig. 2.1 such a behaviour is shown in a normalised flux-velocity diagram for two absorption systems. The left panel shows a system at $z = 3.217$ towards PKS2126-158. The right panel shows another example at $z = 2.444$ towards Q0453-423. The system in the left panel gives an example for a saturated Ly α line becoming unsaturated in Ly δ . Including the higher orders for fitting this system reveals a total H I column density of $\log N_{\text{HI}} = 15.78 \pm 0.04 \text{ cm}^{-2}$ with 7 components. If 7 components are just fitted to the Ly α spectrum without the additional constraints by the higher order transitions, a total

²Carswell et al.: <http://www.ast.cam.ac.uk/~rfc/vpfit.html>

2. The forest CIV at $2 < z < 3.5$

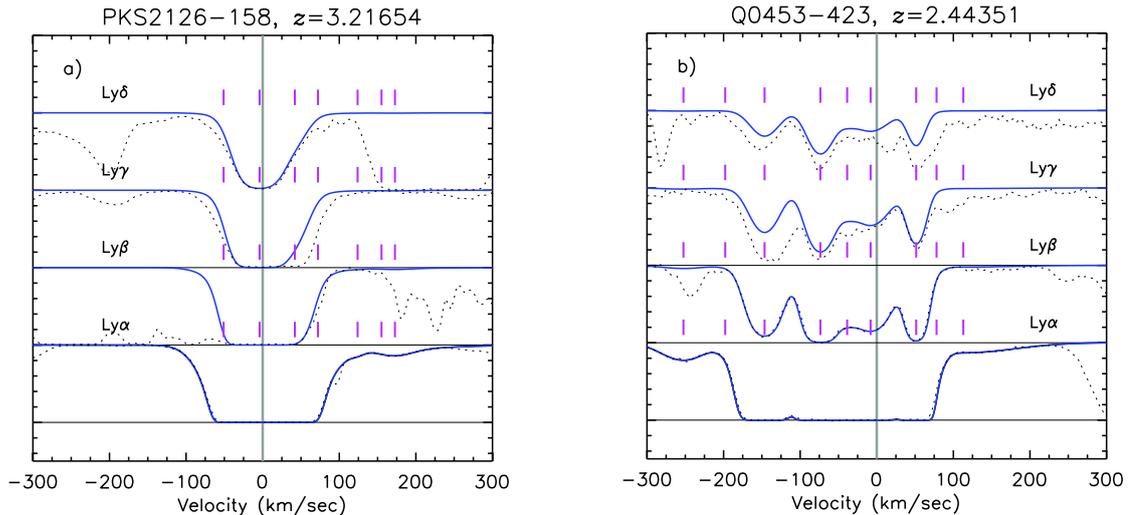


Figure 2.1.: Velocity plot of two systems at $z = 3.217$ towards PKS2126-158 (panel a), and $z = 2.444$ towards Q0453-423 (panel b). The systems are centred on the strongest CIV flux value. The normalised fluxes of HI transitions up to Ly δ are shown, where the thick blue line is the model spectra and the black dotted line the observed normalised flux. The tick marks denote the velocity of each fitted component. Panel a) shows an example of a saturated Ly α line becoming unsaturated at higher transitions. Panel b) illustrates a saturated Ly α line that breaks up into many smaller components in Ly β .

column density of $\log N_{\text{HI}} = 14.94 \pm 0.06 \text{ cm}^{-2}$ is obtained. In the right panel of Fig. 2.1 an example of a saturated Ly α line breaking up into many weaker components is shown, as is clearly seen in Ly β . About 24% of saturated Ly α absorption lines in our sample break up into several weaker higher order components. About 18% of the saturated Ly α lines already break up into several components in Ly β .

Fairly robust N_{HI} within the 1σ fitting errors of the various components can be obtained for $\log N_{\text{HI}} \leq 17.0$ when Ly β and Ly γ lines are included. Using only Ly β additionally to Ly α for a single component system, its fit deviates at most 2σ from results which include Ly γ as well. However fitting errors may increase if a saturated system consists of several components.

For all ground-based optical spectra, a natural redshift limit below which a robust N_{HI} cannot be obtained using higher transitions is imposed by the atmospheric cutoff at 3050\AA . The UVES spectra therefore cover Ly β only for redshifts $z > 1.98$. Therefore, saturated lines at $z < 1.98$ cannot be robustly modelled. We therefore limit our sample only to redshifts larger than $z \geq 2.0$.

for each fitted system we have determined the robustness of the fitted model. We have categorised and rated robustness using the following criteria:

- Class 1: N_{HI} is considered robust, if Ly α or one of the higher order lines is not saturated.
- Class 2: N_{HI} is considered almost robust, if all available Lyman lines higher than Ly γ are saturated or Ly γ corresponds to the highest available order.

- Class 3: N_{HI} is considered reasonable robust, if only Ly α and Ly β are available.
- Class 4: N_{HI} is considered unconstrained, if only a saturated Ly α line is available or if higher order lines are blended with lower- z forest lines.
- Class 5: N_{HI} is considered reasonable robust, if N_{HI} can be constrained using other methods, such as the presence of weak Ly α damping wings of high N_{HI} -end Lyman limit systems.

Systems belonging to class 2 are either Lyman limit systems which are still saturated down to Ly-10, or systems with limited coverage of higher order lines due to an intervening Lyman edge or the atmospheric cutoff. The problem of intervening Lyman edges or the atmospheric cutoff also applies for class 3, mostly for systems at $1.986 < z < 2.146$.

Robust detection of C IV doublets

For C IV, a strong and a weak detection criterion has been applied. A C IV doublet is considered a strong detection, when the strong doublet line C IV $\lambda 1548$ is stronger than 3σ of the spectrum's variance. This implies that the weaker doublet component is detected at $\geq 1.5\sigma$. However whenever the doublet ratio of a C IV component detected at $\geq 3\sigma$ is not roughly 2:1, the line is not considered as a robust detection.

The detection limit depends on the line's Doppler parameter b as well as the S/N ratio. If the Doppler parameter is larger than the instrument resolution, then, for a given N_{CIV} , narrower lines are more easily detected than broader ones. We therefore do not define a minimum N_{CIV} detection limit, but rather list the 1σ r.m.s. values of the normalised flux in the C IV continuum regions in the last column of Table 2.1. Overall (excluding several low S/N spectra) the 3σ detection limit lies roughly at $\log N_{\text{CIV}} \sim 12 \text{ cm}^{-2}$.

It may happen, that a C IV line can be reliably classified as C IV below our strong detection limit, if either the lines are narrow or because additional ions are present at the same redshift. In these cases, the detection limit is relaxed to $2 - 3\sigma$ and we call this a weak detection. The reason for relaxing the detection criterion is to increase the sample of robust C IV detections. Especially at low column densities the most significant redshift evolution of the $N_{\text{HI}} - N_{\text{CIV}}$ relation is expected. It is therefore important to include as many reliable systems as possible.

It might seem, that applying two different criteria results in a highly inhomogeneous C IV sample. By volume averaging over multiple components to obtain a total N_{CIV} of an absorption system of any definition, including the weak criterion C IV components only increases the total column density. This would correspond to the highest possible total N_{CIV} for a system, while not including these weakly constrained components results in a lower limit. Comparing results according to these two criteria shows that only systems with total column density of $N_{\text{CIV}} \leq 12.8 \text{ cm}^{-2}$ are affected by inclusion or non-inclusion of the weaker constrained lines. Of all 127 identified systems (for our definition of a system see Sect. 2.2.2) only 8 systems differ by more than 10% when including weaker constrained absorbers than without them. Also, one has to keep in mind, that in reality, many of the $\log N_{\text{CIV}} \leq 12 \text{ cm}^{-2}$ lines could be spurious and are just noise features. However, the results of this work will not change regardless whether some weak and spurious C IV lines are included or not.

2.2.2. System

In physics, a system is considered a set of interacting or interdependent entities which stand in relationship with each other. Based on this, CIV and HI can be considered a physical system if they are physically dependent or if any process in one of them triggers a reaction in the other. However, even if two CIV and HI absorbers are very close in redshift space, they can still be very far away in real space due to bulk motions and peculiar velocities in the IGM (Rauch et al. 1997a). Therefore, from spectroscopic observations alone, we are unable to apply the physical definition of a system to the absorption lines.

Conventionally, a forest system is considered an apparently isolated absorption feature which can consist of multiple components. However, this definition poses a significant problem in how to associate CIV components to physically related HI components. We therefore adopt a slightly different approach and define a CIV system as a group of individual CIV components lying within a given velocity interval. This definition is extended to any CIV complex spread over a greater velocity range when the separation to the neighbours is less than this velocity interval. This extended velocity interval is then as well used to assign the associated HI complex.

With this definition, depending on the chosen velocity interval, a CIV system can consist of two or more nearly isolated absorption features. With this approach we are methodologically close to volume-averaged quantities commonly derived in numerical simulations. To avoid confusion, the term absorber is only applied when we refer to the conventional definition of an absorption system.

On the basis of observational evidence, we choose to analyse the data using two velocity intervals $[-250, +250]$ km s⁻¹, and $[-600, +600]$ km s⁻¹. Studies of close QSO pairs have found a strong CIV clustering within ~ 200 km s⁻¹ (Rauch et al. 2005; D’Odorico et al. 2006). Also a significant clustering signal is present at a transverse velocity separation of $v_{\perp} \sim 500$ km s⁻¹ (D’Odorico et al. 2006). This is in agreement with the predicted CIV distribution in the forest obtained with numerical simulations using outflow velocities of less than 600 km s⁻¹.

Starting from the longest wavelength of each spectrum, we flag each isolated CIV group as CIV system for a given velocity interval. The system redshift is then defined to be the flux minimum of the CIV group. All CIV absorbers in the velocity interval are then assigned to the system. Whenever the shortest separation to a neighbouring CIV absorption profile (including any shallow broad absorption wing profile with normalised flux $F < 0.95$) is less than half of the velocity interval (i.e. 250 km s⁻¹ or 600 km s⁻¹), the system’s velocity interval is increased towards the absorber to 600 km s⁻¹ or 1200 km s⁻¹. If the newly included absorber contains the CIV flux minimum, the system redshift is adopted and the velocity is redefined accordingly.

The total CIV or HI column density of a system is then determined, by adding all the N_{CIV} or N_{HI} of all robustly determined components inside the velocity range.

An illustration of this assignment algorithm is given in Fig. 2.2, where we show four typical examples of CIV systems. These examples demonstrate how the $[-250, +250]$ km s⁻¹ velocity range is applied to identify CIV systems. The zero velocity is set at the redshift of the CIV flux minimum. The observed normalised spectrum is plotted as a dotted line, and the model spectra is given by the black line. Thin and thick tick marks highlight the

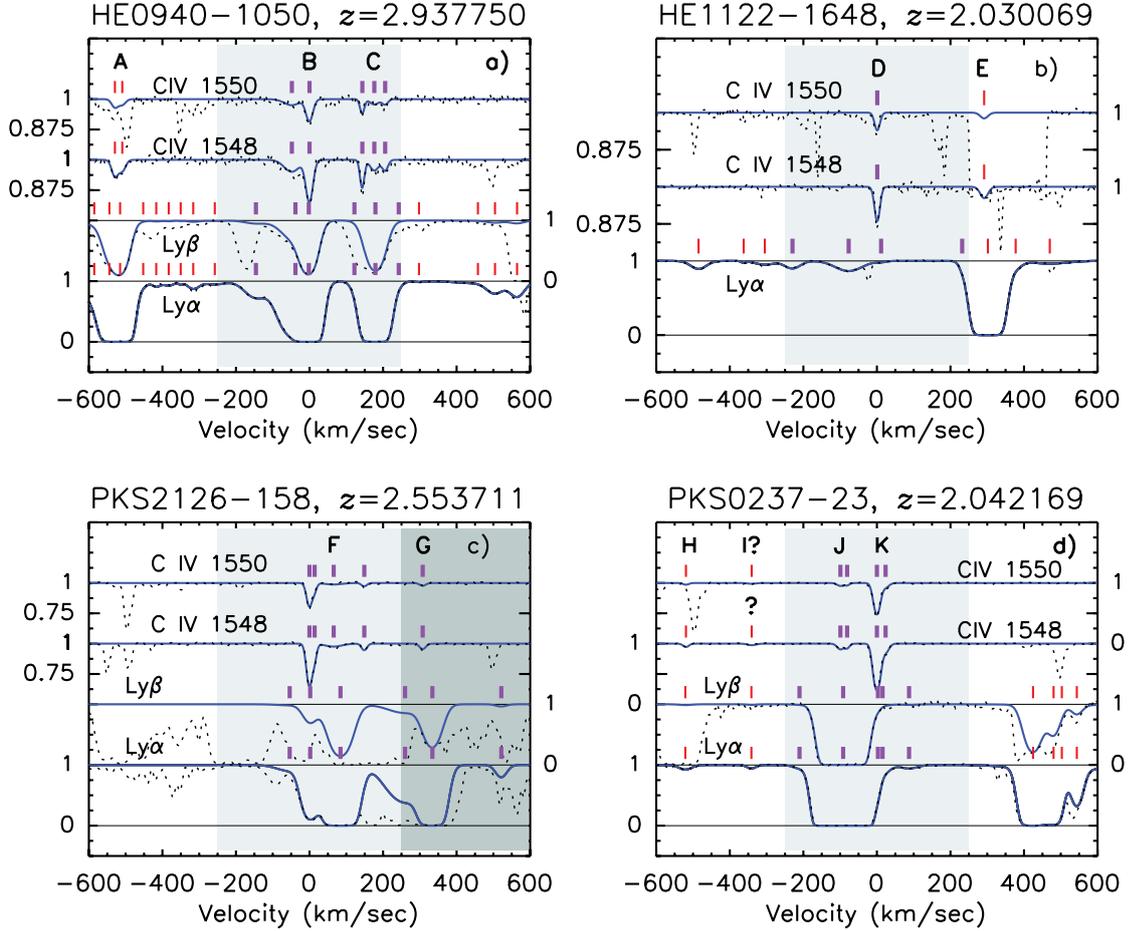


Figure 2.2.: Velocity plot of 4 C IV systems with a velocity interval of $\pm 250 \text{ km s}^{-1}$ (light grey shaded area). The systems are centred on the strongest flux in C IV. The blue solid line gives the model spectra, whereas the black dotted line indicates the observed normalised flux. Tick marks indicate the velocity position of each fitted component, where the thick ticks mark components belonging to a C IV system (see text for further discussion). Capital letters mark absorption features for reference in the discussion.

2. The forest CIV at $2 < z < 3.5$

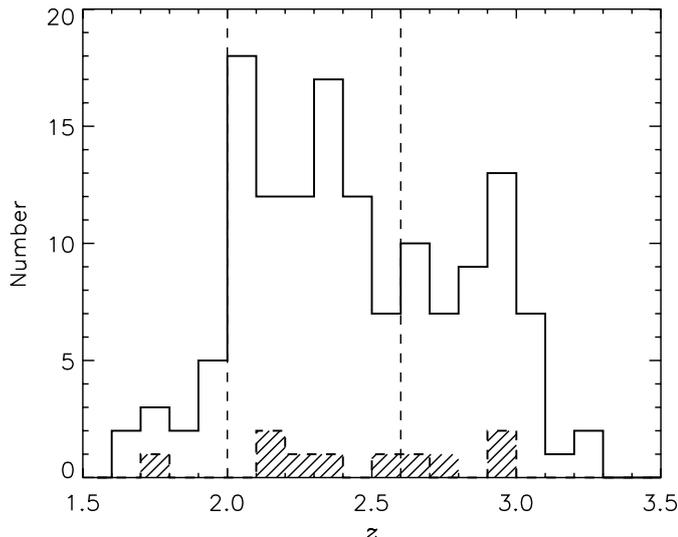


Figure 2.3.: The number of certain CIV system (solid line) as a function of z . The shaded area marks the number of uncertain CIV systems. Uncertain systems are CIV components detected at $2 - 3\sigma$ which do not fulfil the weaker detection criterion. The vertical dashed lines delineate the division of our sample into redshift ranges $z = [2.0, 2.6)$ and $z = [2.6, 3.4]$.

velocities of the fitted components, where the thick tick marks denote components lying within the velocity range. The velocity range is marked by the grey shaded area.

Fig. 2.2 panel **a)** is a good example for illustrating the assignment method. The two CIV groups B and C lie within the $\pm 250 \text{ km s}^{-1}$ range and they can therefore be considered as a system. The redshift of the system is then centred on the CIV flux minimum. Further the separation between the right wing of the CIV profile of group A and the left wing of group B is larger than 250 km s^{-1} , therefore the velocity range is not extended and groups B+C are one system. The HI absorbers assigned to the CIV system are the ones lying strictly inside the velocity range.

Panel **b)** presents also a very clear case. However this system illustrates the limitation of our definition of a system. In our definition, the 7th HI component from the left is assigned to be associated with the CIV system D. However it is more likely, that this component is associated with component E. Our working definition forces us to include this component in the total N_{HI} . On the other hand, volume-averaged parameters used in simulations would suffer from similar limitations.

Panel **c)** gives an example, where the velocity range is extended. The left wing of group G is separated by less than 250 km s^{-1} from the right wing of F. Therefore the velocity range was extended to the right up to 600 km s^{-1} . After the extension, the redshift remains unchanged, since the flux minimum lies still in group F.

In panel **d)**, group I has a $\leq 3\sigma$ detection and the weaker detection limit cannot be applied since the O VI line lies outside the covered wavelength range. Therefore the velocity range is not extended to the left, since group H is too far away from the left wing of group J. If group I would have been a real detection, the velocity range would have been expanded to include H+I+J+K.

In Fig. 2.3 the number of robustly detected CIV systems are compared to the uncertain

detections. These uncertain detections are systems that do not satisfy the weaker detection criterion. Since the z distribution of C IV systems shows a dip at $z \sim 2.7$, we divided the total sample into two redshift ranges $z = [2.0, 2.6)$ (mean redshift $\langle z \rangle = 2.22$) and $z = [2.6, 3.4]$ ($\langle z \rangle = 2.76$). The total redshift path length excluding any wavelength gaps is 19.61 for $\langle z \rangle = 2.22$ and 8.09 for $\langle z \rangle = 2.76$. The number of robust C IV systems are [78, 49] for $\langle z \rangle = [2.22, 2.76]$ and [5, 4] systems are uncertain.

To obtain further insight into the physical nature of the C IV systems, we further classified all C IV systems based on the existence of other ions such as Si IV. Systems without Si IV are denoted as C IV-Si IV systems. Systems with Si IV are marked as C IV+Si IV, including partly blended Si IV. Partly blended components are narrow and weak components with a column density below $N \leq 10^{11.9} \text{ cm}^{-2}$ which are blended with H I forest lines. The position of these partly blended Si IV lines can be estimated from the presence of other ions. When no detection of Si IV is possible due to blending, the system is denoted as C IV+bl system. The total number of C IV-Si IV, C IV+Si IV, and C IV+bl systems is [43, 23, 12] and [24, 17, 8] at $\langle z \rangle = 2.22$ and 2.76.

2.3. The relations between H I and C IV column densities

In Fig. 2.4 we show the $N_{\text{H I}} - N_{\text{C IV}}$ relation for the two redshift intervals and the two system velocity ranges of $\pm 250 \text{ km s}^{-1}$ and $\pm 600 \text{ km s}^{-1}$. In both panels we plot C IV-Si IV (black open circles), C IV+Si IV (red open squares), and C IV+bl systems (blue open triangles). Further green open circles in the lower redshift bin indicate class "4" C IV systems, which have $N_{\text{H I}}$ estimates from Ly α lines only. Even though their column densities cannot be robustly determined, their location in the $N_{\text{H I}} - N_{\text{C IV}}$ plane is not significantly different to the more robustly determined systems at $\log N_{\text{H I}} \leq 17 \text{ cm}^{-2}$. Nevertheless, the scatter at higher redshifts is increased by these class "4" systems. However, to be conservative, these systems are excluded from further analysis. The yellow open diamonds at $\log N_{\text{H I}} \sim 13.7 \text{ cm}^{-2}$ in the $\langle z \rangle = 2.22$ range denote C IV systems that show characteristics of a covering factor of less than 1. These systems are so close to the AGN that it does not cover the emitting source completely. A contribution to the observed flux still remains from the uncovered emitting source and thus the C IV doublet ratio is not 2:1. Such systems have been excluded from further analysis as well, since they do not represent typical intervening IGM systems. In order to be conservative the uncertain C IV systems at the $2 - 3\sigma$ detection limit (grey open upside-down triangles) have been excluded from the sample as well.

We have obtained least-square fits, $\log N_{\text{C IV}} = A \times \log N_{\text{H I}} + B$ for the sample in Fig. 2.4 for systems with $0.001 \leq N_{\text{C IV}}/N_{\text{H I}} \leq 0.01$. The $N_{\text{C IV}}/N_{\text{H I}}$ constraint was applied to omit some outliers. With this we obtain for $\langle z \rangle = [2.22, 2.76]$ a slope of $A = [0.86 \pm 0.05, 1.04 \pm 0.06]$ and a normalisation of $B = [-0.25 \pm 0.83, -3.13 \pm 0.87]$.

At both redshifts, the C IV+Si IV systems are located in the upper right corner of the $N_{\text{H I}} - N_{\text{C IV}}$ plane. This is due to the fact, that Si IV is usually associated with high column density absorbers (Rauch et al. 1997a). At $\langle z \rangle = 2.76$ most of the C IV systems seem to lie on a straight line with a strong scatter around it. This is what can be expected from the optical depth analysis by Schaye et al. (2003), which shows that with decreasing H I optical depth, the corresponding C IV optical depth decreases as well. From this we

2. The forest CIV at $2 < z < 3.5$

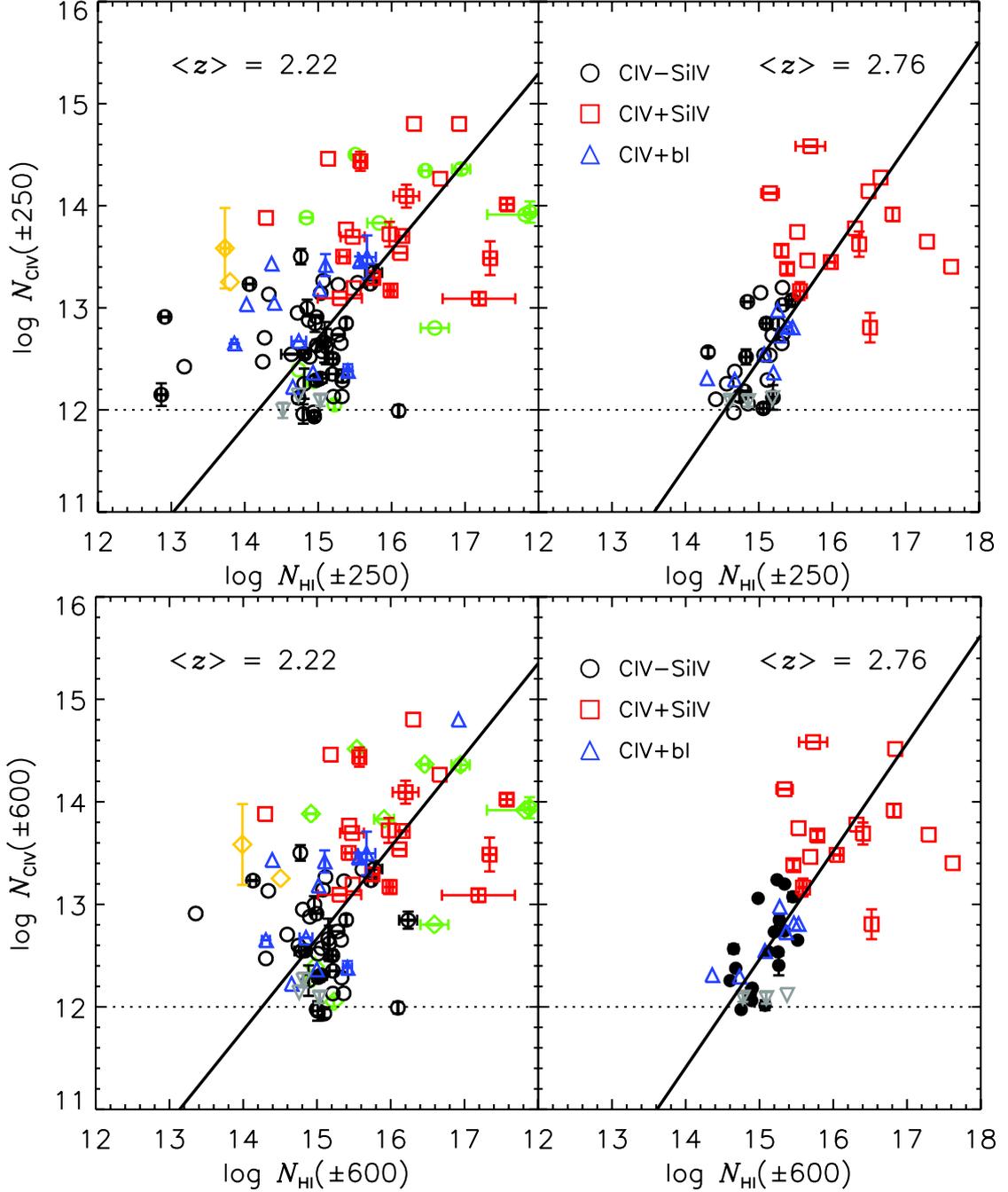


Figure 2.4.: The $N_{\text{HI}} - N_{\text{CIV}}$ relation for a system's velocity range of $\pm 250 \text{ km s}^{-1}$ (upper panel) and $\pm 600 \text{ km s}^{-1}$ (lower panel). The left panels shows results of the $\langle z \rangle = 2.22$ bin, the right one shows $\langle z \rangle = 2.76$. Open circles indicate CIV-SiIV systems, open triangles CIV+SiIV systems, and upside-down triangles CIV+bl systems. Green open circles indicate class "4" systems. Grey open upside-down triangles indicate uncertain systems at the $2 - 3\sigma$ detection limit. Error bars are only plotted when they exceed $\Delta(\log N) \geq 0.1$. The solid line represents the linear regression to the data. The dotted horizontal line marks the $N_{\text{CIV}} = 12.0 \text{ cm}^{-2}$ detection limit.

2.3. The relations between H I and C IV column densities

analogously expect a linear relation between the H I and C IV column density. However at low redshifts, the data clearly deviates from a linear relation. Above $\log N_{\text{HI}} \sim 15.5$ most systems still lie on a linear relation, however for lower H I column densities most systems are situated below such a relation. There seems to be a steep decrease in N_{CIV} at $\log N_{\text{HI}} \leq 15.2 \text{cm}^{-2}$ and $\log N_{\text{CIV}} \leq 13.0 \text{cm}^{-2}$. This behaviour is not dependent on the choice of the system's velocity range. We will therefore focus on the $\pm 250 \text{ km s}^{-1}$ sample.

Further there are many C IV systems at $\langle z \rangle = 2.22$ located to the far left of the power-law fit, which are not seen in the higher redshift bin. These systems show much larger $N_{\text{CIV}}/N_{\text{HI}}$ fractions, than the majority of the sample. These absorbers are not characterised by saturated H I absorption lines, and are thought to be highly ionised and highly metal-enriched. Some of these systems have been previously studied including other ions not only C IV (Carswell et al. 2002; Schaye et al. 2007). Motivated by these studies we will call systems which have unsaturated H I Ly α components that are directly associated with C IV "highly ionised systems". Using this definition we identify [33, 11] highly ionised systems at $\langle z \rangle = [2.22, 2.76]$. We obtain a factor of three more highly ionised systems at low redshifts than for the high redshift bin. This mainly arises from the 2.4 times larger redshift distance coverage of our sample in the low redshift bin.

highly ionised systems are marked in Fig. 2.5 (concentric symbols), where we show the robust data points from Fig. 2.4 for the $\pm 250 \text{ km s}^{-1}$ velocity range, now including further data obtained from the literature. It can be seen, that the highly ionised systems change their location in the $N_{\text{HI}} - N_{\text{CIV}}$ plane. There are more highly ionised systems in the lower-left part of the $N_{\text{HI}} - N_{\text{CIV}}$ plane at lower redshifts than in the higher redshift bin. This is not the result of a selection bias, since these systems are easy to identify due to their strong C IV absorption features.

As shown in Fig. 2.4, a single power-law cannot convincingly describe the $N_{\text{HI}} - N_{\text{CIV}}$ relation at $\langle z \rangle = 2.22$. Numerical simulations have predicted a rather steep decrease in metallicity for N_{HI} lower than a drop-off $N_{\text{HI,drop}}$ (Gnedin 1998a; Aguirre et al. 2001b; Oppenheimer & Davé 2006). Aguirre et al. (2001b) find a drop-off overdensity of $\log(1 + \delta) \sim 0.5$ for a wind model using an outflow velocity of 600 km s^{-1} and $\log(1 + \delta) \sim 1.0$ for an outflow velocity of 300 km s^{-1} . In the model using galaxy mergers as IGM enrichment process, Gnedin (1998a) find a drop-off at overdensities $\log(1 + \delta) \sim 0.0$. Further they predict a slowly increasing metallicity above this drop-off threshold (Springel & Hernquist 2003). In order to verify this, a larger H I column density range is required. In Fig. 2.5 we have therefore included higher column density absorption systems such as Lyman limit / sub-damped Lyman alpha / damped Lyman alpha systems compiled from the literature. Data for forest and Lyman limit systems have been compiled from Levshakov et al. (2003a,b) and Agafonova et al. (2007) (marked as stars). Further data from Cowie et al. (1995) and Songaila & Cowie (1996) has been used (denoted as crosses). Upside-down triangles representing sub-damped Lyman alpha systems are taken from Dessauges-Zavadsky et al. (2003) and Péroux et al. (2007). The damped Lyman alpha systems (diamonds) have been taken from Prochaska's HIRES DLA database webpage (<http://kingpin.ucsd.edu/~hiresdla/>).

The upper x-axis of Fig. 2.5 shows the overdensity corresponding to the H I column

2. The forest CIV at $2 < z < 3.5$

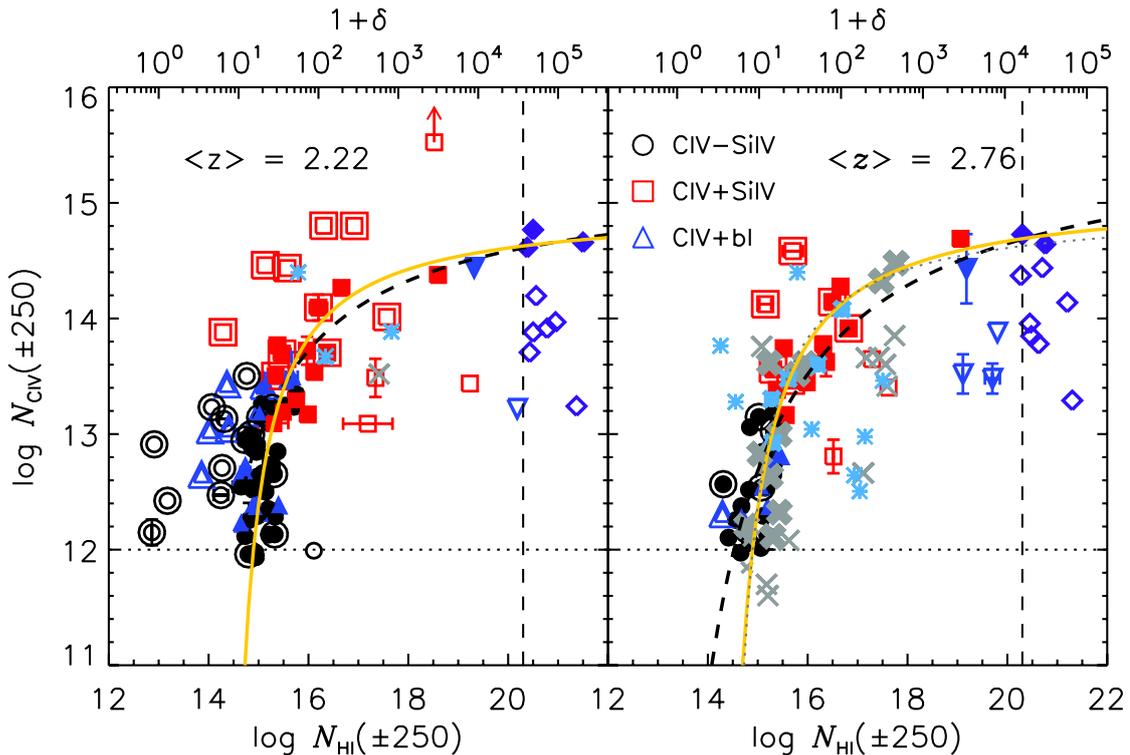


Figure 2.5.: The $N_{\text{HI}} - N_{\text{CIV}}$ relation including Lyman limit / sub-damped Lyman alpha / damped Lyman alpha systems compiled from the literature (see text for details). The symbols are identical to Fig. 2.4, except that concentric symbols mark highly ionised systems. The dashed vertical line marks the boundary above which systems are classified as Lyman limit systems. The dashed, dotted and solid curves represent fits to the data using Eq. 2.2. The dashed curve represents results from the first fitting method (see text for details). The solid line indicates results from the second method. Filled symbols mark data points used for this fit. Further, the dotted curve in the right panel indicates the fit obtained for the lower redshift bin.

density according to Schaye (2001)

$$N_{\text{HI}} \sim 2.7 \times 10^{13} (1 + \delta)^{3/2} T_4^{-0.26} \Gamma_{12}^{-1} \left(\frac{1+z}{4} \right)^{9/2} \left(\frac{\Omega_b h^2}{0.02} \right)^{3/2} \left(\frac{f_g}{0.16} \right)^{1/2} \text{ cm}^{-2}. \quad (2.1)$$

Here $1 + \delta$ denotes the gas overdensity, the gas temperature is $T = T_4 \times 10^4$ K, the photoionisation rate is $\Gamma = \Gamma_{12} \times 10^{-12} \text{ s}^{-1}$ and f_g is the fraction of mass in gas. The gas temperature in the IGM is derived using the effective equation of state $T = T_0(1 + \delta)^{\gamma-1}$ (Hui & Gnedin 1997). For Γ_{12} and γ we interpolated results by Bolton et al. (2008). We thus estimate $\gamma = [0.612, 0.457]$ and $\Gamma_{12} = [1.100, 0.967]$ at $\langle z \rangle = [2.22, 2.76]$. T_0 is assumed to be 2×10^4 K (Schaye 2001), $f_g = 0.16$, and $\Omega_b h^2 = 0.0227$.

The inclusion of the higher column density data from the literature clearly shows that the low-density IGM $N_{\text{HI}} - N_{\text{CIV}}$ relation has a similar shape as the predicted overdensity-metallicity relations by Gnedin (1998a) and Aguirre et al. (2001b). Our observations show a drop-off at $\log N_{\text{HI}} \sim 15.0$, which corresponds to an overdensity

of $\log(1 + \delta) \sim 1.0$, consistent with the findings by Aguirre et al. (2001b). Therefore we chose a simple fitting function motivated by these numerical results that is able to reasonably describe our data. We have fitted the data in Fig. 2.5 using a rectangular hyperbola:

$$\log N_{\text{CIV}} = \left[\frac{C_1}{\log N_{\text{HI}} + C_2} \right] + C_3 \quad (2.2)$$

In order to obtain a fit that compares reasonably to the predicted shape obtained by the numerical simulations as well as to the visual impression of the data having a cut-off column density at around $\log N_{\text{HI}} \sim 15.0$, we pursued two different fitting strategies. First, we obtained a fit by applying a selection criterion on the sample. The dashed curves in Fig. 2.5 represents a fit imposing $0.0005 \leq N_{\text{CIV}}/N_{\text{HI}} \leq 0.025$ and $12.0 \leq \log N_{\text{CIV}} \leq 14.75$. With this we obtain $C_1 = [-4.43 \pm 1.30, -7.57 \pm 1.80]$, $C_2 = [-13.04 \pm 0.44, -12.45 \pm 0.45]$, and $C_3 = [15.22 \pm 0.17, 15.65 \pm 0.19]$ at $\langle z \rangle = [2.22, 2.76]$.

Another fit is obtained by ensuring that the initial fitting parameters values pass through the central part of the cluster of filled and thick points in the plot by hand. For both redshifts, the initial parameters are set to be $C_1 = -1.8$, $C_2 = -14.3$, and $C_3 = 14.9$. A new set of points that are included in the fitting are then obtained by selecting all the data fulfilling $\pm(\log N_{\text{CIV,data}} - \log N_{\text{CIV,fit}}) \leq 0.2$ and $\log N_{\text{CIV}} \geq 12.0$. In an iterative procedure the following parameters are found with this method (solid curves in Fig. 2.5): $C_1 = [-2.01 \pm 0.35, -2.51 \pm 0.47]$, $C_2 = [-14.22 \pm 0.14, -14.08 \pm 0.20]$, and $C_3 = [14.96 \pm 0.05, 15.09 \pm 0.07]$. It is evident that the results strongly depend on the chosen fitting procedure.

From the fitting parameters it is clear, that there is no significant evolution with redshift in the overall form of the $N_{\text{HI}} - N_{\text{CIV}}$ relation for the non highly ionised systems. For comparison, the $\langle z \rangle = 2.22$ fit using the second fitting procedure is plotted as dotted curve in the $\langle z \rangle = 2.76$ panel. Excluding highly ionised systems, a Kolmogorov-Smirnov test gives 97% significance that the two data sets at the two redshifts are drawn from the same population. Including the highly ionised systems yields a probability of 79%.

The sharp decrease in N_{HI} at $\log N_{\text{HI}} \sim 15.2$ suggests that, taking the scatter around the fitted relation into account, the Ly α forest might not be metal-enriched for $\log N_{\text{HI}} < 14$. This does not exclude a metallicity floor with an absolute upper limit of $\sim 10^{-4} Z_{\odot}$ (corresponding to $\log N_{\text{CIV}} \sim 10$) in the forest though (Bromm et al. 2001; Simcoe et al. 2004).

2.4. Comparisons with the optical depth analysis

Studies on the HI optical depth τ_{HI} and its CIV counter part τ_{CIV} (optical depth $\tau = -\ln F$, where F is the normalised flux) have revealed that τ_{CIV} increases monotonically as a function of τ_{HI} above the confusion limit. However, contrary to our findings from the Voigt profile fitting analysis, no hint of a drop-off is present in the results from the optical depth analysis (Ellison et al. 2000; Schaye et al. 2003). Further, no signs indicating the existence of highly ionised systems are obtained with the optical depth analysis. These differences are not a result of the fact that the optical depth analysis does not use volume-averaged values. In a sense, the optical depth at a given pixel (i.e. 0.05 Å in our study) is volume averaged as well, however over a much smaller volume. This

2. The forest CIV at $2 < z < 3.5$

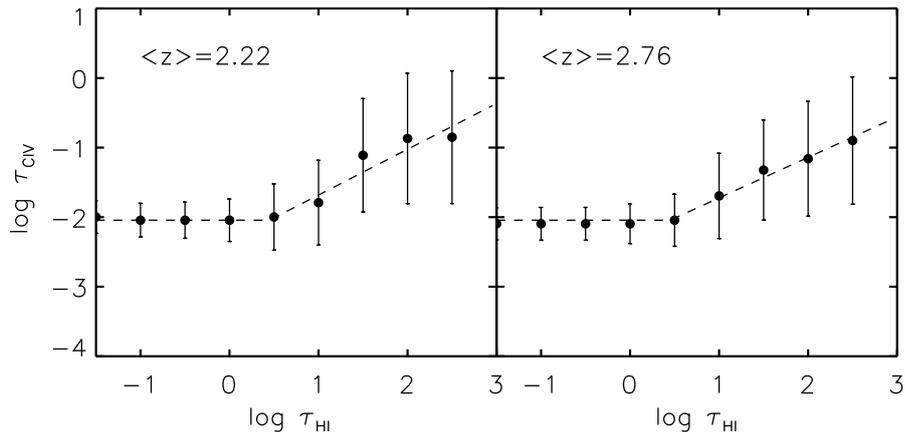


Figure 2.6.: The $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation derived from our QSO sample. Median opacities are used for both H I and C IV. The lowest detectable τ_{CIV} was set to be -2.04 ($F = 0.99$) to account for noises and continuum fitting errors.

corresponds to an averaging of $\sim 3.7 \text{ km s}^{-1}$, compared with our choice of $\pm 250 \text{ km s}^{-1}$ intervals.

In order to understand the differences between the two analysis methods, we have re-analysed the QSO spectra used for the study of the $N_{\text{HI}} - N_{\text{CIV}}$ relation with the optical depth analysis method. This ensures that the same C IV systems uncovered with the Voigt profile fitting method contribute to the $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation. This will also include all the highly ionised systems present in the sample.

2.4.1. The $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation

The $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation derived from the median optical depths in the two redshift bins is shown in Fig. 2.6. The optical depths were calculated from artificial spectra generated with the fitted parameters, including only H I and C IV $\lambda 1548$. By using the model spectra we easily avoid any contamination from other ions and the weaker C IV doublet that might blend with the stronger doublet at different redshifts. No noises were added. Hence, the lowest τ_{CIV} is in theory zero. However, we set the lowest τ_{CIV} to 0.01, which corresponds to $F = 0.99$ or $\log \tau_{\text{CIV}} = -2.04$. With this we imitate errors due to noise and continuum fitting.

In the calculation of the H I optical depth, we use all the available higher order Lyman transitions. Due to the observational detection limits set by continuum fitting errors, noise and contamination by other ions, any relation between τ_{HI} and τ_{CIV} is lost below a critical H I optical depth $\tau_{\text{HI},c}$. From the lowest τ_{CIV} we thus obtain a critical optical depth of $\tau_{\text{HI},c} \sim 0.45$.

From Fig. 2.6 it can be seen, that our results of the optical depth analysis are indeed consistent with previous findings. As in previous studies, τ_{CIV} increases proportional to τ_{HI} above the critical optical depth. In order to compare our results with detailed determination of the $\tau_{\text{HI}} - \tau_{\text{CIV}}$ relation by Schaye et al. (2003), we fit a power-law to the data for $\tau_{\text{HI}} \geq \tau_{\text{HI},c}$. In agreement with Schaye et al. (2003), we find $A = [-2.04, -2.04]$ and $B = [0.66, 0.58]$ at $\langle z \rangle = [2.22, 2.76]$ for $\log \tau_{\text{CIV}} = \log A + B \log (\tau_{\text{HI}}/\tau_{\text{HI},c})$.

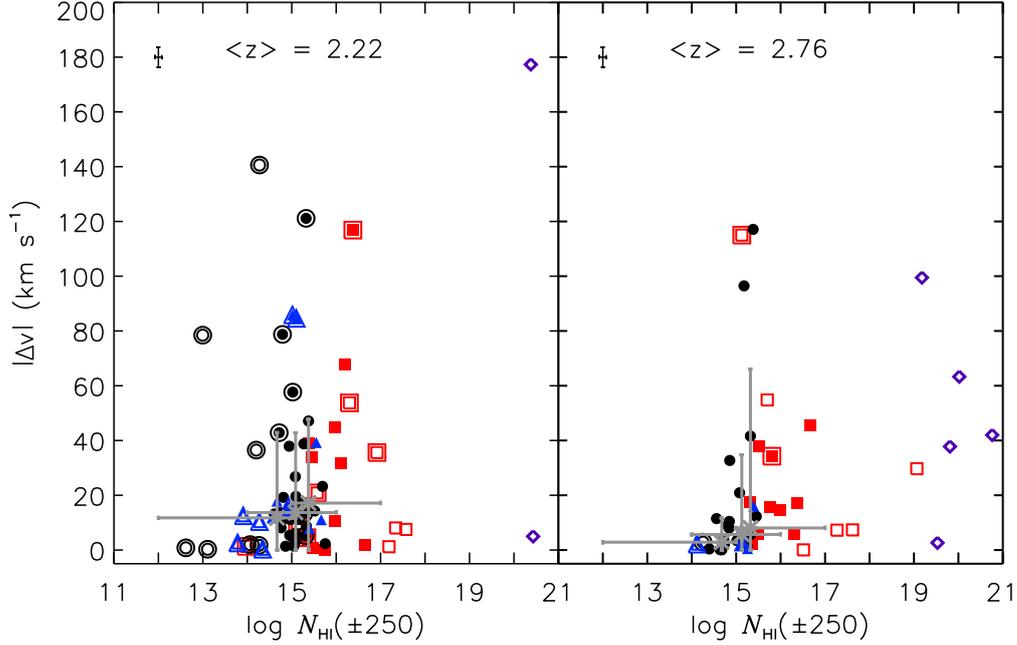


Figure 2.7.: Velocity offset between the absorption centroids of H I and C IV of individual absorbers, using Definition I. The symbols are equal to Fig 2.4. The grey data points with error bars mark the median values given in Table 2.2, with the respective lower and upper quartiles.

Neither the $\tau_{\text{C IV}}$ drop-off nor highly ionised systems are detectable with the optical depth analysis in our data.

2.4.2. Velocity offset between H I and C IV

C IV absorbers are not necessarily centred on H I absorbers and can be shifted in velocity space. Ellison et al. (2000) have determined a velocity offset dispersion of 17 km s^{-1} between H I and C IV absorbers for $\log N_{\text{H I}} \sim 13.6 - 16.0$ at $z \sim 3.45$. If the velocity offset between H I and C IV is significantly large, a one-to-one correspondence between $\tau_{\text{H I}}$ and $\tau_{\text{C IV}}$ does no longer exist. As a result, any signal of the $\tau_{\text{C IV}}$ drop-off could have been smeared out. We therefore determine the velocity offset between H I and C IV in our data.

In order to measure the velocity offset between H I and C IV, the absorbers centroid needs to be defined. We use three different possible definitions for the absorbers centroids:

- **Definition I:** The C IV and the H I absorption centroids are assumed to be the redshift of the C IV flux minimum and the strongest fitted $N_{\text{H I}}$ component for a given system in the $\pm 250 \text{ km s}^{-1}$ sample.
- **Definition II:** The C IV absorption centroid is defined as the flux minimum of an absorber, while the H I centroid is the redshift of the strongest fitted H I component close to the C IV absorber.
- **Definition III:** The C IV centroid is assumed to be the redshift of the C IV flux

2. The forest CIV at $2 < z < 3.5$

minimum, while the H I centroid is the redshift of the closest H I component to the C IV component.

As is evident from the different definitions, Definition I will give the largest and Definition III the smallest median velocity offset. To be conservative and consistent with the $\pm 250 \text{ km s}^{-1}$ sample discussed in Fig. 2.5, we use Definition I to estimate the H I-C IV median velocity offset. The velocity offset for all systems using Definition I is plotted in Fig. 2.7, while in Table 2.2 we provide the median velocity offsets for all the different centroid definitions.

An offset between H I and C IV is clearly present in the sample. At $\log N_{\text{H I}} = 14 - 16$ where most of the non-highly ionised C IV systems are detected, we obtain a median velocity offset of $[13.7 \text{ km s}^{-1}, 5.7 \text{ km s}^{-1}]$ for $\langle z \rangle = [2.22, 2.76]$. This corresponds to 3.6 pixels and 1.5 pixels at equal redshift ranges for our sample. Using Definition III, the velocity offset for equal column densities would correspond to 1.9 pixels and 1.3 pixels for the two redshift bins. In the worst case, this small velocity offset redistributes $\tau_{\text{H I}}$ and $\tau_{\text{C IV}}$ mostly within 4 pixels. We therefore conclude, that this misalignment cannot be the primary reason for the lack of a drop-off and the highly ionised systems in the $\tau_{\text{H I}} - \tau_{\text{C IV}}$ relation.

2.4.3. The fraction of H I pixels associated with C IV

As we have seen above, the C IV drop-off lies at around $\log N_{\text{H I}} \sim 15.2$ for both redshift bins. Assuming a b -parameter of $b = 30 \text{ km s}^{-1}$ for a Voigt profile, the optical depth in the centre of a $\log N_{\text{H I}} = 15.2$ is $\tau_{\text{H I}} \sim 50$. Inspecting Figs. 4 and 5 of Schaye et al. (2003) reveals a hint of a sudden $\tau_{\text{C IV}}$ drop-off at $\tau_{\text{H I}} \sim 50$ in their best-quality spectrum towards Q1422+230. Unfortunately, the errors associated with $\tau_{\text{C IV}}$ are too large to conclusively identify this hint of a drop-off as real. Within the errors, a linear $\log \tau_{\text{H I}} - \log \tau_{\text{C IV}}$ relation is still a good description of the data. However, this suggests, that the discrepancy between the two analysis methods could be caused by the rarity of C IV pixels associated with H I.

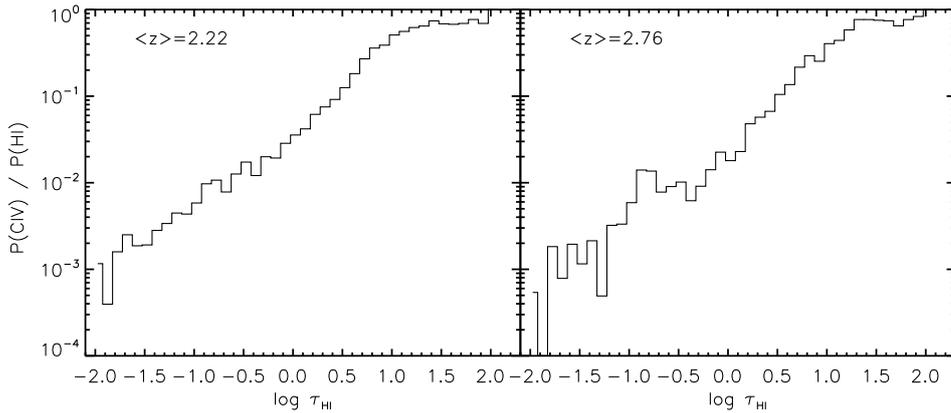
We determine the fraction of H I pixels that are associated with C IV. We relate the number of H I pixels showing associated C IV absorption $P(\text{C IV})$ (pixels with $\tau_{\text{C IV}} > 0$) with the number of pixels showing H I absorption $P(\text{H I})$. In Fig. 2.8 the ratio of $P(\text{C IV})/P(\text{H I})$ as a function of $\tau_{\text{H I}}$ is presented. Both quantities are derived from artificially generated spectra using the fitted parameters, including only H I and C IV $\lambda 1548$. As above, noises were not added to the model spectra. The value of $P(\text{C IV})/P(\text{H I})$ is noisier in the $\langle z \rangle = 2.76$ bin than for the lower redshift bin, due to the lower number of available pixels. Between $0.1 \leq \log \tau_{\text{H I}} \leq 1.5$, the percentage of H I pixels associated with C IV increases monotonically from around 4% (2%) up to 70% (70%) at redshifts $\langle z \rangle = 2.22$ (2.76). Any inclusion of noises in the analysis would decrease these values especially at lower optical depth. Further, the wings of strong absorption systems have lower $\tau_{\text{C IV}}$, additionally smoothing any possible signal (Pieri & Haehnelt 2004).

The drop-off column density of $\log N_{\text{H I}} \sim 15.0$ corresponds to an optical depth of $\log \tau_{\text{H I}} \sim 1.7$. Therefore a drop-off in $\tau_{\text{C IV}}$ for lower H I optical depths should be seen in the $\tau_{\text{H I}} - \tau_{\text{C IV}}$ relation. However, around this H I optical depth, about 60-70% of the H I pixels are associated with C IV. Therefore, a signal should be visible in the median

Table 2.2.: The median H I–C IV velocity offset^a

$\log N_{\text{HI}}$ range	$\langle z \rangle = 2.22$	$\langle z \rangle = 2.76$
Definition I		
12.0–15.0	11.7 (11.7) km s ⁻¹	2.9 (3.7) km s ⁻¹
14.0–16.0	13.7 (13.2) km s ⁻¹	5.7 (5.7) km s ⁻¹
15.0–17.0	17.1 (14.9) km s ⁻¹	8.0 (5.7) km s ⁻¹
Definition II		
12.0–15.0	4.9 (8.3) km s ⁻¹	4.1 (7.9) km s ⁻¹
14.0–16.0	13.4 (13.4) km s ⁻¹	9.4 (9.4) km s ⁻¹
15.0–17.0	16.6 (16.6) km s ⁻¹	9.9 (9.9) km s ⁻¹
Definition III		
12.0–15.0	0.5 (8.3) ^b km s ⁻¹	5.9 (6.3) km s ⁻¹
14.0–16.0	7.0 (6.4) km s ⁻¹	4.9 (4.9) km s ⁻¹
15.0–17.0	6.4 (6.2) km s ⁻¹	6.2 (5.8) km s ⁻¹

- ^a The velocity offsets listed in the parentheses are the ones excluding highly ionised systems. Since most highly ionised systems have $\log N_{\text{HI}} \leq 14$, the median velocity offset at the $\log N_{\text{HI}} = 12.0\text{--}15.0$ range is most significantly affected, depending on whether or not these systems are included.
- ^b Most highly ionised systems have their C IV flux minimum very close to the centre of the corresponding H I absorption cloud which is not necessarily the strongest component. Therefore, the median velocity offset is almost 0 km s⁻¹ when they are included. On the contrary, non-highly ionised C IV absorbers usually have a velocity offset between C IV and H I.


Figure 2.8.: The ratio of the number of pixels having $\tau_{\text{CIV}} > 0$ and the number of H I pixels as a function of τ_{HI} for the two redshift bins derived from our data.

2. The forest CIV at $2 < z < 3.5$

optical depths, since a sufficient number of HI pixels with associated CIV contribute to these optical depth bins. Unfortunately, the remaining 40-30% of CIV-free pixels smooth out any weak signal present in the median optical depths. Thus, any drop-off in τ_{CIV} cannot be determined clearly as long as the associated error bars are still too large.

2.5. Simple shell model for the $N_{\text{HI}} - N_{\text{CIV}}$ relation

From our Voigt profile fitting analysis, we have shown that there is a well-defined $N_{\text{HI}} - N_{\text{CIV}}$ relation for non-highly ionised systems. Similar relations have been predicted by numerical simulations using outflow enrichment models (Aguirre et al. 2001b) or a galaxy merging process (Gnedin & Ostriker 1997; Gnedin 1998a). This raises the question on the physical process behind the observed $N_{\text{HI}} - N_{\text{CIV}}$ and how far CIV has been transported from the stars in the galaxies that produced metals. We now want to address these issues with a simple shell model that is able to reproduce the observational data.

2.5.1. The observed rate of incidence of non-highly ionised CIV absorbers

One of the strongest observational evidences for galactic-scale outflows as the IGM metal enrichment mechanism is the correlation between high-redshift galaxies and CIV absorbers at $2 < z < 3$. By correlating Lyman break galaxies with CIV absorbers, Adelberger et al. (2005) found that most CIV absorbers with $N_{\text{CIV}} \geq 14$ are directly produced by a single galactic halo of $\sim 40 - 80 h^{-1}$ proper kpc in size. Weaker absorbers are not directly related with single galaxies, but are found nearby.

On the other hand, Chen et al. (2001) found that individual galaxies at $z < 0.9$ are surrounded by a CIV halo of $\sim 100 h^{-1}$ kpc down to $\log N_{\text{CIV}} \sim 13.4$ with a covering factor of nearly 1. Further they concluded, that higher equivalent width absorbers stem from brighter galaxies and have smaller impact parameters. All these finding would indicate that CIV halos grow with time, either by infall, or by outflow of metal enriched material.

Let us assume that every non-highly ionised CIV absorber associated with the Ly α forest at $\log N_{\text{HI}} \leq 17$ is produced by a single galaxy with a halo size R . The comoving number density of CIV-producing galaxies is denoted as n_g and their proper cross section πR^2 . The CIV covering factor is denoted by k . The rate of incidence of CIV absorbers $d\mathcal{N}/dz$ is then given by

$$\frac{d\mathcal{N}}{dz} = \pi R^2 n_g k \frac{c}{H(z)} (1+z)^3, \quad (2.3)$$

where c is the speed of light and $H(z)$ is the Hubble constant at z . Thus, the halo size R is

$$R = 43.3 h^{-1} k^{-1/2} \left(\frac{d\mathcal{N}}{dz} \right)^{1/2} \left(\frac{n_g}{0.0025 \text{ Mpc}^{-3}} \right)^{-1/2} \text{ kpc}, \quad (2.4)$$

where the galaxy density is normalised to that of Lyman break galaxies at $z \sim 3$ down to a R magnitude of 25.5 Steidel et al. (1999).

At $z \sim 3$ we find for non-highly ionised CIV absorbers with $\log N_{\text{CIV}} = [12.5 - 13.5, 13 - 14, \geq 14]$ the rate of incidence to be $d\mathcal{N}/dz = [11.9 \pm 2.6, 4.0 \pm 1.5, 1.1 \pm 0.8]$. The halo size R of a typical CIV-producing galaxy is thus about $[150 k^{-0.5}, 87 k^{-0.5}, 45 k^{-0.5}]$ proper kpc.

2.5. Simple shell model for the $N_{\text{HI}} - N_{\text{CIV}}$ relation

If the covering factor is 1, the halo size at $\log N_{\text{CIV}} \sim 14$ is consistent with the results by Adelberger et al. (2005), considering that Eq. 2.4 does not include any possible correlation between the halo size and the brightness of a galaxy (Chen et al. 2001) and the galaxy clustering properties.

Adelberger et al. (2005) on the other hand found that both N_{CIV} and the rate of incidence decreases with increasing impact parameter. This indicates that the covering factor k decreases with decreasing N_{CIV} . Assuming an average k value of 0.5 for all CIV absorber-galaxy pairs, a CIV-producing galaxy can have halo sizes up to 210 kpc for $N_{\text{CIV}} = [12.5 - 13.5]$. With even lower k and CIV column density, the size of the halo increases even further.

All this indicates that high-redshift galaxies are surrounded by a ~ 50 proper kpc sized CIV halo that produces strong CIV absorption. At larger radii CIV absorption is much weaker. If this picture is true, a strong correlation between the impact parameter and the N_{CIV} should exist, which motivates a shell model of CIV absorbers.

2.5.2. Simple shell model

Motivated by the previous subsection, we have considered whether or not the $N_{\text{HI}} - N_{\text{CIV}}$ relation can be reproduced by a simple shell model, similar to the one described by Churchill et al. (2005), e. g. their Fig. 3. In this picture, the size of halos enriched with a given ion around galaxies depends on the ion's column density. For example, a HI halo producing a Lyman limit system is about 35 proper kpc, while a halo producing $\log N_{\text{CIV}} \geq 13.9$ is about 60 proper kpc at $z < 1$.

We assume that the hydrogen gas follows a NFW halo, of mass $M_{\text{halo}} = 10^{11} h^{-1} M_{\odot}$, a concentration of $c = 4.3$ and virial radius $r_{\text{vir}} = 110 \text{ kpc } h^{-1}$ at redshift $z = 2.5$. Hydrogen is assumed to be in photoionisation equilibrium with a UV background having a photoionisation rate of $\Gamma_{\text{UVB}} = 1.2 \times 10^{-12} \text{ s}^{-1}$ based on the Haardt & Madau UV background including quasars and galaxies. Centred in this HI halo, we place a CIV absorber following a Gaussian profile of

$$\rho_{\text{CIV}}(r) = \rho_{0,\text{CIV}} \exp\left(-\frac{r^2}{R_{\text{h,CIV}}^2}\right), \quad (2.5)$$

where $\rho_{\text{CIV}}(r)$ is the CIV density as a function of galactocentric radius r and the CIV scale height $R_{\text{h,CIV}} = 17.3 h^{-1} \text{ kpc}$. The CIV density at the centre $\rho_{0,\text{CIV}}$ is assumed to be 1.2×10^7 times smaller than the characteristic density of the hydrogen NFW profile. The constants are chosen to reproduce the observed N_{CIV} distribution. Assuming photoionisation equilibrium, this model corresponds to a metallicity of $Z \sim 10^{-1.5} Z_{\odot}$ at a galactocentric radius of $r \sim 2 \text{ kpc}$ and $Z \sim 10^{-3} Z_{\odot}$ at $r \sim 20 \text{ kpc}$.

We can determine the column density along a line of sight through HI and CIV absorbers at a given impact parameter by $N = \int_{-\infty}^{\infty} \rho(\sigma, z) dz$, where σ is the impact parameter and z the position along the line of sight. The resulting column density profiles as a function of the impact parameter are shown in Fig. 2.9. The CIV column density profile drops rapidly with increasing impact parameter, since the inner denser core dominates. With smaller impact parameters, the sightline passes closer to the centre where the density of each gas component is higher. As a result, N_{HI} and N_{CIV} are higher

2. The forest CIV at $2 < z < 3.5$

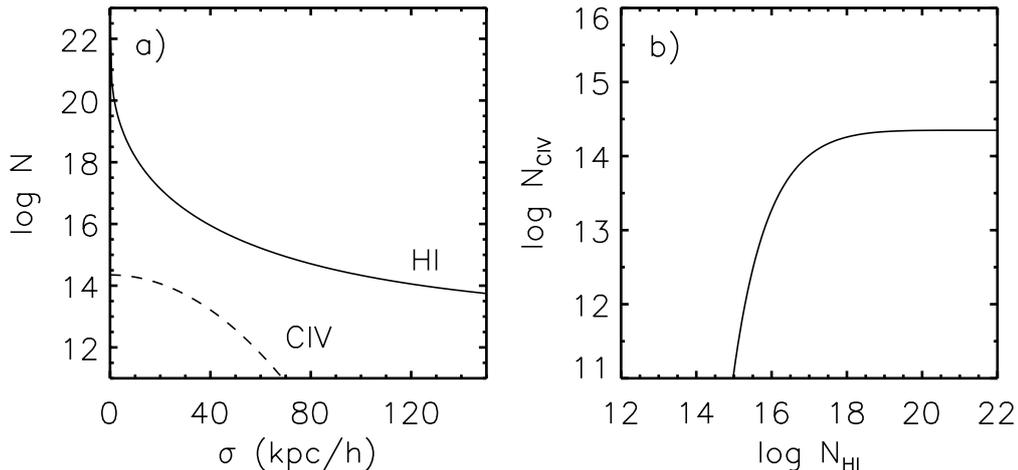


Figure 2.9.: The shell model of a halo of high-redshift galaxies. In this model, H I-producing gas is assumed to follow the NFW profile, while C IV-producing gas is assumed to have an exponential density profile. The C IV density at the centre is chosen to match the observed $N_{\text{HI}} - N_{\text{CIV}}$ relation. a) The H I and C IV profiles as a function of the impact parameter σ used in the model. b) The calculated $N_{\text{HI}} - N_{\text{CIV}}$ relation integrated along the line of sight at various impact parameters.

for small impact parameters, similar to what was found for redshifts smaller than 1 by Chen et al. (2001). A column density of $\log N_{\text{CIV}} = 11$ is already reached at the proper impact parameter of $\sigma = 70 h^{-1} \text{kpc}$. Such an interplay of these two profiles gives rise to a behaviour similar to the $N_{\text{HI}} - N_{\text{CIV}}$ relation.

Note that this model does not assume any metal ejection mechanisms from galactic disks. It simply states that as long C IV gas is Gaussian distributed around a galaxy up to radii of $\sim 70 h^{-1}$ proper kpc, the observed $N_{\text{HI}} - N_{\text{CIV}}$ relation can be re-produced. This C IV halo size is less than the virial radius of our model galaxy with $M_{\text{halo}} = 10^{11} h^{-1} M_{\odot}$, i.e. C IV is located inside the galactic halo and not in the surrounding IGM. Unfortunately, the C IV halo size derived from our shell model is more than a factor of 2 smaller than the size of C IV-producing galaxy halos estimated in the previous subsection. However, this should not be taken against the shell model, since in reality, a collection of galaxies with different masses would contribute to the $N_{\text{HI}} - N_{\text{CIV}}$ relation. Further, our simple assumption on the smoothly decreasing H I and C IV density profiles is likely too simplistic. It is more likely that much of the C IV comes from a halo shell, with the inner regions denser and more neutral. However, the reasonable agreement between this simple shell model and the observed $N_{\text{HI}} - N_{\text{CIV}}$ relation in the zeroth approximation encourages to explore a similar, but more sophisticated model to reproduce the $N_{\text{HI}} - N_{\text{CIV}}$ relation.

2.6. Summary

Numerical simulations based on different metal enrichment mechanisms of the IGM predicted the volume averaged overdensity-metallicity relation. Its shape and evolution constrain the physical processes that results in the enrichment of the IGM. In observations, the overdensity can be related to the column density of neutral hydrogen N_{HI} , while the

metallicity relates to the triply-ionised C IV column density N_{CIV} .

We have analysed 17 high resolution spectra of high-redshift QSOs obtained from the ESO VLT UVES archive covering the Ly α forest at $2 < z < 3.5$. Together with associated metals, all available H I Lyman lines were fitted using Voigt profiles. With this, it was possible to obtain robust column density measurements of H I and C IV. In order to measure volume-averaged N_{HI} and N_{CIV} of C IV absorber groups, we defined and identified C IV systems over $\pm 250 \text{ km s}^{-1}$ velocity intervals centred at the C IV flux minimum. Both N_{HI} and N_{CIV} were then integrated over this $\pm 250 \text{ km s}^{-1}$ interval.

The C IV-selected forest systems show two distinct populations, non-highly ionised systems and highly ionised systems. In the two redshift bin $\langle z \rangle = 2.22$ and $\langle z \rangle = 2.76$, non-highly ionised systems show a well-defined $N_{\text{HI}} - N_{\text{CIV}}$ relation, similar to the numerically predicted overdensity-metallicity relation. The $N_{\text{HI}} - N_{\text{CIV}}$ relation drops very sharply below $N_{\text{HI,drop}} \sim 10^{15.2} \text{ cm}^{-2}$. However, it remains only a weak function of N_{HI} above $N_{\text{HI,drop}}$. The $N_{\text{HI}} - N_{\text{CIV}}$ relation does not show any significant redshift evolution at the two redshift bins.

Based on observations which suggest that high-redshift galaxies are surrounded by a C IV halo with a size of $40 - 80 h^{-1}$ proper kpc, we have constructed a simple shell model which reproduces the observed $N_{\text{HI}} - N_{\text{CIV}}$ relation. In this shell model, a H I halo follows a NFW profile up to around $110 h^{-1}$ kpc, while a C IV halo centred on the H I halo is assumed to follow a Gaussian density profile with a scale height of about $17 h^{-1}$ kpc. At small impact parameters, the sight lines pass through denser areas giving rise to higher N_{HI} and N_{CIV} . In such a model, the C IV gas contributing to $\log N_{\text{CIV}} \geq 12$ is confined within a small radius of around $60 h^{-1}$ kpc. This is a factor of 2 smaller than the impact parameter estimated for C IV systems with $\log N_{\text{CIV}} \geq 12.5$ in the present study.

Further, we find that highly-ionised systems do not follow the $N_{\text{HI}} - N_{\text{CIV}}$ relation defined by the non-highly ionised systems. They all have a lower N_{HI} compared to non-highly ionised systems with similar N_{CIV} . There are only a few such systems, but it appears that their line number density peaks at $z \sim 2$, where the cosmic star formation rate is at its highest. However, more data is needed to support this in a conclusive way.

Contrary to the indications that the forest could be metal enriched down to $\log N_{\text{HI}} \sim 13.5$ from the optical depth analysis, our observations of a sharp cutoff in the $N_{\text{HI}} - N_{\text{CIV}}$ relation suggest that a forest with $N_{\text{HI}} \leq 10^{14} \text{ cm}^{-2}$ might not be metal-enriched. This also indicates, that only forest systems without metals would be truly considered as the intergalactic medium and are not inside galactic halos. However, this does not exclude the possibility of a small metallicity floor in the IGM.

The apparent contradiction between the optical depth analysis and the Voigt profile analysis is mainly caused by the fact that the optical depth analysis deals with all the forest including the C IV free forest, while our analysis only includes the forest with C IV. We conclude that correlating H I pixels to corresponding C IV pixels and taking a median statistical property does not necessary correlate any physically relevant quantities.

There is a small velocity offset between the centroids of H I and C IV. The exact value depends on the definition of the centroids between H I absorbers and the corresponding C IV. Regardless of the definition of both centroids, the velocity offset is largely independent with redshift. The maximum velocity offset is around 14 km s^{-1} for a forest with $\log N_{\text{HI}} = 14 - 16$. The lack of strong velocity offsets suggests that C IV systems are well-mixed with the H I clouds.

The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka!" (I found it!) but "That's funny..."

Isaac Asimov

3

Neutral hydrogen column density evolution of high-redshift QSO absorption line systems at $1.9 < z < 3.2$ ¹

We have investigated the distribution and evolution of neutral hydrogen (H I) column density of the intergalactic medium at $1.9 < z < 3.2$, using the 18 high resolution, high signal-to-noise QSO spectra obtained with the UVES at the VLT. This study is based on the two sets of a Voigt profile fitting analysis, one including all the available high-order Lyman lines to obtain reliable H I column densities, $N_{\text{H I}}$, of saturated lines, and another using only the Ly α transition. We have found that there is no significant difference between the Ly α -only fit sample and the high-order Lyman fit sample. From the high-order fit sample, we have found: 1) The mean number density evolution at $1.9 < z < 3.2$ only from our high-order fit sample can be described as $dn/dz \propto (1+z)^{1.28 \pm 0.20}$ and $dn/dz \propto (1+z)^{3.77 \pm 0.49}$ in the column density range of $\log N_{\text{H I}} = [12.75, 14.0]$ and $[14.0, 17.0]$ respectively, confirming that stronger absorbers evolve more rapidly. Our individual QSO sightlines including the best-quality data in the literature at $z > 1$ give $dn/dz \propto (1+z)^{2.42 \pm 0.15}$ and $dn/dz \propto (1+z)^{2.99 \pm 0.26}$ in the column density range of $\log N_{\text{H I}} = [13.1, 14.0]$ and $[14.0, 17.0]$. 2) The mean dn/dz for $\log N_{\text{H I}} = [14.0, 17.0]$ reveals a significant dip from a power-law at $z \sim 2.15$. 3) The differential distribution function is also well-described with a single power law, which shows a dip at $\log N_{\text{H I}} = [14.5, 18.0]$. At $z > 2.2$, the dip becomes less pronounced. The power-law exponent at $\log N_{\text{H I}} = [12.75, 14.0]$ is $[-1.49 \pm 0.06, -1.43 \pm 0.03, -1.39 \pm 0.05]$ for $z = [1.9 - 2.2, 2.2 - 2.7, 2.7 - 3.2]$. 4) The dip in dn/dz relates to the dip in the differential density distribution function, indicating an identical physical origin. Such an origin may be the increase of the ionising radiation due to the peak of the cosmic star formation rate at $z \sim 2$. 5) The dn/dz evolution of the C IV enriched forest is similar to the entire forest sample, differing only in a smaller normalisation value. 6) The distribution function of the unenriched forest is similar to the one from the entire sample, showing a dip at $\log N_{\text{H I}} = [14, 18]$ which becomes stronger at lower redshifts. However, the distribution function of the C IV enriched forest shows no such dip and it starts to flatten out at $\log N_{\text{H I}} \leq 15.5$. This can be explained as the enriched fraction of the forest decreases as $\log N_{\text{H I}}$ decreases. This also indicates that the C IV enriched forest and the unenriched forest reside in different spatial locations, i.e. the C IV enriched forest in the circum-galactic medium and the unenriched forest in the intergalactic medium.

¹This chapter will be submitted to MNRAS as: A. M. Partl, T.-S. Kim, R. F. Carswell, and V. Müller.

3.1. Introduction

The resonant Ly α absorption by neutral hydrogen (H I) in the warm ($\sim 10^4$ K) photoionised intergalactic medium (IGM) produces rich, narrow absorption features blueward of the Ly α emission line in high-redshift QSO spectra known as the Ly α forest. The Ly α forest contains $\sim 90\%$ of the baryonic matter at $z \sim 3$ and can be observed in a wide range of redshifts up to $z \sim 6$. Semi-analytic models and numerical simulations have been very successful at explaining the observed properties of the Ly α forest mainly at low H I column densities $N_{\text{H I}} \leq 10^{16} \text{ cm}^{-2}$. These theories have shown that the Ly α forest arises by mildly non-linear density fluctuations in the low-density H I gas, which follows the underlying dark matter distribution on large scales. Therefore, the Ly α forest provides a powerful observational constraint on the distribution and evolution of the baryonic matter in the Universe, hence the evolution of galaxies and the large-scale structure (Cen et al. 1994; Rauch et al. 1997b; Theuns et al. 1998a; Davé et al. 1999; Schaye et al. 2000b; Schaye 2001; Kim et al. 2002). In addition, a discovery of triply ionised carbon (C IV) associated with some of the forest absorbers suggests that the forest metal abundances can be utilised to probe early generations of star formation and the feedback between high-redshift galaxies and the surrounding IGM from which they formed (Cowie et al. 1995; Davé et al. 1998; Aguirre et al. 2001b; Schaye et al. 2003; Oppenheimer & Davé 2006).

The physics of the Ly α forest is mainly governed by two competing processes, the Hubble expansion and the ionising UV background radiation. The Hubble expansion, which causes the gas to cool adiabatically, is rather well-constrained by the cosmological parameters and the primordial power spectrum from the latest WMAP observations (Jarosik et al. 2011). On the other hand, the ionising UV background radiation mainly controls the photoionisation heating and the gas ionisation fraction, thus determining the fraction of the observable H I gas compared to the hard to observe H II gas. Unlike the cosmological parameters, the background radiation is not well constrained from both observations and theory. The background radiation is assumed to be provided primarily by the H I ionising UV background from QSOs and in some degree also from star-forming galaxies (Shapley et al. 2006; Siana et al. 2010) and Ly α emitters (Iwata et al. 2009). However, the relative contribution from QSOs and galaxies as a function of redshift is not well known. Note that shock heating is not important in the forest due to its low density (Theuns et al. 1998b; Schaye et al. 2000b).

Two main tools to probe the properties of the forest $N_{\text{H I}}$ are the differential column density distribution function (the number of absorbers per unit absorption path length and per unit column density, an equivalent to the galaxy luminosity function) and the number of absorbers for a given H I column density range per unit redshift, dn/dz . In particular, dn/dz has been shown to be sensitive to the UV background radiation and its evolution as well as the relative contribution from QSOs and galaxies. At high redshift, the Hubble expansion decreases the gas density. A lower gas density results in a strong reduction of the recombination rate, allowing the gas to settle in a photoionisation equilibrium with a high ionisation fraction. With the nearly constant background radiation, this causes a steep number density evolution. At low redshift, the decrease of the QSO number density at $z < 2$ also decreases the available ionising photons. This changes the ionisation fraction in the gas and also counteracts the slow-downed Hubble expansion, resulting

in a constant number density evolution (Theuns et al. 1998a; Davé et al. 1999; Bianchi et al. 2001).

The result from the HST/FOS QSO Key project shows such a slow change in the dn/dz evolution at $z < 1.5$ (Weymann et al. 1998), compared to a much steeper dn/dz evolution at $z > 2$ (Kim et al. 1997, 2001, 2002). The cosmic variance also seems to increase at lower z (Kim et al. 2002). Unfortunately, recent works which based on better-quality data at $z < 1.5$ have shown rather ambiguous results with a large scatter in dn/dz along different sightlines. The only certain observational fact is that all of these studies show a factor of $\sim 2 - 3$ lower number densities than the Weymann et al. values at $z < 1.5$. Considering the lack of constraints in dn/dz from good-quality data at $1 < z < 1.5$, the redshift evolution of dn/dz can be viewed as a single power law without any change in dn/dz at $0 < z < 3.5$. In order to constrain the evolution of the UV ionising radiation, it should be a prerequisite to probe the robust evolution of dn/dz as well as the nature of the cosmic variance at $0 < z < 3.5$.

Here we present the in-depth Voigt profile fitting analysis of 18 high resolution ($R \sim 45\,000$), high signal-to-noise ($\sim 35 - 50$) QSO spectra obtained with the UVES instrument at the VLT, covering the forest at $1.7 < z_{\text{forest}} < 3.5$. Our main scientific aims are to derive the redshift evolution of the absorber number density and the different column density distribution function from the largest and most homogeneous set of data available up to now at $z > 2$. Most previous $N_{\text{H I}}$ studies utilising high-resolution, high signal-to-noise data at $z > 2$ have been based on less than 5 sightlines. Even with a couple of sightlines, the statistics on the weak forest is rather robust due to the large number of weak absorbers (about 400 absorbers at $z \sim 2.5$ per sightline). However, statistics of the strong forest with $N_{\text{H I}} \geq 10^{15} \text{ cm}^{-2}$ are not robust enough (about 35 absorbers at $z \sim 2.5$ per sightline). Cosmic variance also plays an important role at lower redshifts, especially for stronger absorbers (Kim et al. 2002). Therefore, increasing the sample size at $z \sim 2$ is critical in addressing a robust $N_{\text{H I}}$ evolution. In addition to this, we have improved previous results in two senses. First, most previous studies on the forest $N_{\text{H I}}$ have been based on Ly α -only profile fitting. This approach does not provide a reliable $N_{\text{H I}}$ for saturated lines, about $N_{\text{H I}} \geq 10^{14.5} \text{ cm}^{-2}$ for the present UVES data. Therefore, we have included all the available high-order Lyman series to derive a reliable $N_{\text{H I}}$. Second, in recent years there has been some evidence that metals associated with the high-redshift Ly α forest are resided in the circum-galactic medium rather than in the intergalactic space far away from galaxies (Steidel et al. 2010). Therefore, we have divided our sample into two groups, one with C IV (the C IV forest) and another without C IV (the C IV-free forest), in order to probe whether there is any difference in the $N_{\text{H I}}$ evolution between the two samples.

This chapter is organised as follows. Section 3.2 describes the analysed data and both Voigt profile fitting methods. Comparisons with previous studies based on the Ly α -only fit are shown in Section 3.3. The analysis based on the high-order Lyman fit is presented in Section 3.4. Column density distribution and evolution of the Ly α forest containing C IV are found in Section 3.5. Finally, we discuss and summarise the main results in Section 3.6. All the results on the number density and the differential column density distribution from our analysis are tabulated in Section 3.7. Throughout this study, the cosmological parameters are assumed to be the matter density $\Omega_m = 0.3$, the cosmological constant $\Omega_\Lambda = 0.7$ and the current Hubble constant $H_o = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with

3. The IGM H I column density evolution

$h = 0.7$. The logarithm N_{HI} is defined as $\log N_{\text{HI}} = \log(N_{\text{HI}}/1 \text{ cm}^{-2})$.

3.2. Data and Voigt profile fitting

Table 3.1 lists the properties of the 18 high-redshift QSOs analysed in this study. The spectrum of Q1101–264 is the same one as analysed in Kim et al. (2002), while the rest of spectra are the ones from Kim et al. (2007). The raw spectra were obtained from the ESO VLT/UVES archive and were reduced with the UVES pipeline. All of these spectra have a resolution of $R \sim 45\,000$ and heliocentric, vacuum-corrected wavelengths. Readers can refer to Kim et al. (2004) and Kim et al. (2007) for the details of the data reduction. To avoid the proximity effect, the region of $4,000 \text{ km s}^{-1}$ blueward of the QSO’s Ly α emission was excluded.

In order to obtain the absorption line parameters (the redshift z , the column density N in cm^{-2} and the Doppler parameter b in km s^{-1}), we have performed a Voigt profile fitting analysis using VPFIT². The details of its procedure and caveats can be found in Carswell, Schaye & Kim (2002) and Kim et al. (2007). Here, we only give a brief description of the fitting procedure.

First, a localised initial continuum of each spectrum was defined using the CONTINUUM/E-CHELLE command in IRAF. Second, we searched for metal lines in the entire spectrum, starting from the longest wavelengths towards the shorter wavelengths. All the identified metal lines were the first to be fit, correctly fixing the same z and b values for the same ionic transitions. When metal lines were embedded in the H I forest regions, the H I absorption lines blended with metals were also included in the fit. Sometimes the simultaneous fitting of different transitions by the same ion reveals that the initial continuum needs to be adjusted to obtain acceptable ion ratios. In this case, we adjusted the initial continuum accordingly. The rest of the absorption features were assumed to be H I.

A typical $z \sim 3$ IGM absorption feature having $b \sim 30 \text{ km s}^{-1}$ starts to saturate around $N_{\text{HI}} \geq 10^{14.5} \text{ cm}^{-2}$ at the UVES resolution and S/N. Unfortunately, N_{HI} and b values of saturated lines are not well constrained. In order to derive reliable N_{HI} and b values, higher-order Lyman series, such as Ly β and Ly γ , have to be included in the fit, as higher-order Lyman series have smaller oscillator strengths and start to saturate at much larger N_{HI} . Obtaining reliable N_{HI} is important for a study like ours, which deals with saturated lines.

After fitting metal lines, we have fitted the entire spectrum including all the available higher-order Lyman series. During this process, a small amount of the continuum re-adjustment was often required to achieve a satisfactory fit. With this re-adjusted continuum, we re-fitted the entire spectrum. This iteration process of continuum re-adjustments and re-fitting was then repeated several times until satisfactory fitting parameters were obtained. This produces the final set of fitted parameters for each component of the *high-order Lyman fit* analysis.

In addition to this high-order fit, we have also performed the same analysis using only the Ly α transition region, i.e. the wavelength range above the rest-frame Ly β and below the proximity effect zone. This additional fitting analysis was done, since most previous

²Carswell et al.: <http://www.ast.cam.ac.uk/~rfc/vpfit.html>

Table 3.1.: Analysed QSOs

QSO	z_{em}^a	$z_{\text{Ly}\alpha}^b$	$\lambda_{\text{Ly}\alpha}$ (Å)	$z_{\text{Ly}\alpha}^c$	S/N	Lyman limit ^d	notes ^e
Q0055-269	3.655 ^f	2.936-3.605	4785-5598		80-50	2288	
PKS2126-158	3.279	2.815-3.205	4638-5112		50-200	3457	two sub-DLAs at $z=2.768$ & 2.638
Q0420-388	3.116 ^f	2.480-3.038	4231-4909	2.665-3.038	100-140	3754	a sub-DLA at $z=3.087$
HE0940-1050	3.078	2.452-3.006	4197-4870		50-130	≤ 3200	
HE2347-4342 ^g	2.874 ^f	2.336-2.819	4055-4643		100-160	≤ 1160	multiple associated systems
Q0002-422	2.767	2.209-2.705	3901-4504		60-70	3025	
PKS0329-255	2.704 ^f	2.138-2.651	3815-4439		40-60	3157	associated system at 4513.7 Å
Q0453-423	2.658 ^f	2.359-2.588	4084-4362		90-100	3022	a sub-DLA at $z=2.305$
		2.091-2.217	3758-3911				
HE1347-2457	2.609 ^f	2.048-2.553	3705-4319		85-100	2237	
Q0329-385	2.434	1.902-2.377	3528-4105		50-55	≤ 3050	
HE2217-2818	2.413	1.886-2.365	3509-4091	1.971-2.365	65-120	2471	
Q0109-3518	2.405	1.905-2.348	3532-4070	1.974-2.348	60-80	2163	
HE1122-1648	2.404	1.891-2.358	3514-4082	1.974-2.358	70-170	≤ 1629	
J2233-606	2.250	1.756-2.201	3335-3891	1.980-2.201	30-50	≤ 1750	
PKS0237-23	2.223 ^f	1.765-2.179	3361-3865	1.974-2.179	75-110	≤ 3050	a sub-DLA at $z=1.673$
PKS1448-232 ^h	2.219	1.719-2.175	3306-3860	1.974-2.175	30-90	≤ 3050	
Q0122-380	2.193	1.700-2.141	3282-3819	1.977-2.141	30-80	≤ 3052	
Q1101-264	2.141	1.880-2.097	3503-3765	1.970-2.097	80-110	2597	a sub-DLA at $z=1.839$

^a The redshift is measured from the observed Ly α emission line of the QSOs. The redshift based on the emission lines is known to be under-estimated compared to the one measured from the absorption lines of the host galaxies (Tytler & Fan 1992; Vanden Berk et al. 2001).

^b The redshift range of the Ly α forest region analysed for the Ly α -only fit.

^c The redshift range of the Ly α forest region analysed for the high-order Lyman fit. The region is listed only when it is different from the Ly α -only fit region.

^d The wavelength of a Lyman limit of each QSO is defined as the wavelength below which the observed flux becomes 0. The values are taken from Kim et al. (2004). When a Lyman limit is not detected within available data, it is denoted to be less than the lowest available wavelength.

^e When a sub-damped Ly α system exists along the sightline, we discard at least 50Å centred at the sub-DLA each side to exclude its influence on the forest, such as a higher frequency or lack of higher-column density forest.

^f Due to the absorption systems at the peak of the Ly α emission line or to the non-single-peak emission line, the measurement is uncertain.

^g There are very strong O VI absorptions mixed with the two saturated Ly α absorption systems at 4012-4052Å (Fechner et al. 2004). Since the fitted line parameters for these Ly α systems cannot be well constrained (their corresponding Ly β is below the partial Lyman limit produced by the $z \sim 2.738$ systems), we discarded this wavelength region.

^h Part of the continuum uncertainties towards shorter wavelengths are due to the local, non-smooth continuum shape, partly due to a lower S/N ($\sim 30-35$).

3. The IGM H I column density evolution

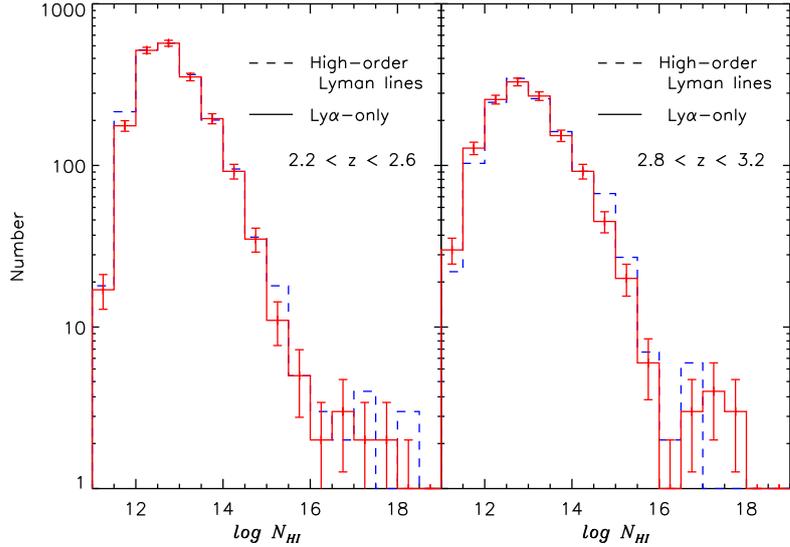


Figure 3.1.: Numbers of absorption lines as a function of N_{HI} at $2.2 < z < 2.6$ and $2.8 < z < 3.2$. The Ly α -only fits are shown as solid lines, while the high-order Lyman fits are marked as dashed lines. Solid errors indicate the 1σ Poisson errors of the Ly α -only fits.

studies on the IGM N_{HI} analysis based on Voigt profile fitting utilised only the Ly α region. For the Ly α -only fit, we kept the same continuum used in the high-order Lyman fitting process. In principle, the difference between two sets of fitted parameters occurs only in the regions where saturated absorption features are included.

In both fitting analyses, we did not tie the fitting parameters, except for the same ion transitions at the same redshift. Fig. 3.1 shows the numbers of absorption lines as a function of N_{HI} for both fitting analyses at the two redshift ranges, $2.2 < z < 2.6$ and $2.8 < z < 3.2$, in order to illustrate the differences at high and low redshifts. The difference between the two samples occurs mostly at $N_{\text{HI}} \geq 10^{14.5} \text{ cm}^{-2}$ and at $N_{\text{HI}} \leq 10^{12} \text{ cm}^{-2}$. This difference in the line number at $N_{\text{HI}} \geq 10^{14.5} \text{ cm}^{-2}$ seems to be stronger at $2.8 < z < 3.2$, although it is still within 2σ Poisson errors. On the other hand, the line number at $N_{\text{HI}} \leq 10^{12} \text{ cm}^{-2}$ is more susceptible to the incompleteness than the difference between the two fitting methods. In fact, the incompleteness becomes severe as N_{HI} is smaller than $N_{\text{HI}} \leq 10^{13} \text{ cm}^{-2}$ and z is higher than $z > 3$ (Kim et al. 1997, 2002). The difference at other column density ranges is smaller, which in turn leads to expect that there is no significant difference between the Ly α -only fit and the high-order Lyman fit. Therefore, we restrict our present analysis to $N_{\text{HI}} \geq 10^{12.75} \text{ cm}^{-2}$ at all redshifts. As we want to have a sample with a reliable N_{HI} , we restrict our analysis to $z > 1.9$. In addition, regions within $\pm 50\text{\AA}$ to the centre of a sub-damped Ly α system ($N_{\text{HI}} \geq 10^{19} \text{ cm}^{-2}$) are excluded, since they are associated with intervening high- z galactic disks.

Note that the availability of the high-order Lyman series depends on the redshift of the QSO and whether the sightline contains a Lyman limit at the rest-frame 912\AA . In addition, the amount of blending affects whether a reliable column density can be measured. At high redshifts $z_{\text{em}} > 3$, line blending becomes severe. However, most UVES spectra also covers down to 3050\AA where Lyman lines higher than Ly δ are available in most

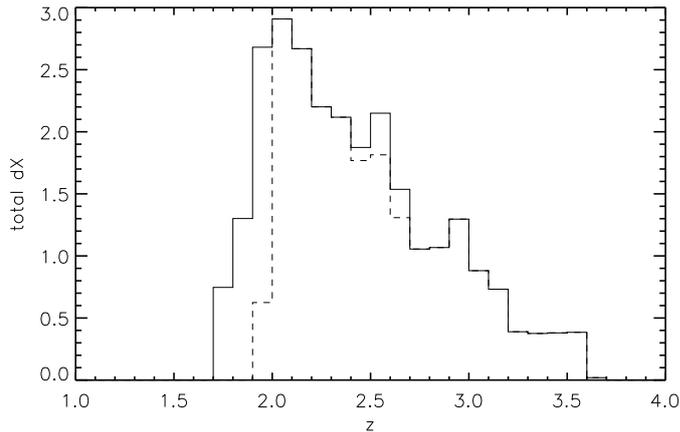


Figure 3.2.: Total redshift path dX covered with our sample of 18 high-redshift QSOs. The solid line is for the $\text{Ly}\alpha$ -only fit, while the dashed line for the high-order fit.

cases. On the other hand, at $z_{\text{em}} < 2.5$ the available high-order Lyman lines are rather limited, with mostly $\text{Ly}\beta$ and $\text{Ly}\gamma$ available. However, line blending is less severe than at higher redshifts. We generated tens of saturated artificial absorption lines and fitted them including and excluding high-order Lyman lines. Our practice shows that *unblended* absorption features at $N_{\text{HI}} \leq 10^{17} \text{ cm}^{-2}$ can be reasonably well constrained with $\text{Ly}\alpha$ and $\text{Ly}\beta$ only. This indicates that our N_{HI} can be also considered reliable even at $z < 2.5$ with $\text{Ly}\alpha$ and $\text{Ly}\beta$ only.

The total redshift path dX covered by the spectra for both $\text{Ly}\alpha$ -only and high-order fits is shown in Fig. 3.2. The redshift path is obtained by integrating the Friedmann equation for a $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ universe, and is given by

$$dX = \int \frac{H_0}{H(z)} (1+z)^2 dz \quad (3.1)$$

(Bahcall & Peebles 1969), where H_0 is the Hubble constant at $z = 0$. The redshift coverage of our sample steadily increases with decreasing redshift until it reaches its maximum at $z \sim 2.1$. For redshifts below $z < 1.9$ the coverage decreases rapidly and our sample ends at $z = 1.7$. Note that the lowest redshift possible for the high-order Lyman fit analysis is $z \sim 1.97$, while the $\text{Ly}\alpha$ -only fit analysis can be possible down to $z \sim 1.7$. The sample coverage of the fit analyses only differ between $2.4 < z < 2.7$, where the coverage of the high-order fit is decreased due to the intervening Lyman limit. At the high redshift end $3.2 < z < 3.6$, the sample only consists of one line of sight. We therefore restrict the sample to redshifts below $z < 3.2$.

In Fig. 3.3 the number of HI absorbers from both fitting methods described above is shown as a function of redshift. The number of absorbers obtained from each fitting analysis is roughly proportional to the redshift path coverage. Therefore, our sample shows the highest HI absorber number around redshift $z \sim 2$ for each fitting analysis, where the sample redshift path coverage also reaches its maximum. The high-order fit analysis reveals a slightly higher number of absorbers between $2 < z < 3$. This is mainly due to the breaking up of some saturated absorbers into multiple components in the

3. The IGM HI column density evolution

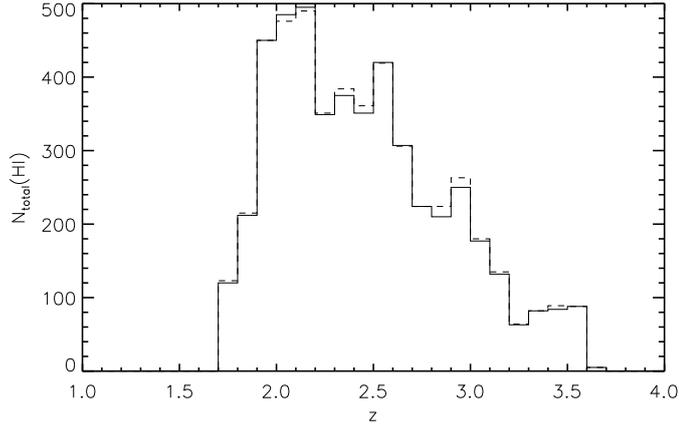


Figure 3.3.: Number of HI absorbers as function of redshift in our our sample of 18 high-redshift QSOs. The solid line is for the Ly α -only fit, while the dashed line for the high-order fit.

high-order fit analysis.

3.3. Comparison with previous studies using Ly α only

In Sect. 3.2 we have shown that including higher order transitions in the fitting process slightly alters the column density statistics at column densities above $\log N_{\text{HI}} \sim 15.0$. In order to compare our QSO sample with previous studies based only on the Ly α transition, we briefly present the column density distribution and evolution derived from the Ly α -only fit in this section. A detailed determination of the differential column density distribution using the high-order fit is presented in Section 3.4.2. All the results from these sections are tabulated in the Section 3.7.

3.3.1. Absorber number density evolution dn/dz

The absorber number density $n(z)$ is measured by counting the number of HI absorption lines for a given column density range and a covered redshift range for each line of sight. The line count n is then divided by the covered redshift range Δz to obtain dn/dz . If forest absorbers have a constant size and a constant comoving number density, its number density evolution due to the Hubble expansion can be described as

$$\frac{dn}{dz} = \pi R^2 N_0 c H(z)^{-1} (1+z)^2, \quad (3.2)$$

where R is the size of an absorber, N_0 is the local comoving number density and c is the speed of light (Bahcall & Peebles 1969). For our assumed cosmology, Eq. 3.2 becomes

$$\frac{dn}{dz} \propto \frac{(1+z)^2}{\sqrt{0.3(1+z)^3 + 0.7}}. \quad (3.3)$$

At $1 < z < 4.5$, Eq. 3.3 has an asymptotic behaviour of $dn/dz \propto (1+z)^{\sim 0.6}$, while at $z < 1$ it becomes $dn/dz \propto (1+z)^{\sim 1.15}$. For higher redshifts the asymptotic behaviour

3.3. Comparison with previous studies using Ly α only

becomes $dn/dz \propto (1+z)^{0.5}$. Any differences in the observed exponent from what is expected from Eq. 3.3 indicate that the absorber size or/and the comoving density are not constant.

Empirically, dn/dz is described as $dn/dz = A(1+z)^\gamma$. It has been known that dn/dz evolves more rapidly at higher column densities. At $z > 1.5$, a $\gamma \sim 2.9$ is found for $N_{\text{HI}} = 10^{14-17} \text{ cm}^{-2}$, and $\gamma \sim 1.4$ for $N_{\text{HI}} = 10^{13.1-17} \text{ cm}^{-2}$ (Kim et al. 2002). At $z < 1.5$, Weymann et al. (1998) found $\gamma \sim 0.16$ and $A \sim 35$ for absorbers with an equivalent width greater than 0.24\AA . Later studies on dn/dz which based on the profile fitting using better-quality data show a factor of $\sim 2-3$ lower dn/dz than the one found by Weymann et al. (1998). These studies also shows a larger scatter in dn/dz at $z < 0.2$ with $A \sim 5 - 22$ (Lehner et al. 2007; Williger et al. 2010). Unfortunately these high-quality data lack a complete z coverage at $z < 1.5$, missing mostly at $0.4 < z < 1.0$. Therefore, it is not possible to derive a robust power-law slope γ at low redshift and its evolution at $0 < z < 3.5$.

The number density evolution is illustrated in Figs. 3.4, 3.5 and 3.6 for three different column density ranges: $\log N_{\text{HI}} = [14, 17]$, $[14.5, 17]$, and $[13.1, 14]$. The redshift range used in Fig. 3.4, 3.5 and 3.6 is given in Table 3.1 as $z_{\text{Ly}\alpha}$. Data compiled from the literature are indicated in the figures: Williger et al. (2010), Lehner et al. (2007), Aracil et al. (2006), Janknecht et al. (2006), Williger et al. (2006), Richter et al. (2004), Sembach et al. (2004), Penton et al. (2000), Savaglio et al. (1999), Weymann et al. (1998), Kim et al. (1997), Kirkman & Tytler (1997), Lu et al. (1996), and Hu et al. (1995).

We performed a linear regression to our data in logarithmic space for the various column density bins, using the maximum likelihood method described in Ripley & Thompson (1987). Their method allows to simultaneously account for the differing line of sight redshift coverage and the uncertainties in the number density. Linear regressions were once obtained from our data including the literature data and once without them. Since for redshifts $z \lesssim 1$ (or $\log(1+z) \lesssim 0.3$) the number density evolution *could* remain constant with redshift (Weymann et al. 1998), only the data with redshift $z > 1$ is used for the fit. The resulting parameters are given in Table 3.3.1. An uncertainty estimate of the fit parameters was obtained using the jackknife method.

Fig. 3.4 shows the dn/dz evolution for the column density interval of $\log N_{\text{HI}} = [14, 17]$. Our results agree well with previous findings at $z > 1.5$ ($\log(1+z) > 0.4$), confirming that there is a real sightline variation in dn/dz . The linear regression to only our results (the solid line) with $\gamma = 4.08 \pm 0.76$ deviates strongly from the fit to all the available data at $z > 1$ ($\log(1+z) > 0.3$) which yields $\gamma = 2.99 \pm 0.26$ (the dashed line). This discrepancy is due to the sparse data of our sample at higher redshift $z > 3.5$ ($\log(1+z) > 0.65$) and the missing constraints at $z < 2.0$. However, the fit to all the available data for redshifts $z > 1$ including the literature values is consistent with the findings by Kim et al. (2002) who found $\gamma = 2.90 \pm 0.25$.

In Fig. 3.5 we present results for the column density interval for stronger absorbers $\log N_{\text{HI}} = [14.5, 17]$. The evolution at $z < 2.5$ continues to follow the empirical power-law, although the scatter increases. There is no significant difference between the power law slopes of $\gamma = 3.55 \pm 0.83$ for our data only, of $\gamma = 3.72 \pm 0.52$ for the entire data set including the literature, of $\gamma = 3.11 \pm 0.42$ found by Kim et al. (2002). Kim et al. (2002) discuss a possibility on whether the column density evolution flattens out at $z < 2.5$ ($\log(1+z) < 0.55$) for this column density interval. Even though more data points are

3. The IGM H I column density evolution

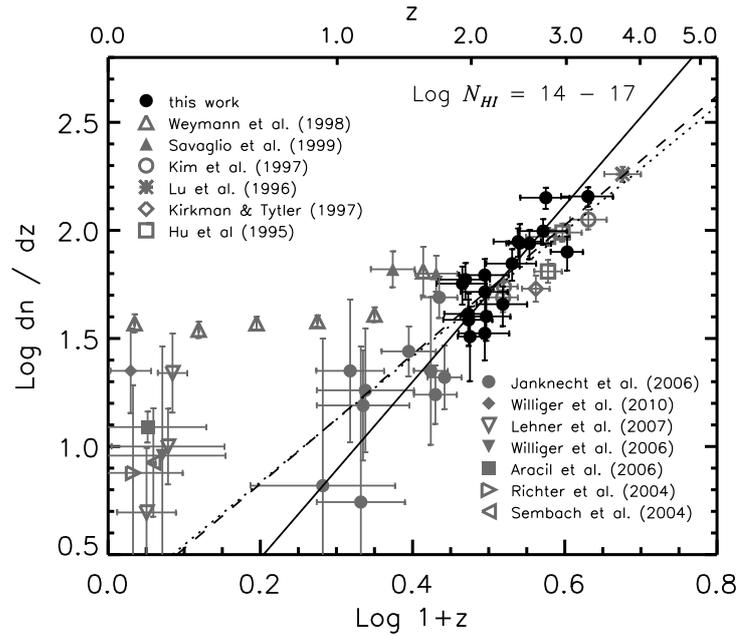


Figure 3.4.: The number density evolution of the Ly α forest in the column density range $\log N_{\text{HI}} = [14, 17]$ of the Ly α -only fits. Black filled circles show results from our data set. Grey data points indicate various results obtained from the literature. The solid line show the fit to our data only, while the dashed line are the results including the literature data for $z > 1$ ($\log(1+z) > 0.3$). The dotted line gives the fit given in Kim et al. (2002). The vertical error bars give the Poisson 1σ error.

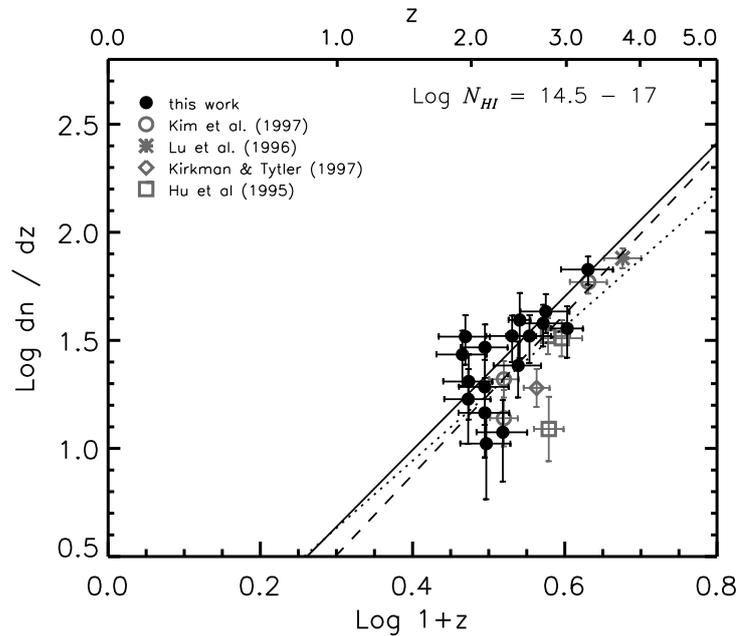


Figure 3.5.: The number density evolution of the Ly α forest in the column density range $\log N_{\text{HI}} = [14.5, 17]$. All the symbols have the same meaning as in Fig. 3.4.

3.3. Comparison with previous studies using Ly α only

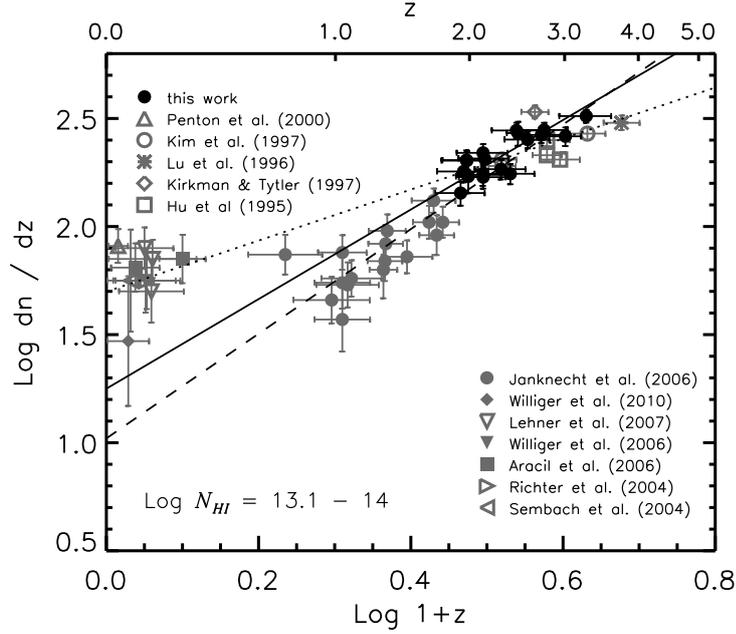


Figure 3.6.: The number density evolution of the Ly α forest in the column density range $\log N_{\text{HI}} = [13.1, 14]$. All the symbols have the same meaning as in Fig. 3.4.

available in this study, this question cannot be conclusively answered and more data covering lower redshifts are required. The scatter is similar to the lower column density interval shown in Fig. 3.4, where an increase in the scatter is seen in the data of Janknecht et al. (2006) at redshifts below $z \sim 2.0$ ($\log(1+z) \sim 0.45$). This is also the region in which the flattening out of the dn/dz evolution is seen. All this indicates that higher column density systems evolve more rapidly than low column density systems and its number density decreases faster with decreasing redshifts. The increase in the scatter at redshifts $z < 2.5$ might indicate the transition point where the evolving number density changes into a non-evolving one, as is predicted in numerical simulations by Theuns et al. (1998a); Davé et al. (1999).

The line number density evolution for low column density systems in the range of $\log N_{\text{HI}} = [13.1, 14]$ is given in Fig. 3.6. Again the results from our sample agree well with previous results found in the literature. The linear regression to our data gives comparable results to the fit including all available data points. However, these results do not compare well with the linear regression obtained by Kim et al. (2002), which show with $\gamma = 1.18 \pm 0.14$, a more shallow evolution of the line number density. This discrepancy arises due to line blending of low column density systems with denser systems at high redshifts and a lack of any data points at $z < 1.5$ in their sample. Line blending at high redshifts makes the detection of weak systems in the Voigt profile analysis difficult. This effect has been shown to underestimate the line number density of low column density systems by as much as $\sim 30\%$ at $z \sim 3$ ($\log(1+z) \sim 0.6$) and above (Giallongo et al. 1996). In addition, the robust estimate of the exponent γ requires a large z range. Note that the exponent $\gamma = 1.18 \pm 0.14$ found by Kim et al. (2002) matches well with the data at $z < 0.3$, although they did not include any data at $z < 1.5$. Based on the

3. The IGM H I column density evolution

Table 3.2.: Table of linear regression results for the number density evolution dn/dz using the Ly α -only fits. Given are fits to our data only and fits including literature data for $z > 1$.

$\Delta \log N_{\text{HI}}$	QSO by QSO		QSO by QSO with lit.	
	$\log(dn/dz)_0$	γ	$\log(dn/dz)_0$	γ
13.64 – 17.0	0.53 ± 0.29	3.01 ± 0.55	0.69 ± 0.17	2.65 ± 0.31
14.0 – 17.0	-0.33 ± 0.40	4.08 ± 0.76	0.23 ± 0.14	2.99 ± 0.26
14.5 – 17.0	-0.43 ± 0.44	3.55 ± 0.83	-0.61 ± 0.29	3.72 ± 0.52
13.1 – 14.0	1.25 ± 0.21	2.07 ± 0.39	1.02 ± 0.08	2.42 ± 0.15
12.75 – 14.0	1.69 ± 0.17	1.70 ± 0.31		

data points at $0.5 < z < 2.0$ by Janknecht et al. (2006), dn/dz could be considered to flatten out at $z < 1$ ($\log(1+z) \sim 0.3$)³.

3.3.2. The differential column density distribution function

The differential column density distribution function is defined as the number of absorbers per unit redshift path $X(z)$ and per unit column density N_{HI} . The redshift path is calculated using Eq. 3.1. The differential column density distribution is empirically described as a power-law,

$$\frac{dN}{dN_{\text{HI}} dX} = \left(\frac{dN}{dN_{\text{HI}} dX} \right)_0 N_{\text{HI}}^\beta, \quad (3.4)$$

where $(dN/(dN_{\text{HI}} dX))_0$ gives the normalisation point of the distribution function and β denotes its slope. This power law behaviour has been shown to persist over more than ten orders of magnitudes in column density (Petitjean et al. 1993).

In Fig. 3.7 we present the results using the Ly α -only fits for the entire sample redshift range. The top x-axis is in units of the gas overdensity which was computed according to Eq. 10 by Schaye (2001)

$$N_{\text{HI}} \sim 2.7 \times 10^{13} (1 + \delta)^{3/2} T_4^{-0.26} \Gamma_{12}^{-1} \left(\frac{1+z}{4} \right)^{9/2} \left(\frac{\Omega_b h^2}{0.02} \right)^{3/2} \left(\frac{f_g}{0.16} \right)^{1/2} \text{cm}^{-2}. \quad (3.5)$$

The gas temperature T is assumed to be $T = T_4 \times 10^4$ K, the photoionisation rate $\Gamma = \Gamma_{12} \times 10^{-12} \text{sec}^{-1}$. The parameter f_g denotes the fraction of mass in gas. The IGM gas temperature is assumed to be governed by the effective equation of state $T = T_0(1 + \delta)^{\gamma-1}$ (Hui & Gnedin 1997). For Γ_{12} and γ we interpolated results obtained by Bolton et al. (2008). We assumed that T_0 is 2×10^4 K, $f_g = 0.16$, and $\Omega_b h^2 = 0.0227$ (Schaye 2001).

Supplementary to our results, we include observations by Storrie-Lombardi & Wolfe (2000) and O’Meara et al. (2007) for increased column density coverage. Further we

³The fitted line parameters by Janknecht et al. (2006) show many lines with $b < 20 \text{ km s}^{-1}$, about 25% of all lines. They attribute this to their low signal-to-noise data of less than 10. Note that N_{HI} obtained from the profile fitting analysis for low signal-to-noise data is not reliable.

compare our results with the ones by Petitjean et al. (1993) and Hu et al. (1995). Our results are in good agreement with the Petitjean et al. (1993) data over the whole column density range down to $\log N_{\text{HI}} = 12.75 \text{ cm}^{-2}$. At smaller column densities, our results start to deviate from a power law due to the sample incompleteness for weak systems (Kim et al. 1997).

We have fitted the differential column density distribution with a power law in the range of $\log N_{\text{HI}} = [12.75, 14] \text{ cm}^{-2}$, since the data slightly deviates from the power law form at higher column densities. This behaviour has previously been observed by Petitjean et al. (1993) and Kim et al. (1997). We will address this issue in more detail in the next section using results from the high-order fit sample. From the linear regression we find $\log(dN/(dN_{\text{HI}} dX))_0 = 7.23 \pm 0.23$ and a slope of $\beta = -1.44 \pm 0.02$. These results are slightly lower than the findings by Hu et al. (1995) who found $\beta = -1.46$ and the ones by Petitjean et al. (1993) obtaining $\beta = -1.49 \pm 0.02$.

3.4. Analysis using higher-order transition lines

In the last section, we checked the Ly α absorber number density evolution and the differential column density distribution obtained from the Ly α -only fits for consistency with previous studies. The analysis is now revisited with the results from the Voigt profile analysis including the higher order transitions, hence a sample with a reliable N_{HI} . As seen in Fig. 3.1, better constraints on the column density can be obtained with the inclusion of high-order Lyman lines in the fitting process. Therefore it can be established whether the dip seen in the differential column density distribution is a physical feature or just an imprint of uncertainties in N_{HI} . All the results from this section are tabulated in the Section 3.7.

3.4.1. The mean number density evolution

We now revisit the line number density evolution using the high-order Lyman sample, as described in the previous section. On a QSO by QSO basis we determine dn/dz for a low column density range of $\log N_{\text{HI}} = [12.75, 14]$ and for high column densities of $\log N_{\text{HI}} = [14, 17]$. The lower column density range is chosen in such a way that the part of the differential column density distribution function which follows a power-law is covered, whereas the $\log N_{\text{HI}} = [14, 17]$ interval covers those systems responsible for the dip in $dN/(dN_{\text{HI}} dX)$.

The results are presented in Fig. 3.8. Linear regressions from the data are obtained and the resulting parameters are summarised in Table 3.3. Similar to the previous analysis, the line number density shows a decrease with decreasing redshift. No significant differences between the two fits of the Ly α -only and the higher-order data are visible. In the case of the $\log N_{\text{HI}} = [14, 17]$ interval, the slope of the power law steepens from the Ly α -only slope of $\gamma = 4.08 \pm 0.76$ to $\gamma = 4.68 \pm 0.75$ for the high-order fit. However, the slopes of the two samples are still in the 1σ uncertainty range, rendering the two results consistent to each other. Similar results are obtained for the $\log N_{\text{HI}} = [12.75, 14]$ range. The slope decreased from $\gamma = 1.70 \pm 0.31$ to $\gamma = 1.48 \pm 0.29$ for the high-order fit. Again the results from the two samples agree within the 1σ uncertainty range.

3. The IGM H I column density evolution

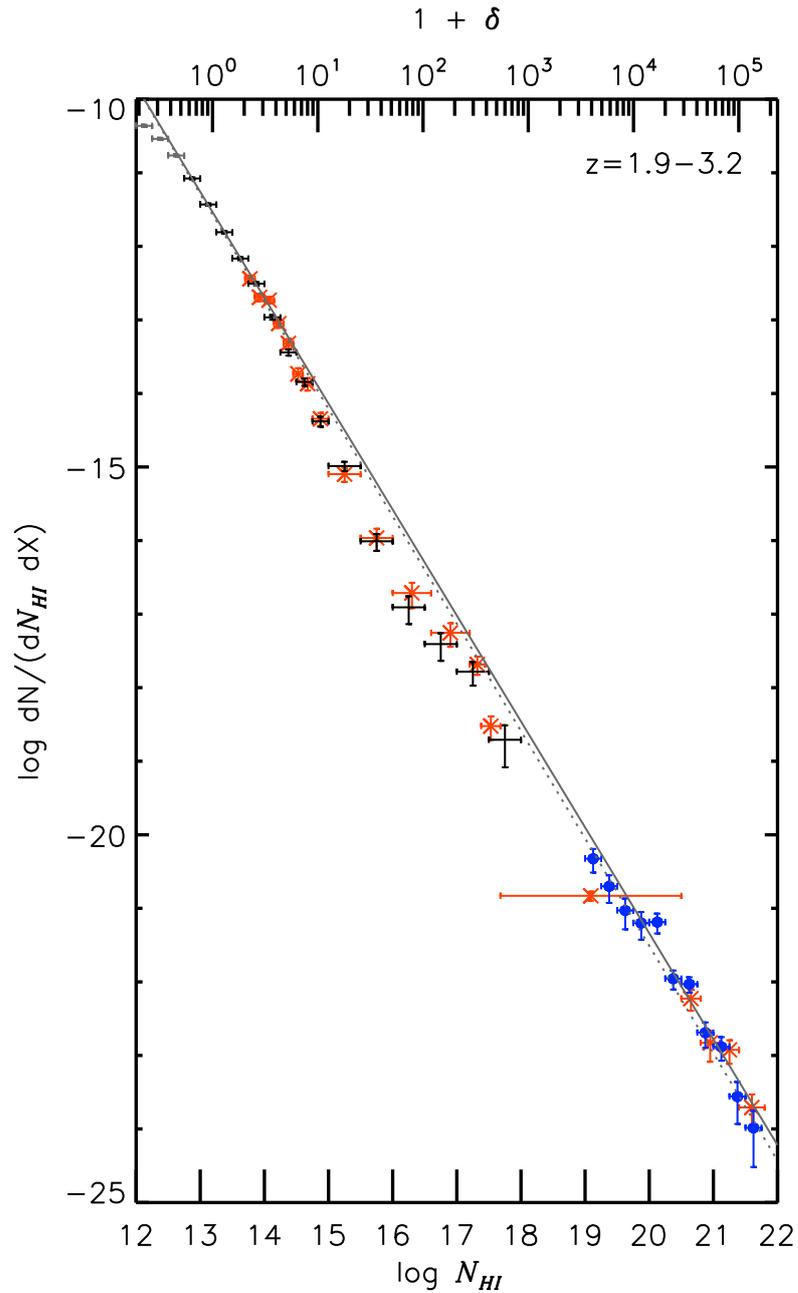


Figure 3.7.: The differential column density distribution for the entire redshift range of our QSO sample using the Ly α -only fits. Both black and grey data points show the results from our QSO sample. The grey data points at $\log N_{\text{HI}} < 12.75 \text{ cm}^{-2}$ mark the column densities that are affected by incompleteness. The stars are the data points obtained by Petitjean et al. (1993). The circles indicate data points derived from data given by Storrie-Lombardi & Wolfe (2000) and O’Meara et al. (2007). The solid line gives the power law fit for $\log N_{\text{HI}} = [12.75, 14] \text{ cm}^{-2}$. The dotted line represent the fit obtained by Hu et al. (1995). The vertical error bars indicate 1σ Poisson errors. Gas overdensities are computed using Eq. 10 by Schaye (2001) (see text for details).

Up to now, the number density evolution has been only derived on a QSO by QSO basis. Previous studies did not have enough QSO sight lines available to sample the number density evolution at smaller redshift interval Δz , without suffering from small number statistics. However, our sample of 18 high-redshift QSOs is characterised by a large redshift distance coverage in the redshift range of $1.9 < z < 3.2$ (see Fig. 3.2). As a result, a large number of absorption lines is available for small redshift intervals to combine the individual QSO line lists into one big sample. From this combined sample, the evolution of the *mean* number density is derived in redshift bins of $\Delta z = 0.1$.

Results of the combined line number density evolution are shown in Fig. 3.9 for identical column density ranges as used in the QSO by QSO analysis. Error bars have been determined using the bootstrap technique. For comparison, results using the Ly α -only fits are overplotted for the high column density bin. For the low column density range, the results from the two samples are undistinguishable. The high column density results are similar to the ones obtained from the Ly α -only fits, with a slight difference at the low redshift end. The number density is higher in the high-order fits, since some saturated systems break up into multiple components in the high-order Lyman transition. This increases the absorber number density. However the differences between the two samples are smaller than the statistical uncertainties. Only at $z \sim 2.9$ ($\log(1+z) \sim 0.59$) the difference between the two fitting methods is in the order of 2σ . At this redshift bin, three Lyman limit systems with a column density above $\log N_{\text{HI}} > 17$ are found in the Ly α -only fit sample (two in HE0940–1050 and one in Q0420–388). These systems therefore do not contribute to the Ly α -only results. However, these Lyman limit systems break up into multiple smaller components with the high-order fit and suddenly contribute substantially to the number count.

Again, linear regressions have been obtained and its parameters are given in Table 3.3. With the combined analysis the scatter in the number density evolution is reduced and we obtain smoother relations than with the QSO by QSO case. This combined analysis produces dn/dz closer to the *mean* number density evolution than the QSO by QSO analysis. The power law nature of the evolution can thus be more clearly seen.

At low column densities, no differences between the two samples are observed, as expected. The slope of the combined sample is with $\gamma = 1.28 \pm 0.20$ somewhat shallower than for the QSO by QSO analysis. On the other hand, for the high column density sample, the differences in the slopes are larger. The slope of the combined sample decreases to $\gamma = 3.77 \pm 0.49$, which is more consistent with the slope derived in the previous section for the combined literature results of $\gamma = 2.99 \pm 0.26$.

For the high column density sample, a significant deviation from a power law by more than 2σ is seen at $z = 2.2$ ($\log(1+z) = 0.5$). For higher redshifts, the evolution shows the smooth power law form. At $z \sim 2.3$ – 2.4 ($\log(1+z) \sim 0.52$ – 0.53), the evolution starts to deviate from a power law and the number density decreases in a dip down to $z \sim 2.2$ ($\log(1+z) = 0.5$). At $z < 2$ ($\log(1+z) < 0.48$), the number density increases again and the evolution seems to regain its power law form. This feature is unlikely a result of small number statistics, since our sample coverage is at its largest in the redshift range of the dip. However, keep in mind that this dip happens only at the edge of our sample redshift coverage at $z \sim 2.15$ ($\log(1+z) \sim 0.49$), where the sightline variation is fairly large. It might also be, that the power-law form is not regained for redshifts below $z < 2$ and that the strong decline in the number density only marks the transition into the regime

3. The IGM H I column density evolution

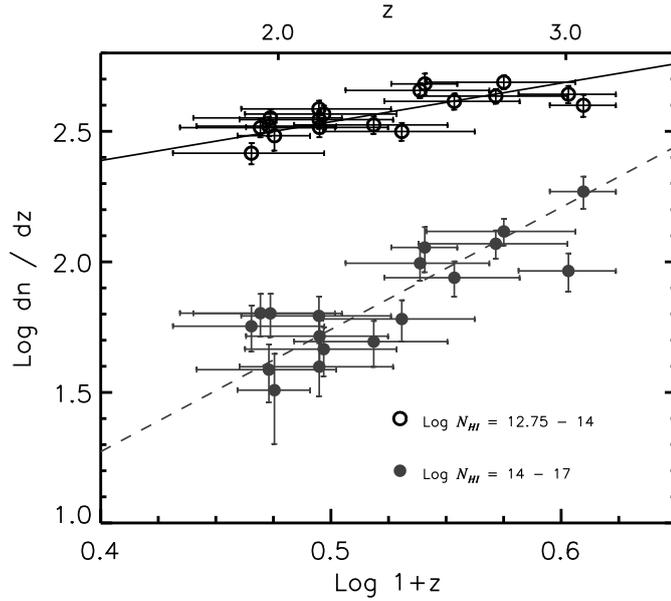


Figure 3.8.: The line number density evolution derived on a QSO by QSO basis using the high-order Lyman sample for column density intervals of $\log N_{\text{HI}} = [12.75, 14]$ and $\log N_{\text{HI}} = [14, 17]$. The vertical error bars indicate 1σ Poisson errors. The straight lines denote results from a linear regression to the data with parameters given in Table 3.3.

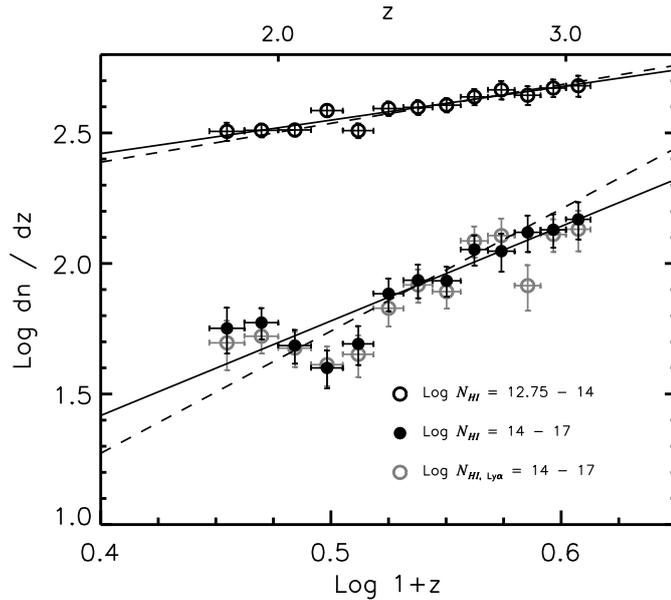


Figure 3.9.: The line number density evolution of the combined sample as a function of redshift using the high-order Lyman-series sample for column density intervals of $\log N_{\text{HI}} = [12.75, 14]$ and $\log N_{\text{HI}} = [14, 17]$. The sample is binned in redshift with $\Delta z = 0.1$. For comparison reasons the results of the Ly α -only fits (grey open circles) are shown for the $\log N_{\text{HI}} = [14, 17]$ interval. The vertical error bars indicate 1σ bootstrap errors. The straight solid lines denote results from a linear regression to the binned data. The dashed line represent the fits of the uncombined sample shown in Fig. 3.8. The parameters of the fits are given in Table 3.3.

Table 3.3.: Table of linear regression results for the number density evolution dn/dz using the high-order Lyman-series dataset. Power law fits to the QSO by QSO data are given, as well as to the results from the combined sample.

$\Delta \log N$	QSO by QSO		Combined sample	
	$\log(dn/dz)_0$	γ	$\log(dn/dz)_0$	γ
12.75 – 14.0	1.79 ± 0.15	1.48 ± 0.29	1.91 ± 0.10	1.28 ± 0.20
14.0 – 17.0	-0.60 ± 0.41	4.68 ± 0.75	-0.21 ± 0.27	3.77 ± 0.49

where the number density does not evolve anymore. A similar dip is also present in the Ly α -only fits (grey open circles) and is thus not a result of the improved constraints from the high-order fits.

By carefully inspecting the lower column density sample at the redshift of the dip, we find the number density to slightly decrease at $z \sim 2.3$ ($\log(1+z) \sim 0.52$). It then slightly increases again at $z \sim 2.2$ ($\log(1+z) \sim 0.5$). The low column density sample regains its original form faster than in the high column density sample. The physical process responsible for the dip cannot be identified from the number density evolution itself. Whatever the process is, it has to substantially diminish the number of high column density systems at $z \sim 2.15$ ($\log(1+z) \sim 0.5$). These high column density systems have been found to originate in filaments and in the vicinity of galaxies (Miralda-Escudé et al. 1996). Thus it might be possible that an increase in UV photons around massive galaxies due to a peak in the star formation rate at $z \sim 2$ (Madau et al. 1998) is responsible for providing more ionising photons, thus decreasing the amount of neutral hydrogen in these systems. In fact, a similar trend is seen in Fig. 3 of Theuns et al. (1998a), although their simulation contains a UV background dominated by QSOs only and the result is due to the decrease of the UV ionising photons at $z < 2$. Therefore, the beginning of a change in the ionising background might result in a dip at $z \sim 2$.

3.4.2. The differential column density distribution function

Using the high-order fits, we have derived the differential column density distribution function for $1.9 < z < 3.2$ analogous to Section 3.3.2. In Fig. 3.10 we show the results for the entire redshift range. As with the Ly α -only fits, we include observations by Storrie-Lombardi & Wolfe (2000) and O’Meara et al. (2007). Our results show a power law relation which is identical to the results of the Ly α -only fits. As with the Ly α -only fits, the differential column density distribution function shows a deviation from the empirical power law for column densities between $15.0 < \log N_{\text{HI}} < 19.0$.

We fit a power law to the data in the column density range of $\log N_{\text{HI}} = [12.75, 14]$. The linear regression yields a normalisation point of $\log(dN/(dN_{\text{HI}} dX))_0 = 7.46 \pm 0.22$ and a slope of $\beta = -1.44 \pm 0.02$. This result is almost identical to the Ly α -only fit, since differences between the Ly α -only and the high-order fits only appear at $\log N_{\text{HI}} > 15$ (see Fig. 3.10). The high-order fits show a larger number of systems between $15 < \log N_{\text{HI}} < 17$ than the Ly α -only fits. However at higher column densities, the number of absorbers is lower for the higher-order fits than for the Ly α -only fits. This again indicates the breaking up of high column density systems into multiple smaller ones when looking at higher transitions than Ly α .

3. The IGM H I column density evolution

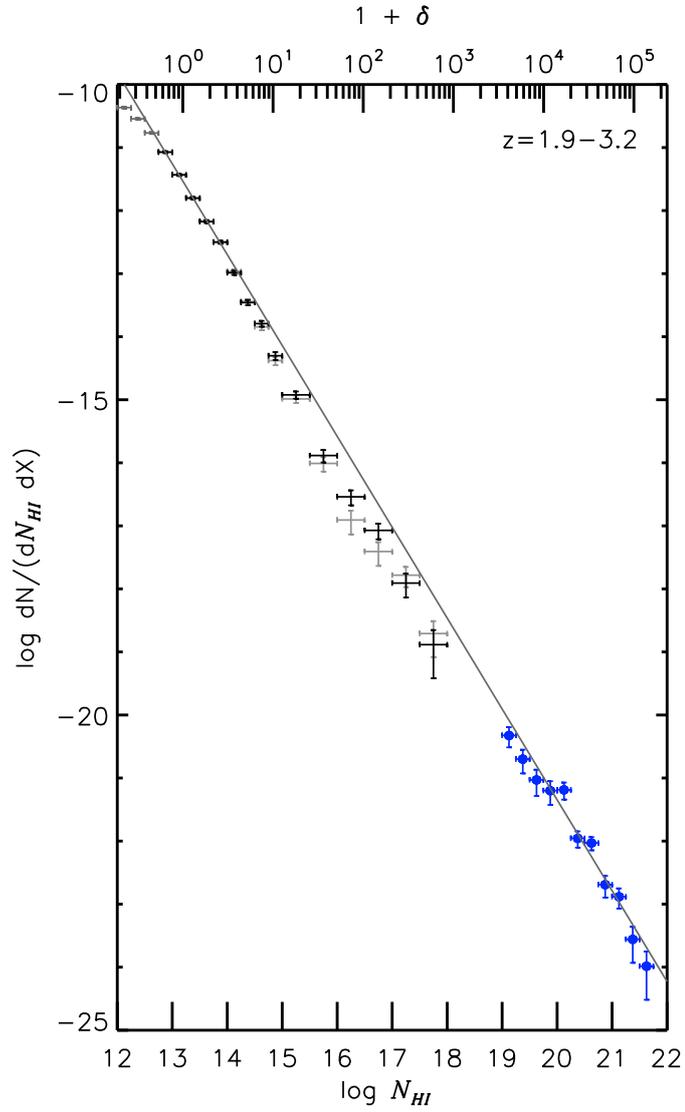


Figure 3.10.: The differential column density distribution for the entire redshift range of our QSO sample using the high-order Lyman-series dataset. Black and grey data points below show the results from our QSO sample. The grey data points below $\log N_{\text{H I}} < 12.75$ mark the column densities that are affected by incompleteness. The light grey data points above $\log N_{\text{H I}} > 12.75$ mark the results from the Ly α -only fit. The vertical error bars indicate 1σ Poisson errors. The circles indicate data points from the data given by Storrie-Lombardi & Wolfe (2000) and O’Meara et al. (2007). The solid line gives the power law fit for $\log N_{\text{H I}} = [12.75, 14]$ as given in the text.

3.4. Analysis using higher-order transition lines

Table 3.4.: Table of linear regression results for the differential column density distribution as a function of redshift and column density using the high-order Lyman dataset. Here the normalisation point $\log(dN/(dN_{\text{HI}}dX))_0$ is denoted by A .

z	$\log N_{\text{HI}} = 12.75 - 14.0$		$\log N_{\text{HI}} = 14.0 - 15.0$	
	A	β	A	β
1.9 – 2.2	8.14 ± 0.74	-1.49 ± 0.06	9.54 ± 2.72	-1.61 ± 0.19
2.2 – 2.7	7.35 ± 0.34	-1.43 ± 0.03	13.13 ± 2.26	-1.84 ± 0.16
2.7 – 3.2	6.85 ± 0.63	-1.39 ± 0.05	9.92 ± 2.08	-1.61 ± 0.15
1.9 – 3.2	7.48 ± 0.42	-1.44 ± 0.03	11.09 ± 1.13	-1.71 ± 0.08
z	$\log N_{\text{HI}} = 15.0 - 18.0$		$\log N_{\text{HI}} = 12.75 - 18.0$	
	A	β	A	β
1.9 – 2.2	7.31 ± 3.35	-1.47 ± 0.21	10.08 ± 0.45	-1.64 ± 0.03
2.2 – 2.7	7.85 ± 4.29	-1.51 ± 0.26	10.37 ± 0.70	-1.66 ± 0.05
2.7 – 3.2	8.11 ± 0.59	-1.50 ± 0.04	9.01 ± 0.48	-1.55 ± 0.03
1.9 – 3.2	8.10 ± 1.29	-1.51 ± 0.08	9.71 ± 0.28	-1.61 ± 0.02

In order to determine the redshift evolution of the differential column density distribution, we split the sample into three redshift bins: $z = [1.9, 2.2]$, $[2.2, 2.7]$, and $[2.7, 3.2]$. The lowest redshift interval covers the dip seen in the high column density sample of the dn/dz evolution. Since the column density distribution deviates from a single power law at around $\log N_{\text{HI}} = 14$, we have individually fitted power laws to four column density intervals of $\log N_{\text{HI}} = [12.75, 14]$, $[14, 15]$, $[15, 18]$, and $[12.75, 22]$, characterising the shape of the distribution function. The resulting parameters are listed in Table 3.4. Unfortunately the uncertainties in the fit parameters do not allow us to constrain the shape of the distribution and its dip as a function of redshift very well. Low uncertainties in the slopes can only be achieved with the results of the whole redshift range. From these results, however, it becomes evident that the lowest and highest column density ranges have similar slopes, and that the intermediate column density range is significantly steeper. In order to better constrain the evolution of the differential column density distribution in the various column density ranges, a much larger observational sample is needed.

The differential column density distributions as a function of redshift are presented in Fig. 3.11, where we overplot the fit of the $\log N_{\text{HI}} = [12.75, 14]$ interval for each redshift bin. Further we overplot the results of the fit to the entire redshift range $1.9 < z < 3.2$ as the dashed line. Comparing the slope of the linear relations shows, that the differential column density distribution becomes slightly steeper with decreasing redshift. In the highest redshift bin, the lowest column density range has a slope of $\gamma = -1.39 \pm 0.05$, which increases to $\gamma = -1.49 \pm 0.06$ for the lowest redshift bin. Further the scatter in the high column density regime decreases with decreasing redshift due to the larger sample size (compare Fig. 3.2). This is especially true for higher column densities. In order to thoroughly constrain the high column density part and to reduce its variance, the sample size needs to be further increased in future studies. However for low column densities of $\log N_{\text{HI}} < 15.0$, the differential column density can be reliably determined in all redshift bins.

Let us now focus on column densities above $\log N_{\text{HI}} > 14.0$. From Fig. 3.10 we have seen that the differential column density distribution deviates from the power law form

3. The IGM H I column density evolution

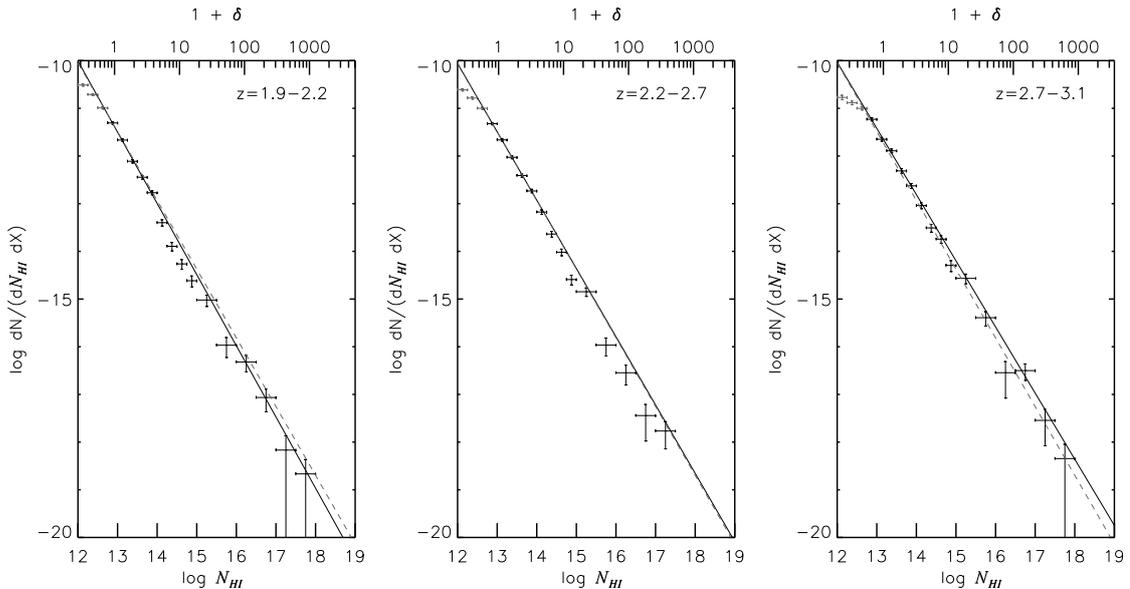


Figure 3.11.: The differential column density distribution as a function of redshift. Black and grey data points show the results from the high-order Lyman sample. The vertical error bars indicate 1σ Poisson errors. The grey data points mark the column densities that are affected by incompleteness. The black solid line gives the power law fit for $\log N_{\text{HI}} = [12.75, 14]$ whereas the grey dashed line resembles the fit to the $z = [1.9, 3.2]$ redshift range (see Fig. 3.10).

for column densities above $\log N_{\text{HI}} > 14.0$. By looking at the highest redshift bin in Fig. 3.11, it becomes evident that no such deviation is present there. A small deviation by more than 2σ starts to show for the intermediate redshift bin at column densities above $\log N_{\text{HI}} > 14.5$. This deviation increases for the lowest redshift bin and the deviation point shifts to a lower column density of $\log N_{\text{HI}} = 14.0$. This depression at high column densities is directly related to the dip at $z \sim 2.1$ seen in the number density evolution (compare Fig. 3.9). Thus the mechanism for reducing dn/dz is also responsible for the deviations in the differential column density distribution. Since the evolution in dn/dz is seen to regain its original power law form below $z = 2.0$, the deviation in the differential column density distribution is expected to diminish for lower redshifts. Unfortunately our QSO sample does not include many lines of sight for a good redshift coverage below $z = 2$ and thus does not allow the derivation of a well constrained $dN/(dN_{\text{HI}} dX)$.

3.5. Characteristics of the metal-enriched forest

The discovery of metal lines which are associated to H I absorbers in the Ly α forest, such as C IV or O VI (Cowie et al. 1995; Songaila 1998; Schaye et al. 2000a), have raised the question of how the IGM has been metal enriched. As the forest is characterised by a high temperature and a low gas density, it is not likely to form stars in-situ. Metals should be transferred from galaxies in some mechanisms such as, for example, galactic outflows (Aguirre et al. 2001b; Schaye et al. 2003; Oppenheimer & Davé 2006). In recent

years, studies on galaxy-galaxy pairs at high redshift have revealed some evidence that metals associated with the Ly α forest reside in the circum-galactic medium (Adelberger et al. 2005; Steidel et al. 2010). In this case, the metal-enriched forest cannot be called the IGM in the conventional sense and is likely to show a different evolutionary behaviour compared to the metal-free forest. In order to learn more about these enriched hydrogen absorbers, we characterise C IV enriched H I systems in this section by determining their number density evolution and differential column density distribution. All the results from this section are tabulated in the Section 3.7.

3.5.1. Method

In order to determine whether an H I absorption line is associated with C IV or not, we apply a simple algorithm to our fitted absorber line lists. The algorithm is designed in such a way, that it can be easily applied to model spectra generated from cosmological simulations investigating the IGM metal enrichment.

We consider an H I absorber to be metal enriched, if a C IV line is present in a specific velocity bin $\pm\Delta v_{\text{metal}}$ centred on the H I absorber. For each H I absorption line in the line list, we identify all the C IV lines available in the vicinity of the H I line that fall into the velocity bin Δv_{metal} . Then the search range is extended by a factor n_{metal} , and any additional C IV systems lying in the extended velocity range are identified. From all identified C IV lines a noise free transmission spectrum is generated and we check whether any transmission falls below a value of c_{metal} in the velocity range $\pm\Delta v_{\text{metal}}$ (centred on the H I line). If the transmission falls below c_{metal} the H I absorber is flagged as C IV enriched.

The parameter n_{metal} ensures that we include any contribution of C IV lines lying outside the $\pm\Delta v_{\text{metal}}$ range, which might still contribute to the transmission inside the considered velocity range. This can occur if the wings of a strong C IV absorber fall within $\pm\Delta v_{\text{metal}}$. Further the threshold value of c_{metal} excludes all possible spurious detections and is chosen to be on the order of the noise level. We have thoroughly checked the dependency of our results on c_{metal} and found a value of $c_{\text{metal}} = 0.95$ to be reasonable with our QSO sample. In the following, two choices of Δv_{metal} are considered: a conservative narrow range of $\pm 10 \text{ km s}^{-1}$ and a more generous interval of $\pm 100 \text{ km s}^{-1}$.

3.5.2. Results

C IV enriched absorber number density evolution

Similar to the analysis of all the H I absorbers, we calculate the absorber number density evolution $dn_{\text{C IV}}/dz$ on a QSO by QSO basis for all the C IV enriched H I absorbers. Unfortunately the number of enriched systems in our sample is too small to obtain the mean evolution from the combined sample. The resulting $dn_{\text{C IV}}/dz$ evolution is shown in Fig. 3.12 for the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ and $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ sample from the high-order Lyman fit samples. To increase the redshift coverage, the sample was extended outside the range covered by the high-order sample with the results from the Ly α -only fits at $1.7 < z < 1.9$.

As for the entire Ly α forest analysis, the $dn_{\text{C IV}}/dz$ evolution resembles a power law. Therefore linear regressions have been obtained from the data set and its results are

3. The IGM H I column density evolution

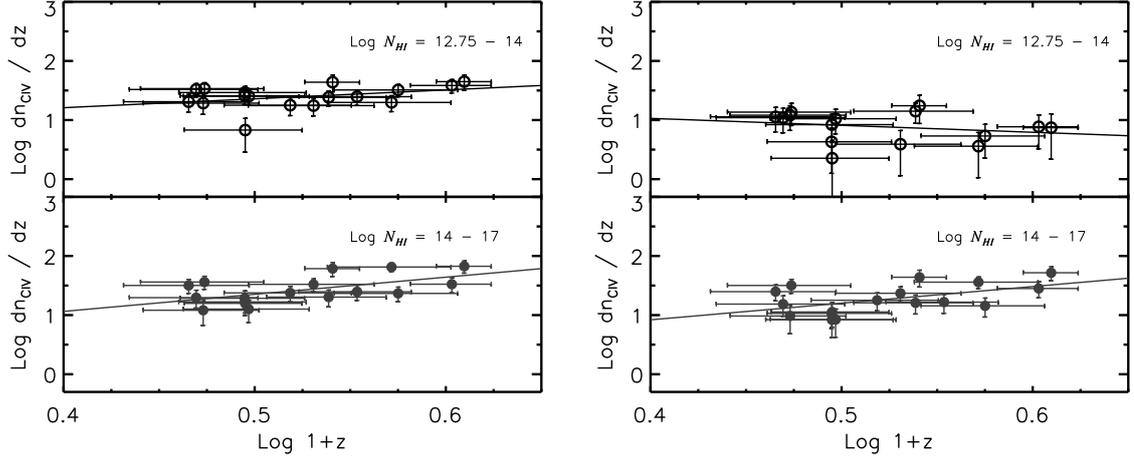


Figure 3.12.: Line number density evolution of all C IV enriched H I systems. The black open circles in the upper panel represent a column density interval of $\log N_{\text{HI}} = [12.75, 14.0]$ and the grey filled circles in the lower panel represent $\log N_{\text{HI}} = [14, 17]$. The error bars mark 1σ Poisson errors. The solid lines represent linear regressions to the data, using the parameters summarised in Table 3.5. The left panels are derived from a $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$, while the right panels represent $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$.

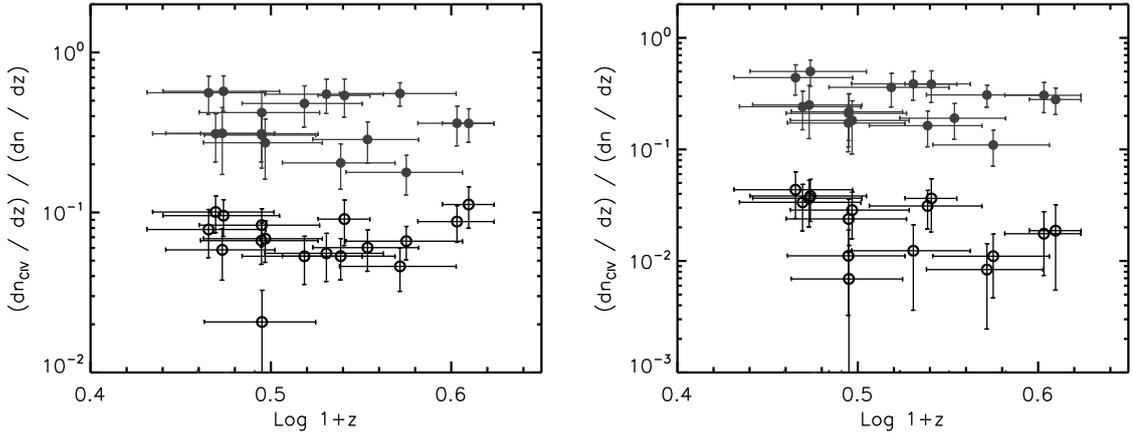


Figure 3.13.: The fraction of the C IV enriched H I absorber number density to the total absorber number density as a function of redshift. The black open circles represent a column density interval of $\log N_{\text{HI}} = [12.75, 14.0]$ and the grey filled circles represent $\log N_{\text{HI}} = [14, 17]$. The error bars mark 1σ Poisson errors. The left panels are derived from the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample, while the right panels represent the $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ sample.

Table 3.5.: Table of linear regression results for the number density evolution dN/dz of the C IV enriched H I forest absorbers.

$\Delta \log N$	$\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$		$\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$	
	$\log(dn/dz)_0$	γ	$\log(dn/dz)_0$	γ
12.75 - 14.0	0.61 ± 0.37	1.50 ± 0.68	1.50 ± 0.48	-1.17 ± 0.89
14.0 - 17.0	-0.09 ± 0.52	2.89 ± 0.98	-0.21 ± 0.59	2.81 ± 1.10

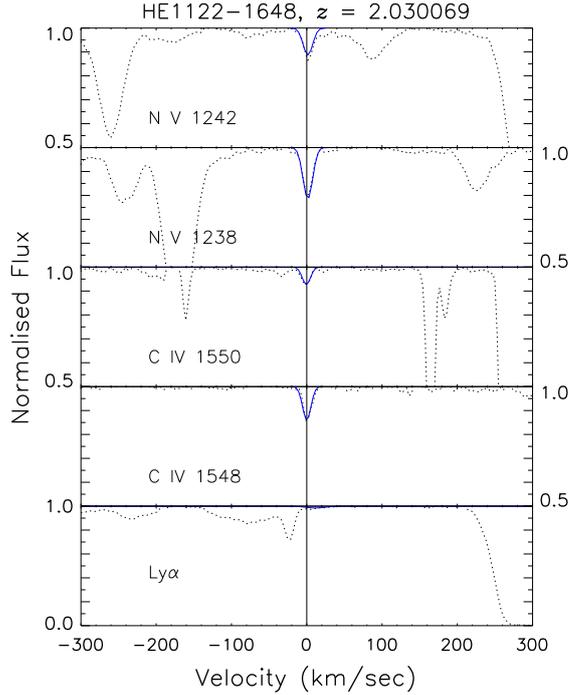


Figure 3.14.: A velocity plot of a highly ionised C IV absorber at $z = 2.030069$ towards HE1122–1648. The zero velocity is centred at $z = 2.030069$. Although there is no obvious Ly α absorption seen at the zero velocity, both O VI and N V doublets are present to secure the existence of this absorber. Note that the y-axis range for each ion is different: from the normalised flux 0 to 1 for H I and from 0.5 to 1 for the rest of the ions.

summarised in Table 3.5. Similar to the entire H I sample, the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample shows a decline in the C IV enriched absorber number density with decreasing redshift. This behaviour is present in both column density ranges. Comparing these results with the dn/dz of the entire sample shows that the resulting evolution of the C IV enriched absorbers has a similar slope as the evolution of all H I absorbers. The only difference is that the number densities are shifted to lower values at all redshifts.

This becomes evident in Fig. 3.13, where the fractions between the number densities of the C IV enriched systems $dn_{\text{C IV}}/dz$ and the number density of the entire sample dn/dz are shown. The results for the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample shows that there is no significant evolution of the C IV enrichment fraction for both column density ranges with redshift. For the low column density sample we find that around 7% of all the H I absorbers show C IV enrichment. The C IV enrichment fraction is higher for larger column densities, where around 50% of the absorbers are C IV enriched.

This picture changes slightly for the $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ sample. The $dn_{\text{C IV}}/dz$ evolution in Fig. 3.12 for the high column densities is almost identical to the one of the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample, except a slight shift to lower number densities. Now around 30% to 40% of the high column density H I absorbers in the forest are C IV enriched, showing a slight increase with decreasing redshift. The low column density results, however, indicate a completely different behaviour than in the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$

3. The IGM H I column density evolution

case. Now the number density of the C IV enriched forest does not decline with decreasing redshift, but increases instead. Looking at the enrichment fractions it is seen, that only 2% of the low column density lines are C IV enriched. However, the fraction of enriched systems increases slightly at $z < 2.2$ ($\log(1+z) < 0.5$) to 4%, although it is still consistent with no redshift evolution given the large scatter among different sightlines.

The two different velocity intervals in which we look for C IV absorption pick up two distinct groups of C IV absorbers. One group contains strong, saturated high column density H I absorbers. These absorbers are usually accompanied with smaller column density absorbers within a velocity range of $\Delta v < 200 \text{ km s}^{-1}$. In these systems, the C IV absorption is usually found within 20 km s^{-1} to the strongest H I absorber (Kim et al. 2011, *in preparation*, also see Section 2.4.2 and Fig. 2.7). Therefore these accompanied low column density systems get associated with these C IV absorbers if the velocity range Δv_{metal} is large. With a small velocity search interval, however, only H I systems that have C IV in their direct vicinity are flagged as C IV enriched. This means that the aforementioned low column density systems around strong absorbers are not considered C IV enriched.

Another C IV enriched group consists of usually isolated, low column density H I absorbers associated with C IV absorption, *i. e.* the same forest population studied by (Schaye et al. 2007). An example of such a system is shown in Fig. 3.14 towards HE1122–1648. In this velocity plot (the relative velocity centred at the redshift of an absorber vs normalised flux), an H I absorption feature is hardly recognisable, while the strong O VI and N V doublets are present. The existence of both doublets makes the identification of this absorber secure. Due to the low $N_{\text{H I}}$ and high N_{metals} , these systems show a higher ionisation and a higher metallicity compared to a typical absorber with similar $N_{\text{H I}}$ (Carswell et al. 2002; Schaye et al. 2007). Since photoionisation models indicate that the sizes of these absorbers are rather small at $\sim 100\text{pc}$, Schaye et al. (2007) speculate that these systems result from gas streams that could be responsible for transporting metals from galaxies to the surrounding IGM. The number of such *highly ionised systems* seems to increase with decreasing redshift in the same sample of QSOs (Kim et al. 2011, *in preparation*, also compare Section 2.3). Since most of the low column density systems picked up by the $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ interval are such *highly ionised systems*, the evolution of these systems can be quantified with the power-law parameters given in Table 3.5.

Differential column density distribution function of the C IV enriched forest

In Fig. 3.15, the differential column density distribution function for C IV enriched H I absorbers is shown for $1.9 < z < 3.2$. The results of the two different search velocity ranges $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ and $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ are shown. For $\log N_{\text{H I}} > 15.5$, the results are not sensitive to our choice of the search velocity. In Fig. 3.15, this manifests itself the filled circles and stars to be identical at $\log N_{\text{H I}} > 17$. For these column densities, the differential column density distribution function of the enriched systems is slightly shifted to lower values when compared to the entire sample. However the functional form remains similar to the entire sample, showing a power-law with similar slopes as in the entire sample. This indicates that a large fraction of high column density systems are C IV enriched.

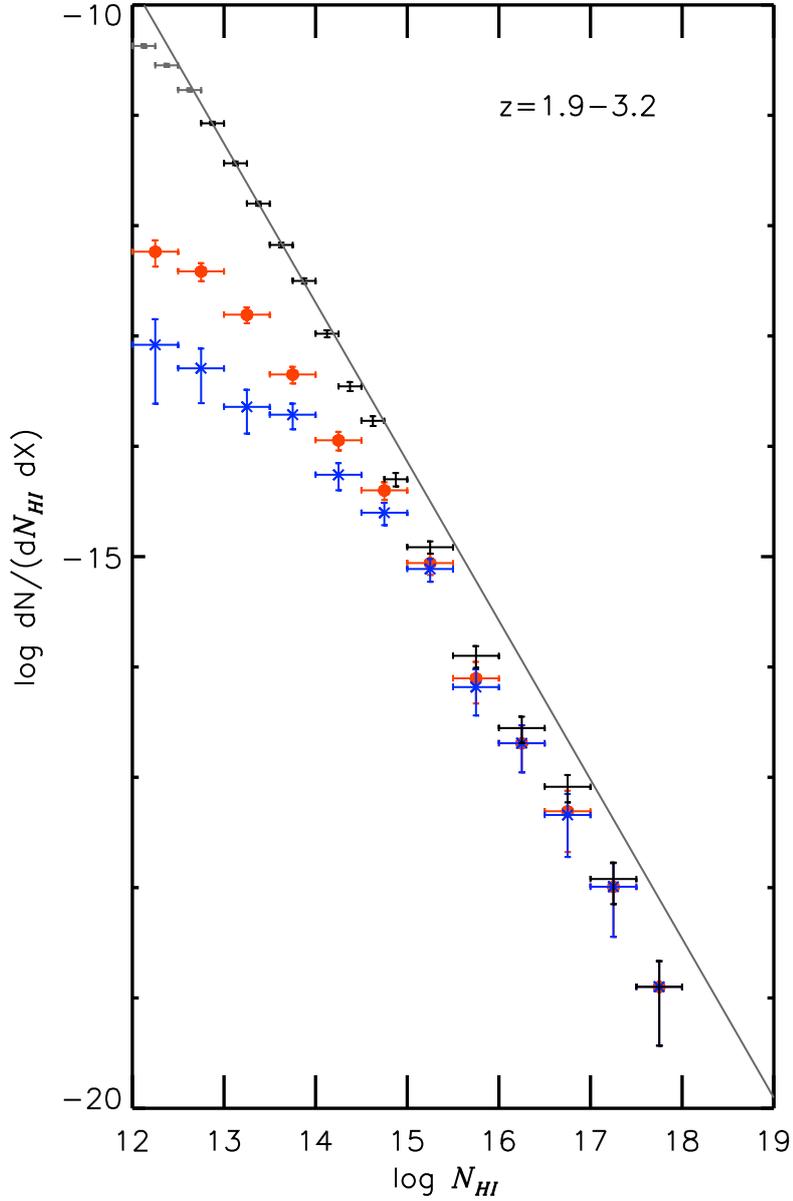


Figure 3.15.: Differential column density distribution function for all HI Ly α absorbers in our sample (black data points) at $1.9 < z < 3.2$. Grey data points below $\log N_{HI} < 12.75$ indicate the column densities affected by sample incompleteness. The solid line indicates the fit to the data as in Fig. 3.10. Further the distribution function for C IV enriched HI lines is shown for $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ (filled circles) and $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ (stars). The vertical errors indicate 1σ Poisson errors.

3. The IGM H I column density evolution

Below $\log N_{\text{HI}} < 15.5$ the distribution function of the C IV enriched systems starts to deviate strongly from the entire sample. The power-law slope becomes more shallow for the enriched systems than for the entire sample. Furthermore the slope of the enriched systems depends strongly on the choice of Δv_{metal} . The larger search velocity results in a steeper slope than the small one. Again this is due to the $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ sample being predominantly sensitive to highly ionised systems, whereas the larger velocity range is somewhat sensitive to low column density systems in the vicinity of strong systems as well.

The flattening of the distribution function seen at $\log N_{\text{HI}} < 15.5$ by C IV enriched systems cannot be caused by the incompleteness of the H I sample. Incompleteness would result in a similar flattening as is seen at $\log N_{\text{HI}} < 12.75$ for the entire sample. However, our sample of H I absorbers is complete for column densities larger than $\log N_{\text{HI}} > 12.75 \text{ cm}^{-2}$. Therefore the flattening at $\log N_{\text{HI}} < 15.5 \text{ cm}^{-2}$ is a property physically related to C IV enriched systems only.

The redshift evolution of the distribution function of C IV enriched and unenriched systems is shown in the upper panel of Fig. 3.16 for the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample. The distribution function is calculated for the redshift ranges $z = [1.5, 1.9]$, $[1.9, 2.2]$, $[2.2, 2.5]$, $[2.5, 3.2]$. The small spacing in redshift is chosen to detect any evolution in the distribution function, especially around the redshift of the dn/dz dip discussed in Sect. 3.4.1. Unfortunately the uncertainties are large due to the small coverage of the redshift path in each bin. For the redshift intervals of $z = [1.5, 1.9]$, $[1.9, 2.2]$, $[2.2, 2.5]$, $[2.5, 3.2]$, the redshift paths are $dX = 1.2, 8.3, 5.1, 8.0$. The lowest redshift bin is mainly used as an indicator of how the evolution of the distribution function continues towards lower redshifts, although the redshift path for this redshift bin is very small.

In all redshift bins most of the systems are C IV enriched above $\log N_{\text{HI}} > 15.5$ with only a few systems showing no enrichment. However at lower column densities, the unenriched systems dominate. Furthermore the enriched systems do not show any strong evolution with redshift. There is a hint that the C IV enriched system show a shallower exponent of the distribution function than the unenriched systems at $\log N_{\text{HI}} > 15.5$. The number of unenriched systems becomes lower than the one of the enriched systems with increasing column density and seems to eventually drop off at even higher densities. However, the error bars are still large enough to conclude that there is no significant difference in the distribution function between the two samples at $\log N_{\text{HI}} > 15.5$.

It is interesting to note that the unenriched forest shows a distinct dip at all redshifts around a column density of $\log N_{\text{HI}} \sim 14 - 15$, which evolves with redshift. At $2.5 < z < 3.2$ the dip starts at a column density of $\log N_{\text{HI}} \sim 15.5$. This off-set point shifts towards smaller column densities with decreasing redshift and lies at $\log N_{\text{HI}} \sim 14$ for $1.9 < z < 2.2$. This is very similar to the evolution of the distribution function of the entire forest sample. On the other hand, there is no noticeable dip in the distribution function of the C IV enriched forest.

The C IV enriched systems remain unchanged as a single power law above $\log N_{\text{HI}} > 15$, before their distribution function flattens out. However, a larger QSO sample is needed in order to determine more conclusively the evolution of the differential column density distribution for enriched and unenriched systems.

In the lower panel of Fig. 3.16 we show the evolution of the column density distribution function as a function of redshift, now using the $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ sample. Similar

3.5. Characteristics of the metal-enriched forest

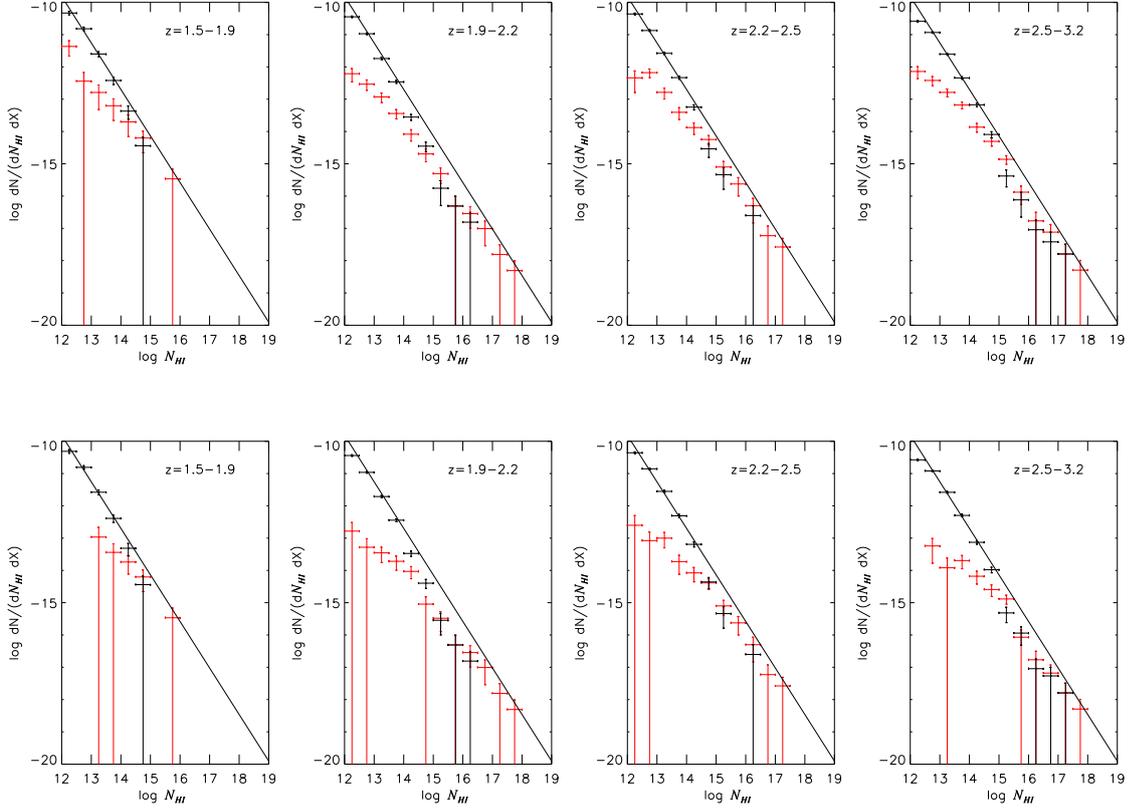


Figure 3.16.: Differential column density distribution function for enriched and unenriched H I Ly α absorbers in our high-order fit sample using $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ (upper panels) and $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$ (lower panels) as a function of redshift, except that the Ly α -only fit sample is used for the $1.5 < z < 1.9$ bins. The redshift increases from left to right. The black data points mark systems without any signs of C IV enrichment, whereas red data points indicate C IV enriched systems. The solid black line indicates the fit to the entire sample and the whole redshift range as in Fig. 3.10. The vertical errors indicate 1σ Poisson errors.

3. The IGM H I column density evolution

to the sample using the large Δv_{metal} , there is no apparent evolution with redshift for the C IV enriched systems for $\log N_{\text{HI}} > 15$. As with the $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ sample, the unenriched systems show a similarly evolving dip in the distribution function at $\log N_{\text{HI}} = 14 - 15$. Thus the differential column density distribution is only sensitive to the search velocity range at column densities below $\log N_{\text{HI}} = 14 \text{ cm}^{-2}$, which we already concluded from the dn/dz results in Section 3.5.2.

The fact that the differential column density distribution of the C IV enriched systems flattens at low column densities can be easily explained by that the enrichment fraction becomes smaller as N_{HI} decreases, as is evident from Fig. 3.13. At $\log N_{\text{HI}} = [13.0, 13.5, 14.0, 14.5, 15.0, 15.5]$, the fraction of the metal enriched forest is roughly $[0.7, 2, 7, 15, 51, 61]\%$, respectively. On the face value, this result seems to be contradictory to the ones based on the pixel optical depth method. Relating the H I pixel optical depths with the *median* C IV optical depth, the pixel optical depth method infers that statistically the C IV abundance decreases with decreasing H I optical depth in the form of a power law above a characteristic optical depth of around $\tau_{\text{HI}} \sim 1.5$ (or $\log N_{\text{HI}} \sim 13.7$ with $b = 30 \text{ km s}^{-1}$) (Ellison et al. 2000; Schaye et al. 2003). The relation between H I optical depths and C IV optical depths becomes constant at $\tau_{\text{HI}} \leq 1.5$. This apparent change is claimed to be caused by a washing out of the signal by noises, continuum fitting errors, and other contaminations, and is not a real observational fact. This in turn suggests that the IGM could be enriched at much lower column density ranges than $N_{\text{HI}} \leq 10^{13.5} \text{ cm}^{-2}$. On the face value, the optical depth result implies that both the enriched and the unenriched forest would show a similar column density distribution function down to about $N_{\text{HI}} \sim 10^{13.5-14.0} \text{ cm}^{-2}$, since the optical depth method based on the median value *statistically* shows a linear power-law between τ_{HI} and τ_{CIV} down to the characteristic $\tau_{\text{HI}} \sim 1.5$. If the number of pixels associated with C IV drops far below 50% of the total number of available pixels, such as our enriched forest, the *median* C IV optical depth method is not capable of picking up these minor systems. In fact, if we only take the distribution function at $\log N_{\text{HI}} \geq 10^{13.5-14}$, we could consider that the C IV enriched forest with $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ can be described with a single power law within the errors. Therefore, our Voigt profile analysis has the advantage that it enables to probe the enrichment of the forest at much smaller column densities. Additionally, it enables the study of the enriched forest in regimes where the population density is much smaller than the ones probed by the optical depth method.

Numerical simulations and theories have shown that the dominating physical process of the forest evolution changes at $z \sim 2$ from the Hubble expansion to the evolution of the ionising background (Theuns et al. 1998a; Davé et al. 1999; Bianchi et al. 2001). Based on this, we can assume that the likely origin of the dip shown both in dn/dz and the differential column density distribution function is caused by the change in the ionising background radiation. The fact that solely the unenriched forest shows the dip, indicates that it in general follows the theoretical prediction of the general IGM evolution. On the other hand, since no dip is present in the enriched forest, this indicates that the enriched systems are not dominated by the meta-galactic UV background, i.e. their ionising radiation does not undergo a significant change. If we assume that metals are only found close to galaxies in the circum-galactic medium as indicated by (Steidel et al. 2010) for $2 < z < 4$ and not far away from galaxies, this lack of a dip in the enriched forest can be easily explained. In the circum-galactic medium, the ionisation state of

the enriched forest is mainly dominated by the local flux stemming from nearby galaxies. The local flux of each star forming galaxy could be seen analogous to the star formation mechanism in different galaxies as an universal process. Therefore, the enriched forest shows a different distribution function compared to the unenriched forest.

In addition, numerical simulations have found that stronger absorbers reside closer to galaxies, while weak absorbers are mostly found far from galaxies. At high column densities of $\log N_{\text{HI}} > 15$ which probe regions near galaxies, almost all of the absorbers show associated C IV (see Fig. 3.15). Also the circum-galactic medium would be exposed to a stronger ionising background, which decreases the number of weak absorbers found in the circum-galactic medium. However, far away from the galaxies in to the IGM, where the number of the low column density systems is larger, a large fraction of absorbers are still unenriched. Therefore the column density distribution of the enriched forest flattens out at lower $\log N_{\text{HI}}$.

Note that the distribution function of the entire forest sample is dominated by the unenriched forest, as the number of the enriched forest becomes lower than the unenriched forest by up to a factor of 100 as z decreases. Therefore, the distribution function of the entire forest shows a similar behaviour as the unenriched forest, i.e. showing a dip.

3.6. Conclusions

Based on an in-depth Voigt profile fitting analysis of 18 high-redshift QSOs obtained from the ESO VLT/UVES archive, we have investigated the H I absorber number density evolution and the differential column density distribution function at $1.9 < z < 3.2$ and for $\log N_{\text{HI}} = [12.75 - 17.0]$. Two methods of the Voigt profile fitting analysis have been applied, one by fitting absorption profiles only to the Ly α transition and another by including higher order Lyman transitions such as Ly β and Ly γ . These higher order transitions provides a reliable column density measurement of saturated absorption systems, since saturated and blended lines often become unsaturated at higher order transitions. This also enables to resolve the structure of absorbers better. This study is based on by far the largest and most uniform sample of high-quality QSO spectra at $z > 2$, increasing the sample size by a factor of 3. In addition, we have also investigated whether there exist any differences in the N_{HI} evolution between the C IV enriched forest and the unenriched forest.

We have found that our results based on the Ly α -only fit are in good agreement with previous results on a QSO by QSO basis. For our data only (values in brackets indicate results including data from the literature at $z > 1$), the number density dn/dz is $dn/dz = 1.25 \pm 0.21 (1+z)^{2.07 \pm 0.39} (1.02 \pm 0.08 (1+z)^{2.42 \pm 0.15})$ and $dn/dz = -0.33 \pm 0.40 (1+z)^{4.08 \pm 0.76} (0.23 \pm 0.14 (1+z)^{2.99 \pm 0.26})$ at $\log N_{\text{HI}} = [13.1 - 14.0]$ and $[14.0 - 17.0]$, respectively. The large difference in the exponent between our sample and the sample including the data from the literature is caused by the fact that our sample does not cover a large redshift range and that the evolution of dn/dz is more pronounced for stronger absorbers. The scatter between different sightlines becomes larger at lower redshifts and stronger absorbers due to the evolution of the structure formation. The differential column density distribution function is also consistent with previous results. The single power-law exponent is $\sim -1.44 \pm 0.02$ at $1.9 < z < 3.2$ and $\log N_{\text{HI}} = [12.75 - 14.0]$,

3. The IGM H I column density evolution

with a deviation from the power law at $\log N_{\text{HI}} = [14.5 - 18.0]$.

The high-order Lyman fits do not show any significantly different results from the ones based on the Ly α -only fits. The dn/dz evolution based on a QSO by QSO case yields a very similar result to the Ly α -only fit. The *mean* dn/dz based on the combined sample from all the QSOs is $dn/dz = 1.91 \pm 0.10 (1+z)^{1.28 \pm 0.20}$ and $dn/dz = -0.21 \pm 0.27 (1+z)^{3.77 \pm 0.49}$ at $\log N_{\text{HI}} = [13.1 - 14.0]$ and $[14.0 - 17.0]$, respectively.

However, deriving the evolution of dn/dz from the combined sample shows a dip with a $2-3\sigma$ significance at $\log N_{\text{HI}} = [14, 17]$ at redshifts $2.0 < z < 2.3$. As our QSO sample shows the highest redshift path coverage around the redshift of this dip, the depression cannot be caused by the small redshift coverage of our sample. This dip is also present in the *combined* Ly α -only sample.

Using the high-order fits, we have derived the differential column density distribution function at $1.9 < z < 3.2$ and confirm the existence of a dip in the distribution function between $14 < \log N_{\text{HI}} < 18$ as already seen in Petitjean et al. (1993). At $1.9 < z < 3.2$, the power-law exponent of the differential column density distribution function is -1.44 ± 0.03 , -1.71 ± 0.08 and -1.51 ± 0.08 at $\log N_{\text{HI}} = [12.75 - 14.0]$, $[14.0 - 15.0]$ and $[15.0 - 18.0]$, respectively. At lower z , the power-law seems to be steeper for the column density range $\log N_{\text{HI}} = [12.75 - 14.0]$ in which the distribution function follows a perfect single power-law. By obtaining the differential column density distribution function for three redshift bins $z = [1.9, 2.2]$, $[2.2, 2.7]$, and $[2.7, 3.2]$, we observe that a dip at high column densities is clearly visible in the lowest redshift bin, while it becomes less pronounced at higher redshifts. The physical process responsible for the dip in the number density evolution is thus also responsible for the dip in the differential distribution function. This dip might be caused by the enhanced ionising background radiation due to the higher cosmic star formation rates at $z \sim 2$, resulting in a change in the ionisation fraction of stronger absorbers. However, the exact cause of this reduction in high column density systems around $z \sim 2$ cannot be constrained from these observations alone.

Further we have associated C IV absorption with H I absorbers using a simple algorithm that can be easily applied to synthetic spectra from metal enrichment simulations of the IGM. For this assigning algorithm, we used two search velocity intervals within which we look for associated C IV around H I absorbers, $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$ and $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$. We have split our absorber sample into two groups: absorbers associated with C IV tracing the metal enriched forest, and absorbers associated with no C IV tracing the unenriched forest. At $1.9 < z < 3.2$, about 2 – 7% of all absorbers show direct association with C IV at $\log N_{\text{HI}} = [12.75, 14]$, while about 40–50% are metal enriched at $\log N_{\text{HI}} = [14, 17]$. However, these numbers are strongly dependent on the search velocity distance.

At $\log N_{\text{HI}} = [14, 17]$, we find the C IV enriched H I systems to evolve similar to the entire Ly α forest, with a decrease in number with decreasing redshift. However, the enriched systems at $\log N_{\text{HI}} = [12.75, 14]$, show the contrary. The number of enriched absorbers increases with decreasing redshift. This is due to the existence of highly ionised systems at low redshifts. They are characterised by their large N_{CIV} compared to their N_{HI} which is usually less than $\log N_{\text{HI}} \leq 14$. These systems are similar to the ones extensively studied by Carswell et al. (2002) and Schaye et al. (2007).

The differential column density distribution function for the enriched and unenriched systems show a significant difference at all of the 3 redshift bins, $z = [1.9, 2.2]$, $[2.2, 2.7]$,

Table 3.6.: Number density evolution data for each QSO.

QSO	z	Δz	Ly α -only fit				higher-order fit			
			$\log N_{\text{HI}} = [12.75 - 14]$ dn/dz	$\Delta dn/dz$	$\log N_{\text{HI}} = [14 - 17]$ dn/dz	$\Delta dn/dz$	$\log N_{\text{HI}} = [12.75 - 14]$ dn/dz	$\Delta dn/dz$	$\log N_{\text{HI}} = [14 - 17]$ dn/dz	$\Delta dn/dz$
Q0055–269	3.27	0.33	2.672	0.025	2.157	0.044	2.600	0.042	2.269	0.062
PKS2126–158	3.01	0.2	2.637	0.033	1.900	0.079	2.642	0.033	1.965	0.073
Q0420–388	2.76	0.28	2.685	0.026	2.151	0.049	2.688	0.026	2.117	0.051
HE0940–1050	2.73	0.28	2.628	0.028	1.997	0.059	2.635	0.028	2.069	0.054
HE2347–4342	2.58	0.24	2.619	0.031	1.939	0.068	2.615	0.031	1.939	0.068
Q0002–422	2.46	0.25	2.651	0.029	1.948	0.066	2.657	0.029	1.995	0.062
PKS0329–255	2.39	0.26	2.483	0.035	1.846	0.073	2.499	0.034	1.781	0.079
Q0453–423	2.47	0.11	2.689	0.041	1.941	0.099	2.682	0.042	2.055	0.086
HE1347–2457	2.3	0.25	2.506	0.034	1.658	0.092	2.525	0.033	1.695	0.088
Q0329–385	2.14	0.24	2.571	0.033	1.602	0.101	2.566	0.033	1.666	0.094
HE2217–2818	2.13	0.24	2.575	0.032	1.524	0.111	2.545	0.034	1.598	0.101
Q0109–3518	2.13	0.22	2.518	0.036	1.715	0.092	2.515	0.036	1.715	0.092
HE1122–1648	2.12	0.23	2.588	0.032	1.793	0.082	2.586	0.032	1.793	0.082
J2233–606	1.98	0.22	2.535	0.035	1.586	0.107	2.551	0.035	1.803	0.083
PKS0237–23	1.97	0.21	2.520	0.037	1.613	0.107	2.520	0.037	1.587	0.111
PKS1448–232	1.95	0.23	2.514	0.036	1.772	0.085	2.514	0.036	1.803	0.082
Q0122–380	1.92	0.22	2.435	0.040	1.754	0.088	2.416	0.041	1.754	0.088
Q1101–264	1.99	0.11	2.483	0.054	1.509	0.173	2.483	0.054	1.509	0.173

and [2.7, 3.2]. The distribution of C IV enriched systems is found to flatten out at $\log N_{\text{HI}} \leq 15$, while the unenriched systems show a power-law distribution similar to the entire forest sample. Depending on the search velocity interval, the number of enriched systems is a factor of 15 ($\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$) to 126 ($\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$) lower than the one of the unenriched systems at $\log N_{\text{HI}} = 13$. This is caused by a decrease in the enriched fraction of the Ly α forest as $\log N_{\text{HI}}$ decreases.

The unenriched systems are found to show a drop in number above a characteristic column density between $14 < \log N_{\text{HI}} < 15.5$. The starting point of this dip in the unenriched absorbers shifts from $\log N_{\text{HI}} = 15.5$ at redshift $2.5 < z < 3.2$ to lower column densities at lower redshifts. The same physical process causing the dn/dz dip is responsible for this, probably due to the increased UV radiation due to the peak of the cosmic star formation rate at $z \sim 2$. The fact that the unenriched systems show an evolution with redshift, whereas the enriched ones do not, suggests that the two systems are likely to reside in different spatial locations. The C IV enriched forest is found to trace the circum-galactic medium, where the radiation from the galaxies dominate over the meta-galactic ionising background. Therefore this population is not sensitive to any changes in the background radiation. On the other hand, the unenriched forest systems trace the intergalactic medium far away from galaxies and is thus sensitive to the evolution in the ionising background. Besides this evolution, the two groups show no other redshift evolution and the distribution function seems to stay constant at $1.9 < z < 3.2$. However, tighter constraints from larger sample sizes are needed to better understand the characteristics and evolution of the C IV enriched Ly α forest.

3.7. Data

In this appendix we give all the data for the QSO by QSO number density evolution (Table 3.6), the mean number density evolution (Table 3.7), the differential column density distribution for the entire sample (Table 3.8) and for the C IV enriched systems (Table 3.9 and 3.10).

3. The IGM H I column density evolution

Table 3.7.: Mean number density evolution data for $\Delta z = 0.1$

z	Ly α -only fit			higher-order fit		
	$\log N_{\text{H I}} = [12.75 - 14]$ dn/dz	$\log N_{\text{H I}} = [14 - 17]$ dn/dz	$\log N_{\text{H I}} = [14 - 17]$ $\Delta \text{dn}/\text{dz}$	$\log N_{\text{H I}} = [12.75 - 14]$ dn/dz	$\log N_{\text{H I}} = [14 - 17]$ dn/dz	$\log N_{\text{H I}} = [14 - 17]$ $\Delta \text{dn}/\text{dz}$
1.85	2.506	0.037	1.696	0.094	2.506	0.037
1.95	2.512	0.026	1.721	0.064	2.510	0.026
2.05	2.511	0.025	1.676	0.065	2.511	0.025
2.15	2.587	0.024	1.613	0.074	2.586	0.024
2.25	2.501	0.029	1.652	0.079	2.509	0.029
2.35	2.585	0.027	1.829	0.066	2.593	0.027
2.45	2.596	0.029	1.918	0.064	2.598	0.029
2.55	2.610	0.027	1.892	0.062	2.606	0.027
2.65	2.640	0.031	2.086	0.059	2.638	0.031
2.75	2.659	0.037	2.107	0.070	2.665	0.037
2.85	2.647	0.037	1.915	0.088	2.644	0.038
2.95	2.651	0.034	2.111	0.064	2.672	0.033
3.05	2.681	0.040	2.131	0.076	2.681	0.040

Table 3.8.: Differential column density evolution $f = \log(dN/(dN_{\text{H I}} dX))$

$\log N_{\text{H I}}$	$z = 1.9 - 3.2$			$z = 1.9 - 2.2$			$z = 2.2 - 2.7$			$z = 2.7 - 3.2$		
	f	$+\Delta f$	$-\Delta f$									
12.125	-10.366	0.016	0.016	-10.277	0.023	0.024	-10.378	0.024	0.025	-10.537	0.044	0.049
12.375	-10.545	0.014	0.015	-10.479	0.022	0.023	-10.550	0.022	0.023	-10.650	0.038	0.041
12.625	-10.770	0.014	0.014	-10.758	0.022	0.024	-10.766	0.021	0.022	-10.766	0.032	0.035
12.875	-11.071	0.015	0.015	-11.072	0.024	0.026	-11.088	0.023	0.024	-10.996	0.032	0.034
13.125	-11.434	0.017	0.017	-11.431	0.027	0.029	-11.428	0.025	0.026	-11.414	0.038	0.042
13.375	-11.800	0.019	0.020	-11.876	0.034	0.037	-11.793	0.028	0.030	-11.657	0.038	0.042
13.625	-12.176	0.022	0.023	-12.211	0.037	0.041	-12.176	0.033	0.036	-12.078	0.046	0.051
13.875	-12.498	0.024	0.025	-12.537	0.040	0.045	-12.499	0.036	0.039	-12.391	0.049	0.055
14.125	-12.987	0.031	0.034	-13.165	0.061	0.071	-12.941	0.044	0.049	-12.805	0.059	0.068
14.375	-13.459	0.040	0.044	-13.660	0.079	0.097	-13.403	0.055	0.064	-13.274	0.074	0.089
14.625	-13.793	0.044	0.049	-14.029	0.090	0.113	-13.784	0.064	0.075	-13.509	0.073	0.088
14.875	-14.302	0.058	0.067	-14.382	0.100	0.130	-14.357	0.090	0.113	-14.060	0.100	0.130
15.250	-14.922	0.054	0.061	-15.106	0.103	0.135	-14.932	0.079	0.097	-14.651	0.090	0.113
15.750	-15.885	0.088	0.110	-16.053	0.161	0.257	-16.051	0.149	0.228	-15.476	0.125	0.176
16.250	-16.540	0.103	0.135	-16.407	0.139	0.206	-16.631	0.161	0.257	-16.629	0.232	0.533
16.750	-17.072	0.106	0.141	-17.150	0.176	0.301	-17.528	0.232	0.533	-16.585	0.139	0.206
17.250	-17.908	0.149	0.228	-18.252	0.301		-17.852	0.198	0.374	-17.629	0.232	0.533
17.750	-18.885	0.232	0.533	-18.752	0.301					-18.430	0.301	

Table 3.9.: Differential column density evolution $f = \log(dN/(dN_{\text{H I}} dX))$ of C IV enriched systems in $\Delta v_{\text{metal}} = 100 \text{ km s}^{-1}$

$\log N_{\text{H I}}$	$z = 1.9 - 3.2$			$z = 1.9 - 2.2$			$z = 2.2 - 2.5$			$z = 2.5 - 3.2$		
	f	$+\Delta f$	$-\Delta f$									
12.25	-12.236	0.103	0.135	-12.209	0.158	0.254	-12.338	0.215	0.454	-12.140	0.149	0.228
12.75	-12.414	0.074	0.089	-12.533	0.132	0.191	-12.177	0.114	0.156	-12.418	0.119	0.165
13.25	-12.807	0.066	0.077	-12.929	0.121	0.174	-12.787	0.133	0.207	-12.788	0.105	0.138
13.75	-13.349	0.069	0.082	-13.442	0.121	0.169	-13.405	0.145	0.228	-13.180	0.094	0.120
14.25	-13.945	0.076	0.093	-14.084	0.140	0.211	-13.877	0.140	0.211	-13.865	0.114	0.157
14.75	-14.399	0.073	0.088	-14.695	0.155	0.243	-14.247	0.123	0.174	-14.306	0.109	0.148
15.25	-15.058	0.086	0.108	-15.305	0.173	0.301	-15.099	0.173	0.301	-14.861	0.113	0.153
15.75	-16.103	0.149	0.228	-16.308	0.301		-15.625	0.198	0.374	-15.879	0.190	0.381
16.25	-16.691	0.162	0.264	-16.544	0.215	0.454	-16.301	0.232	0.533	-16.768	0.267	
16.75	-17.308	0.186	0.370	-17.007	0.232	0.533	-17.227	0.301		-17.117	0.239	
17.25	-17.992	0.215	0.454	-17.808	0.301		-17.576	0.267		-17.793	0.301	
17.75	-18.899	0.232	0.533	-18.308	0.301					-18.293	0.301	

Table 3.10.: Differential column density evolution $f = \log(dN/(dN_{\text{H I}} dX))$ of C IV enriched systems in $\Delta v_{\text{metal}} = 10 \text{ km s}^{-1}$

$\log N_{\text{H I}}$	$z = 1.9 - 3.2$			$z = 1.9 - 2.2$			$z = 2.2 - 2.5$			$z = 2.5 - 3.2$		
	f	$+\Delta f$	$-\Delta f$									
12.25	-13.080	0.232	0.533	-12.783	0.267		-12.602	0.301				
12.75	-13.293	0.179	0.316	-13.283	0.267		-13.076	0.267		-13.242	0.232	0.533
13.25	-13.642	0.155	0.243	-13.456	0.176	0.301	-13.000	0.176	0.301	-13.918	0.301	
13.75	-13.714	0.101	0.132	-13.717	0.161	0.275	-13.727	0.196	0.395	-13.695	0.158	0.254
14.25	-14.258	0.106	0.140	-14.030	0.149	0.228	-14.076	0.168	0.279	-14.180	0.155	0.243
14.75	-14.602	0.090	0.113	-15.044	0.225		-14.386	0.140	0.209	-14.592	0.143	0.217
15.25	-15.112	0.091	0.116	-15.482	0.204	0.417	-15.099	0.173	0.301	-14.884	0.116	0.159
15.75	-16.182	0.161	0.257	-16.308	0.301		-15.625	0.198	0.374	-16.069	0.231	
16.25	-16.691	0.162	0.264	-16.544	0.215	0.454	-16.301	0.232	0.533	-16.768	0.267	
16.75	-17.341	0.190	0.381	-17.007	0.232	0.533	-17.227	0.301		-17.180	0.249	
17.25	-17.992	0.215	0.454	-17.808	0.301		-17.576	0.267		-17.793	0.301	
17.75	-18.899	0.232	0.533	-18.308	0.301					-18.293	0.301	

Ich habe wenig Geduld mit Wissenschaftlern, die ein Brett dahernehmen, sich die dünnste Stelle aussuchen, und dort einen Haufen Löcher bohren, wo es sich mühelos bohren lässt.

Albert Einstein

4

Enabling parallel computing in CRASH¹

We present the new parallel version (pCRASH2) of the cosmological radiative transfer code CRASH2 for distributed memory supercomputing facilities. The code is based on a static domain decomposition strategy inspired by geometric dilution of photons in the optical thin case that ensures a favourable performance speed-up with increasing number of computational cores. Linear speed-up is ensured as long as the number of radiation sources is equal to the number of computational cores or larger. The propagation of rays is segmented and rays are only propagated through one sub-domain per time step to guarantee an optimal balance between communication and computation. We have extensively checked pCRASH2 with a standardised set of test cases to validate the parallelisation scheme. The parallel version of CRASH2 can easily handle the propagation of radiation from a large number of sources and is ready for the extension of the ionisation network to species other than hydrogen and helium.

4.1. Introduction

The field of computational cosmology has developed dramatically during the last decades. Especially the evolution of the baryonic physics has played an important role in the understanding of the transition from the smooth early universe to the structured present one. We are especially interested in the development of the intergalactic ionising radiation field and the thermodynamic state of the baryonic gas. First measurements of the epoch of reionisation are soon expected from new radio interferometers such as LOFAR² or MWA³, which will become operative within one year. The interpretation of these measurements requires, among others, a treatment of radiative transfer coupled to cosmological structure formation.

Over the last decade, many different numerical algorithms have emerged, allowing the continuum radiative transfer equation to be solved for arbitrary geometries and density distributions. A substantial fraction of the codes solve the radiative transfer (RT) equation on regular or adaptive grids (for example see Gnedin & Abel (2001); Abel & Wandelt (2002); Razoumov et al. (2002); Mellema et al. (2006)). Other codes developed schemes that introduce RT into the SPH formalism (Pawlik & Schaye 2008; Petkova & Springel 2009; Hasegawa & Umemura 2010) or into unstructured grids (Ritzerveld et al. 2003; Paardekooper et al. 2010) The large amount of different numerical strategies prompted a comparison of the different methods on a standardised problem set. For

¹Accepted in MNRAS for publication as Partl, A.M.; Maselli, A.; Ciardi, B.; Ferrara, A.; and Müller, V. MNRAS 2011.

²<http://www.lofar.org/>

³<http://www.mwatelescope.org/>

4. Enabling parallel computing in CRASH

results of this comparison project we refer the reader to Iliev et al. (2006) and Iliev et al. (2009), where the performance of 11 cosmological RT and 10 radiation hydrodynamic codes are systematically studied.

Our straightforward and very flexible approach is based on a ray-tracing Monte Carlo (MC) scheme. Exploiting the particle nature of a radiation field, it is possible to solve the RT equation for arbitrary three-dimensional Cartesian grids and an arbitrary distribution of absorbers. By describing the radiation field in terms of photons, which are then grouped into photon packets containing a large number of photons each, it is possible to solve the RT along one dimensional rays. With this strategy the explicit dependence on direction and position can be avoided. Instead of directly solving for the intensity field only the interaction of photons with the gas contained in the cells needs to be modelled, as done in long and short characteristics algorithms (Mellema et al. 2006; Rijkhorst et al. 2006; Whalen & Norman 2006). MC ray-tracing schemes differ from short and long characteristic methods. Instead of casting rays through the grid to each cell in the computational domain, the radiation field is described statistically by shooting rays in random directions from the source. However unlike in fully Monte Carlo transport schemes where the location of the photon matter interaction is determined by sampling the packet's mean free path, the packets are propagated and attenuated through the grid from cell to cell along rays until all the photons are absorbed or the packets exits the computational domain. This allows for an efficient handling of multiple point sources and diffuse radiation fields, such as recombination radiation or the ultraviolet (UV) background field. Additionally this statistical approach easily allows for sources with anisotropic radiation. A drawback of any Monte Carlo sampling method however is the introduction of numerical noise. By increasing the number of rays used for the sampling of the radiation field though, numerical noise can be reduced at cost of computational resources.

Such a ray-tracing MC scheme has been successfully implemented in our code `CRASH2`, which is, to date, one of the main references among RT numerical methods used in cosmology. `CRASH` was first introduced by Ciardi et al. (2001) to follow the evolution of hydrogen ionisation for multiple sources under the assumption that hydrogen has a fixed temperature. Then the code has been further developed by including the physics of helium chemistry, temperature evolution, and background radiation (Maselli et al. 2003; Maselli & Ferrara 2005). In its latest version, `CRASH2`, the numerical noise by the MC sampling has been greatly reduced through the introduction of coloured photon packets (Maselli et al. 2009).

The problems that are being solved with cosmological RT codes become larger and larger, in terms of computational cost. Especially, the study of reionisation is a demanding task, since a vast number of sources and large volumes are needed to properly model the era of reionisation (Baek et al. 2009; McQuinn et al. 2007; Trac & Cen 2007; Iliev et al. 2006). Furthermore the addition of more and more physical processes to `CRASH2` requires increased precision in the solution. To study such computationally demanding models with `CRASH2`, the code needs support for parallel distributed memory computers. In this chapter we present the parallelisation strategy adopted for our MPI parallel version of the latest version of the serial `CRASH2` code, which we call `pCRASH2`.

The chapter is structured as follows. First we give a brief summary of the serial `CRASH2` implementation in Section 4.2. In Section 4.3 we review the existing parallelisation

strategies for MC ray-tracing codes and describe the approach taken by pCRASH2. In Section 4.4 we extensively test the parallel implementation against standardised test cases. We further study the scaling properties of the parallel code in Section 4.4.3 and summarise our results in Section 4.5. Throughout this chapter we assume $h = 0.7$.

4.2. CRASH2: Summary of the algorithm

In this Section we briefly summarise the CRASH2 code. A complete description of the algorithm is found in Maselli et al. (2003) and in Maselli et al. (2009), with an additional detailed description of the implementation for the background radiation field given in Maselli & Ferrara (2005). We refer the interested reader to these papers for a full description of CRASH2.

CRASH2 is a Monte-Carlo long-characteristics continuum RT code, which is based on ray-tracing techniques on a three-dimensional Cartesian grid. Since many of the processes involved in RT, like recombination emission or scattering processes, are probabilistic, Monte Carlo methods are a straight forward choice in capturing these processes adequately. CRASH2 therefore relies heavily on the sampling of various probability distribution functions (PDFs) which describe several physical processes such as the distribution of photons from a source, reemission due to electron recombination, and the emission of background field photons. The numerical scheme follows the propagation of ionising radiation through an arbitrary H/He static density field and captures the evolution of the thermal and ionisation state of the gas on the fly. The typical RT effects giving rise to spectral filtering, shadowing and self-shielding are naturally captured by the algorithm.

The radiation field is discretised into distinct energy packets, which can be seen as packets of photons. These photon packets are characterised by a propagation direction and their spectral energy content $E(\nu_j)$ as a function of discrete frequency bins ν_j . Both the radiation fields arising from multiple point sources, located arbitrarily in the box, and from diffuse radiation fields such as the background field or radiation produced by recombining electrons are discretised into such photon packets.

Each source emits photon packets according to its luminosity L_s at regularly spaced time intervals Δt . The total energy radiated by one source during the total simulation time t_{sim} is $E_s = \int_0^{t_{\text{sim}}} L_s(t_s) dt_s$. For each source, E_s is distributed in N_p photon packets. The energy emitted per source in one time step is further distributed according to the source's spectral energy distribution function into N_ν frequency bins ν_j . We call such a photon packet a coloured packet. Then for each coloured packet produced by a source in one time step, an emission direction is determined according to the angular emission PDF of the source. Thus N_p is the main control parameter in CRASH2 governing both the time resolution as well as the spatial resolution of the radiation field.

After a source produced a coloured packet, it is propagated through the given density field. Every time a coloured packet traverses a cell k , the length of the path within each crossed cell is calculated and the cell's optical depth to ionising continuum radiation τ_c^k is determined by summing up the contribution of the different absorber species (H I, He I, He II). The total number of photons absorbed in cell k per frequency bin ν_j is thus

$$N_{A,\gamma}^{(k)} = N_{T,\gamma}^{(k-1)}(\nu_j) \left[1 - e^{-\tau_c^k(\nu_j)} \right] \quad (4.1)$$

4. Enabling parallel computing in CRASH

where $N_{T,\gamma}^{(k-1)}$ is the number of photons transmitted through cell $k-1$. The total number of absorbed photons is then distributed to the various species according to their contribution to the cell's total optical depth. Before the packet is propagated to the next cell, the cell's ionisation fractions and temperature are updated by solving the ionisation network for Δx_{HI} , Δx_{HeI} , Δx_{HeII} , and by solving for changes in the cells temperature ΔT due to photo-heating and the changes in the number of free particles of the plasma. The number of recombining electrons N_{rec} is recorded as well and is used for the production of the diffuse recombination radiation. In addition to the discrete process of photoionisation, CRASH2 includes various continuous ionisation and cooling processes in the ionisation network (bremsstrahlung, Compton cooling/heating, collisional ionisation, collisional ionisation cooling, collisional excitation cooling, and recombination cooling).

After these steps, the photon packet is propagated to the next cell and these steps are repeated until the packet is either extinguished or, if periodic boundary conditions are not considered, until it leaves the simulation box. At fixed time intervals Δt_{rec} , the grid is checked for any cell that has experienced enough recombination events to reach a certain threshold criteria $N_{\text{rec}} \geq f_{\text{rec}} N_a$, where N_a is the total number of species "a" atoms and $f_{\text{rec}} \in [0, 1]$ is the recombination threshold. If the reemission criteria is fulfilled, a recombination emission packet is produced by sampling the probability that a photon with energy larger than the ionisation threshold of H or He is emitted. The spectral energy distribution of the photon packet is determined by the Milne spectrum (Mihalas & Weibel Mihalas 1984). After the reemission event, the cell's counter for recombination events is put to zero and the photon packet is propagated through the box.

For further details on the algorithm and its implementation, we again refer the reader to the papers mentioned above.

4.3. Parallelisation strategy

Monte Carlo radiation transfer methods are a powerful and easy to implement class of algorithms that enable a determination of the radiation intensity in a simulation grid or on detectors (such as CCDs or photographic plates). Photons originating at sources are followed through the computational domain, i.e. the whole simulation box, up to a grid cell or the detector in a stochastic fashion (Jonsson 2006; Juvela 2005; Bianchi et al. 1996). If only the intensity field is of interest, a straight forward parallelisation strategy is to mirror the computational domain and all its sources on multiple processors, so that every processor holds a copy of the same data set. Then each node (a node can consist of multiple computational cores) propagates its own subset of the global photon sample through the domain until the grid boundary or the detector plane is reached. At the end, the photon counts which were determined independently on each node or core are gathered to the master node and are summed up to obtain the final intensity map (Marakis et al. 2001). This technique is also known as reduction. This strategy however only works if the memory requirement of the problem setup fits the memory available to each core. What if the computational core's memory does not allow for duplication of the data?

To solve this problem, hybrid solutions have been proposed, where the computational domain is decomposed into sub-domains and distributed to multiple task farms (Alme

et al. 2001). Each task farm is a collection of nodes and/or cores working on the same sub-domain. A task farm can either reside on just one computational node, or it can span over multiple nodes. Each entity in the task farm propagates photons individually through the sub-domain until they reach the border or a detector. If photons reach the border of the sub-domain, they are communicated to the task farm containing the neighbouring sub-domain. The cumulated intensity map is obtained by first aggregating the different contributions of the computational cores in each task farm, and then by merging the solutions of the individual task farms. The hybrid use of distributed (multiple nodes per task farm) and shared memory concepts (domain is shared between all cores per task farm) allows to balance the amount of communication that is needed between the various task farms, and the underlying computational complexity.

However this method has two potential drawbacks. If, in a photon scattering process, the border of the sub-domain lies unfavourably in the random walk, and the photon crosses the border multiple times in one time step, a large communication overhead is produced, slowing down the calculation. This eventuality arises in optically thick media. Further, in an optically thin medium, the mean free path of the photons can be larger than the sub-domain size. If the photons need to pass through a number of sub-domains during one time step, they need to be communicated at every border crossing event. This causes a large synchronisation overhead. Sub-domains would need to communicate with their neighbours often per time step, in order to allow a synchronous propagation of photons through the domain. Since in **CRASH2** photons are propagated instantly through the grid, each photon might pass through many sub-domains, triggering multiple synchronisation events. This important issue has to be taken into account in order to avoid inefficient parallelisation performance and scalability.

These task farm methods usually assume that the radiation intensity is determined separately from additional physical processes, while in cosmological radiative transfer methods the interest is focused on the coupling of the ionising radiation transfer and the evolution of the chemical and thermal state of the gas in the IGM. In **CRASH2** a highly efficient algorithm is obtained by coupling the calculation of the intensity in a cell with the evaluation of the ionisation network. The ionisation network is solved each time a photon packet passes through a cell, altering its optical depth. However, if on distributed architectures the computational domain is copied to multiple computational nodes, one would need to ensure that each time the optical depth in a cell is updated, it has to be updated on all the nodes containing a copy of the cell. This would produce large communication overheads. Further, special care needs to be taken to prevent two or more cores from altering identical cells at the same time, again endangering the efficiency of the parallelisation.

One possibility to avoid the problem of updating cells across multiple nodes would be to use a strict shared memory approach, where a task farm consists of only one node and all the cores have access to the same data in memory. Such an OpenMP parallelisation of the **CRASH2** Monte Carlo method has been feasible for small problems (Partl et al. 2010) where the problem size fits one shared computational node, but would not perform well if the problem needs to be distributed over multiple nodes. However the unlikely case of multiple cores accessing the same cell simultaneously still remains with such an approach.

We have decided to use the more flexible approach, by parallelising **CRASH2** for dis-

4. Enabling parallel computing in CRASH

tributed memory machines using the MPI library⁴. Using a distributed MPI approach however requires the domain to be decomposed into sub-domains. Each sub-domain is assigned to only one core, which means that the sub-domains become rather small when compared to the task farm approach. Since in a typical simulation the photon packets will not be homogeneously distributed, the domain decomposition needs to take this into account, otherwise load imbalances dominate over performance. This can be addressed by adaptive load balancing, which is technically complicated to achieve, though. An alternative approach is to statically decompose the grid using an initial guess of the expected computational load. This is the method we adopted for pCRASH2.

In order to optimise the CRASH2 code basis for larger problem sizes, the routines in CRASH2 handling the reemission of recombination radiation had to be adapted. Since pCRASH2 greatly extends the maximum number of photon packets that can be efficiently processed by the code, we revert changes introduced in the recombination module of the serial version CRASH2 to reduce the execution time. To handle recombination radiation in CRASH2 effectively, photon packets produced by the diffuse component were only emitted at fixed time steps. At these specific time steps, the whole grid was searched for cells that fulfilled the reemission criterium and reemission packets were emitted. This approach allowed the sampling resolution with which the diffuse radiation field was resolved to be controlled, depending on the choice of the recombination emission time interval.

Searching the whole grid for reemitting cells becomes a bottle neck when larger problem sizes are considered. In pCRASH2 we therefore reimplemented the original prescription for recombination radiation as described in the Maselli et al. (2003) paper. Diffuse photons are emitted whenever a cell is crossed by a photon packet and the recombination threshold in the cell has been reached. This results in a more continuous emission of recombination photons and increases the resolution with which this diffuse field is sampled. The two methods converge when the time interval between reemission events in CRASH2 is chosen to be very small.

4.3.1. Domain decomposition strategy

An intuitive solution to the domain decomposition problem is to divide the grid into cubes of equal volume. However in CRASH2 this would result in very bad load balancing and would deteriorate the scaling performance. The imbalance arises firstly from the large surface through which packets need to be communicated to the neighbouring domain, resulting in large communication overhead. A decomposition strategy that minimises the surface of the sub-domains reduces the amount of information that needs to be passed on to neighbouring domains. Secondly, the main contribution to the imbalance stems from the fact that the sub-domain where the source is embedded has to process more rays than sub-domains further away, a problem common to all ray tracing techniques. Good scaling can thus only be achieved with a domain decomposition strategy that addresses these two issues simultaneously. The issue of minimal communication through small sub-domain surfaces can be solved with the right choice of how the sub-domains are constructed. The problem of optimal computational load can only be handled with a good work load estimator which we now want to address.

⁴<http://www.mpi-forum.org>

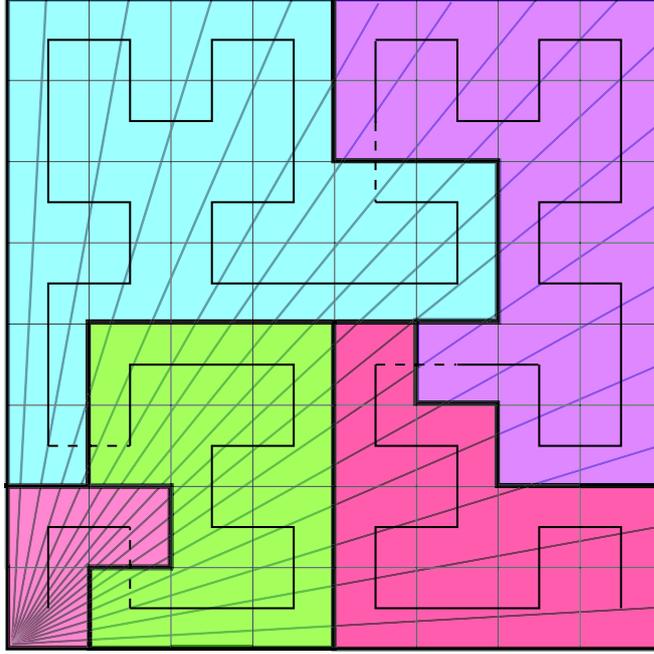


Figure 4.1.: Schematic view of the decomposition strategy with one source in the lower left corner. The ray density decreases with r^{-2} from the source. The Hilbert curve (black) is integrated until the threshold of the expected workload for each core is reached and the domain is cut (indicated by dashed lines). See text for details.

In an optically thin medium the average number of rays to be processed at radius r is proportional to the ray density $\rho_\gamma(r) \propto r^{-2}$, as a consequence of geometric dilution. As we do not have an a-priori knowledge of the evolution of the optical depths distribution throughout a given run, we estimate the ray density in the optically thin approximation and use this as a work load estimator for the domain decomposition.

For each cell at position \mathbf{r}_{cell} and each source, the expected ray density ρ_γ

$$\rho_\gamma(\mathbf{r}_{\text{cell}}) = \sum_{\text{sources}} |\mathbf{r}_{\text{source}} - \mathbf{r}_{\text{cell}}|^{-2} \quad (4.2)$$

is evaluated, where $\mathbf{r}_{\text{source}}$ gives the position vector of the source and \mathbf{r}_{cell} the position of the cell. This yields a direct estimate of the computational time spent in each cell. It has to be stressed that, if for a given run large portions of the volume remain optically thick, the chosen estimator does not produce an optimal load balancing. Nevertheless for these cases it is not possible to predict the evolution of the opacity distribution before running the radiative transfer calculation, so we retain the optically thin approximation as a compromise.

In order to achieve a domain decomposition that requires minimum communication needs (i.e. the surface of the sub-domain is small which reduces the amount of information that needs to be communicated to the neighbours) and assures the sub-domains to be locally confined (i.e. that communication between nodes does not need to be relayed far through the cluster network over multiple nodes), pCRASH2 implements the widely

4. Enabling parallel computing in CRASH

used approach of decomposing along a Peano-Hilbert space filling curve (Teyssier 2002; Springel 2005; Knollmann & Knebe 2009). The Peano-Hilbert curve is used to map the cartesian grid from the set of normal numbers \mathbb{N}_0^3 to a one dimensional array in \mathbb{N}_0^1 . To construct the Peano-Hilbert curve mapping of the domain, we use the algorithm by Chenyang et al. (2008) (see Section 4.6 for details on how the Peano-Hilbert curve is constructed). Then the work estimator is integrated along the space-filling curve, and the curve is cut whenever the integral exceeds $\sum_i \rho_{\gamma,i}/N_{\text{CPU}}$. The sum gives the total work load in the grid and N_{CPU} the number of CPUs used. This process is repeated until the curve is partitioned into consecutive segments and the grid points in each segment are assigned to a subdomain. Because the segmentation of the Peano-Hilbert curve is consecutive, the sub-domains by construction are contiguously distributed on the grid and are mapped to nodes that are adjacent on the MPI topology. The mapping of the sub-domains onto the MPI topology first maps the sub-domains to all the cores on the node, and then continues with the next adjacent node, filling all the cores there. The whole procedure is illustrated in Fig.4.1.

4.3.2. Parallel photon propagation

One of the simplifications used in many ray-tracing radiation transfer schemes is that the number of photons at any position along a ray is only governed by the absorption in all the preceding points. This approximation is generally known as the static approximation to the radiative transfer equation (Abel et al. 1999). In CRASH2 we recursively solve eq. 4.1 until the ray exits the box or all its photons are absorbed. Depending on the length and time scales involved in the simulation, the ray can reach distances far greater than $c \times \Delta t$ and the propagation can be considered to be instantaneous. In such an instant propagation approximation, each ray can pass through multiple sub-domains. At each sub-domain crossing, rays need to be communicated to neighbouring sub-domains. In one time step, there can be multiple such communication events, each enforcing some synchronisation between the different sub-domains. Such a scheme can be efficiently realised with a hybrid characteristics method (Rijkhorst et al. 2006), but has the drawback of a large communication overhead.

To minimise the amount of communication phases per time step, we follow a different approach by truncating the recursive solution of eq. 4.1 at the boundaries of the sub-domains. Instead of letting rays pass through multiple sub-domains in one time-step, rays are only processed until they reach the boundary of the enclosing sub-domain. At the border, they are then passed on to the neighbouring sub-domain. Once received, however, they will not be immediately processed by the neighbour. Propagation of the ray is continued in to the next time step. The scheme is illustrated in Fig. 4.2. In this way, each sub-domain needs to communicate only once per time step with all its neighbours, resulting in a highly efficient scheme, assuming that the computational load is well distributed.

Propagation is thus delayed and a ray needs at most $\approx 2(N_{\text{CPU}})^{1/3}$ time steps to pass through the box (assuming equipartition). The delay with which a ray is propagated through the box depends on the size of the time-step and the number of cores used. The larger the time-step is, the bigger the delay. The same applies for the number of cores.

If the crossing time of a ray in this scheme is well below the physical crossing time,

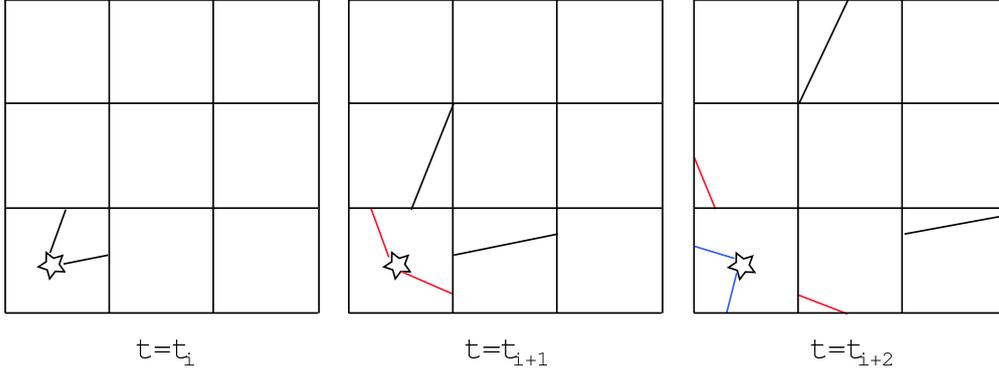


Figure 4.2.: Illustration of rays propagating through the distributed computational domain at three time-steps t_i, t_{i+1} , and t_{i+2} . To simplify the illustration, a box domain decomposition strategy is shown, where each square resembles one sub-domain. Further we assume that the source only emits two rays per time-step. The rays are propagated to the edges of each sub-domain during one time step. Then they are passed on to the neighbouring sub-domain to be processed in the next time step.

the instant propagation approximation is considered retained. However rays will have differing propagation speeds from sub-domain to sub-domain. In the worst case scenario a ray only passes through one cell of a sub-domain. Therefore the minimal propagation speed is $v_{\text{prop,min}} \approx 0.56\Delta l/\Delta t$, where Δl is the size of one cell, Δt the duration of one time step, and the factor 0.56 is the median distance of a randomly oriented ray passing through a cell of size unity as given in Ciardi et al. (2001). If the propagation speed of a packet is $v_{\text{prop,min}} \gg c$, the propagation can be considered instantaneous, as in the original CRASH2. Even if this condition is not fulfilled and $v_{\text{prop,min}} < c$, the resulting ionisation front and its evolution can still be correctly modelled (Gnedin & Abel 2001), if the light crossing time is smaller than the ionisation timescale, i.e. when the ionisation front propagates at velocities much smaller than the speed of light. However the possibility exists, that near to a source, the ionisation timescale is shorter than the crossing time, resulting in the ionisation front to propagate at speeds artificially larger than light (Abel et al. 1999). With the segmented propagation scheme it has thus to be assured that the simulation parameters are chosen in such a way that the light crossing time is always smaller than the ionisation time scale.

The adopted parallelisation scheme can only be efficient, if the communication bandwidth per time step is saturated. Each time two cores need to communicate, there is a fixed overhead needed for negotiating the communication. If the information that is transferred in one communication event is small, the fixed overhead will dominate the communication scheme. It is therefore important to make sure that enough information is transferred per communication event for the overhead not to dominate the communication scheme. The original CRASH2 scheme only allowed for the propagation of one photon packet per source and time-step. In order to avoid the problem described above, this restriction has been relaxed and each source emits multiple photon packets per time step (Partl et al. 2010). Therefore in addition to the total number of photon packet produced per source N_p a new simulation parameter governing the total number of time steps N_t is

4. Enabling parallel computing in CRASH

introduced. The number of packets emitted by a source in one time step is thus N_p/N_t .

As in CRASH2, the choice of the global time step should not exceed the smallest of the following characteristic time scales: ionisation time scale, recombination time scale, collisional ionisation time scale, and the cooling time scale. If this condition is not met, the integration of the ionisation network and the thermal evolution is sub-sampled.

4.3.3. Parallel pseudo-random number generators

Since CRASH2 relies heavily on a pseudo-random numbers generator, here we have to face the challenging issue of generating pseudo-random numbers on multiple CPUs. Each CPU needs to use a different stream of random numbers with equal statistical properties. However this can be limited by the number of available optimal seeding numbers. A large collection of parallel pseudo random number generators is available in the SPRNG library ⁵ (Mascagni & Srinivasan 2000). From the library we are using the Modified Lagged Fibonacci Generator, since it provides a huge number of parallel streams (in the default setting of SPRNG $\approx 2^{39648}$), a large period of $\approx 2^{1310}$, and good quality random numbers. On top each stream returns a distinct sequence of numbers and not just a subset of a larger sequence.

4.4. Performance

In this Section we present the results of an extensive comparison of the parallel scheme with the serial one. We follow the tests described in Maselli et al. (2003) and Iliev et al. (2006). Subsequently we discuss the speed performance of the parallel code and its scaling properties.

4.4.1. Test 0: Convergence test of a pure-H isothermal sphere

First we study the convergence behaviour of a Strömngren sphere as a function of the number of photon packets N_p and the number of time steps N_t . The setup of this test is equivalent to Test 1 in the code comparison project (Iliev et al. 2006). For this test, we distribute the computational domain over 8 CPUs.

The time evolution of an H II region produced by a source having a 10^5 K black-body spectrum with constant intensity expanding in a homogeneous medium is followed. The hydrogen density of the homogeneous medium is fixed at $n_H = 10^{-3} \text{ cm}^{-3}$, with no helium included. The temperature is fixed at $T = 10^4$ K and kept constant throughout the calculation. The initial ionisation fraction is initialised with $x_{\text{H II}} = 1.2 \times 10^{-3}$. The grid has a linear size of $L_{\text{box}} = 6.6 \text{ kpc } h^{-1}$ and is composed of $N_c^3 = 128^3$ cells. The ionising source produces $\dot{N}_\gamma = 5 \times 10^{48} \text{ photons s}^{-1}$. Recombination radiation is followed and photons are emitted whenever 10% of the electrons in a cell have recombined. The simulation time is set at $t_s = 5 \times 10^8 \text{ yr}$ which is approximately four times the recombination timescale.

In Fig. 4.3 we present the resulting spherically averaged profile of the neutral hydrogen fraction where we use $N_p = 10^6, 10^7, 10^8, 10^9$ photon packets. The number of time steps

⁵<http://sprng.cs.fsu.edu/>

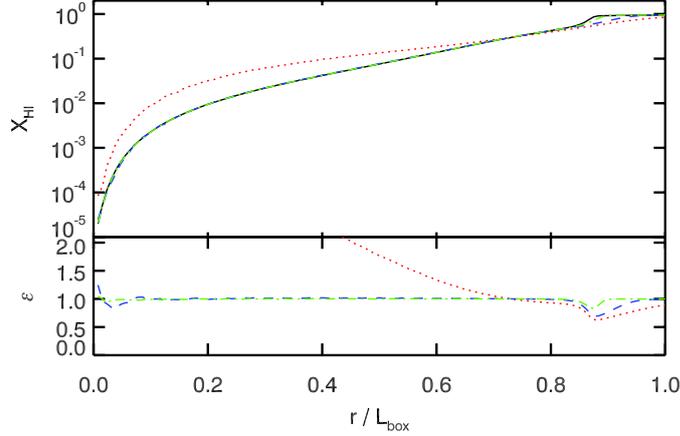


Figure 4.3.: Test 0: Convergence study varying the number of photon packets $N_p = 10^6, 10^7, 10^8, 10^9$ used to simulate the evolution of an H II region. Shown are the spherically averaged neutral hydrogen fraction profiles of the sphere as a function of radius at the end of the simulation. The dotted red line gives $N_p = 10^6$, the blue dashed line $N_p = 10^7$, the green dash dotted line $N_p = 10^8$, and the black solid line $N_p = 10^9$. The test used 8 CPUs. The lower panel shows the relative deviation ϵ of the pCRASH2 run from the highest resolution run.

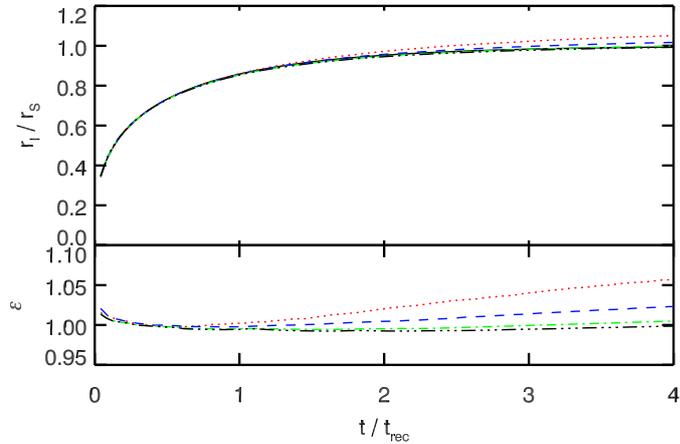


Figure 4.4.: Test 0: Evolution of the H II region's size as a function of time normalised to the recombination time. The upper panel shows results for the convergence study where the number of photon packets was varied. The dotted red line gives $N_p = 10^6$, the blue dashed line $N_p = 10^7$, the green dash dotted line $N_p = 10^8$, and the black dash triple-dotted line $N_p = 10^9$. The solid black curve represents the analytic evolution $r_S(t)$. The test used 8 CPUs. The lower panel shows the relative deviation ϵ of the pCRASH2 runs from the analytic expression.

4. Enabling parallel computing in CRASH

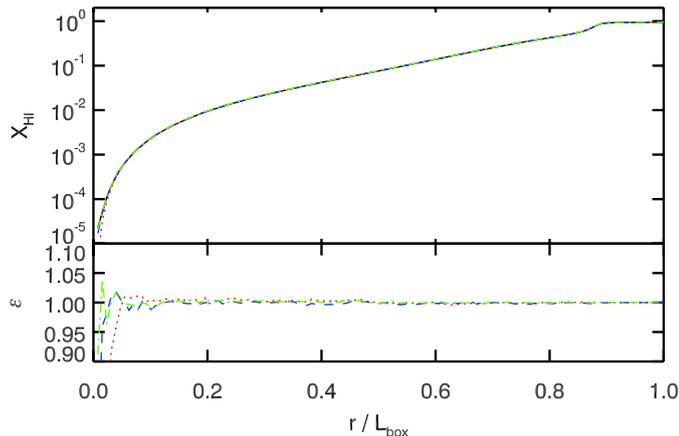


Figure 4.5.: Test 0: Convergence study varying the number of time steps used to simulate the evolution of an H II region. Shown are the spherically averaged neutral hydrogen fraction profile of the sphere as a function of radius at the end of the simulation. The dotted red line gives the results using $N_t = 10^5$ time steps, the blue dashed one $N_t = 10^6$ time steps, the green dash dotted one $N_t = 10^7$, and the black solid line $N_t = 10^8$ time steps. A constant number of $N_p = 10^8$ has been used. The test was run on 8 CPUs. The lower panel shows the relative deviation ϵ of the pCRASH2 run from the highest time resolution run.

N_t is equal to the number of packets in order to obtain the identical scheme used in the serial version of CRASH2. $N_p = 10^6$ packets are unable to reproduce the correct neutral hydrogen fraction profile. Using $N_p = 10^7$ packets instead already yields satisfactory results, however we consider the solution to have converged with at least $N_p = 10^8$ packets. For this case the mean relative deviation $\langle \epsilon(r) \rangle = \langle x_{\text{HI}}(r) / x_{\text{HI,ref}}(r) \rangle$ of the neutral fraction to the $N_p = 10^9$ run is 0.7%. The largest relative deviation is found in a small region in the ionisation front itself. Its width is sensitive to the sampling resolution, where the solution for the $N_p = 10^8$ run differs by 18% from the $N_p = 10^9$ run. In Fig. 4.4 the time evolution of the H II region’s size is given. The radius is calculated using the volume of the H II region, which is determined with $V = \sum_{\text{cells}} x_{\text{HI},i} (\Delta l)^3$, where Δl denotes the physical width of one cell. pCRASH2 is able to resolve the time evolution of the H II region up to 2.5% accuracy for $N_p = 10^7$ when compared with the analytic expression $r_{\text{I}}(t) = r_{\text{S}} (1 - \exp(-t/t_{\text{rec}}))$, where t_{rec} is the recombination time. Deviations from the analytic solution at $4t_{\text{rec}}$ decrease to 0.5% for $N_p = 10^8$ and 0.2% for $N_p = 10^9$. Because of these findings, we use $N_p = 10^9$ photon packets whenever the serial version allows for such a high sampling resolution, otherwise $N_p = 10^8$ packets will be used. We further note that the impact of the sampling resolution on the structure of the ionisation front is very important when the formation and destruction of molecular hydrogen is considered. It can significantly affect the stability of the ionisation front in cosmological simulations, in particular by introducing inhomogeneities in the star formation (Ricotti et al. 2002; Susa & Umemura 2006; Ahn & Shapiro 2007).

For the test where the number of time steps is varied (and thus the number of photons emitted per time step), we use $N_p = 10^8$ packets. The lower sampling resolution is used to reduce the accuracy with which the ionisation network is solved. Any influence of the time step length should be more easily seen if the accuracy in the solution of the

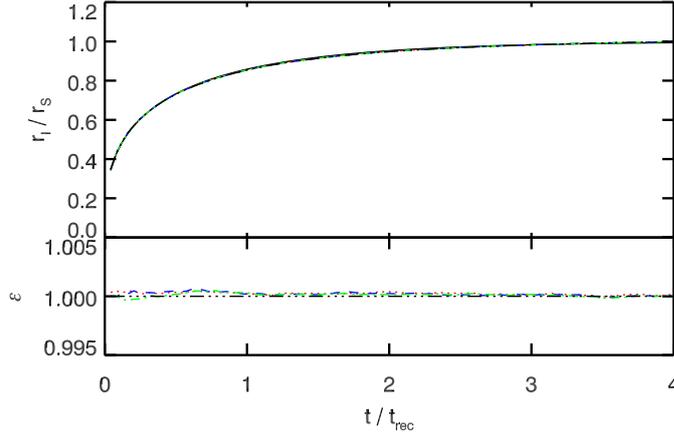


Figure 4.6.: Test 0: Evolution of the H II region’s size as a function of time normalised to the recombination time. The upper panel shows results for the convergence study where the number of time steps used in the simulation was varied. The number of packets used was fixed at $N_p = 10^8$. The dotted red line gives the results using $N_t = 10^5$ time steps, the blue dashed one $N_t = 10^6$ time steps, the green dash dotted one $N_t = 10^7$, and the black dashed triple-dotted line $N_t = 10^8$ time steps. The solid black curve represents the analytic evolution $r_S(t)$. The lower panel shows the relative deviation ϵ of the pCRASH2 runs from the solution using $N_t = 10^8$ time steps.

ionisation network is lower. We evolve the H II region using $N_t = 10^5, 10^6, 10^7, 10^8$ time steps. The resulting neutral hydrogen fraction profiles are given in Fig. 4.5. Overall, varying the number of time steps does not alter the resulting profile significantly. Only in the direct vicinity of the source, the solution is sensitive to the number of time steps used. In Fig. 4.6 the time evolution of the H II region’s radius is shown for this test. The fluctuations between the results for various numbers of time steps are small, when compared to the $N_t = 10^8$. This is especially true for the equilibrium solution. At early times when the H II region grows rapidly, the largest deviations can be seen. However for none of the cases the difference exceeds 0.05%. Since the choice of N_t directly determines the propagation speed of the photon packets, we confirm that the choice of propagation speed is not affecting the resulting H II region and that the new ray propagation scheme is a valid approach.

4.4.2. Comparing pCRASH2 with CRASH2

Test 1: Pure-H isothermal sphere

Now we compare the results obtained from pCRASH2 with those from the serial CRASH2 code. Since pCRASH2 evolves recombination radiation smoother (remember CRASH2 emits recombination photons at specific time steps while pCRASH2 emits diffuse radiation continuously), we expect differences between the two codes in regions where recombination radiation dominates. In order to show that pCRASH2’s solver performs identically to the one in CRASH2, we use the same setup as in Test 0, but we do not allow for recombination radiation to be emitted. For both codes, we set $N_p = 10^9$. The pCRASH2 solution is obtained on 8 CPUs, using $N_t = 10^7$ time steps.

The resulting neutral hydrogen fraction profile is shown in Fig. 4.7. The two curves

4. Enabling parallel computing in CRASH

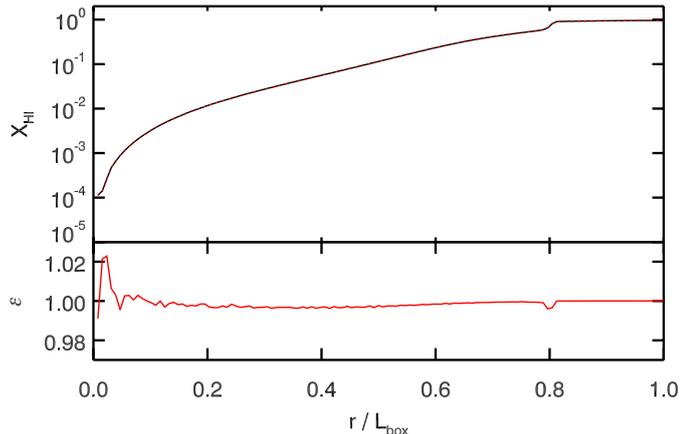


Figure 4.7.: Test 1: Comparing the solvers of pCRASH2 with CRASH2. Shown are the spherically averaged neutral hydrogen fraction profile of the sphere as a function of radius at the end of the simulation. The red dotted line gives the pCRASH2 results and the black solid line the solution obtained with CRASH2. pCRASH2 was run on 8 CPUs. The lower panel shows the relative deviation ϵ of the pCRASH2 run from the serial CRASH2 run.

of the CRASH2 and the pCRASH2 runs are barely distinguishable. The fluctuations in the relative deviations from the CRASH2 results near to the source are due to fluctuations in the least significant decimal. Only at the ionisation front, a small deviation of around 0.5% is seen. In the convergence study we have seen that the radial profile is sensitive to the sampling resolution at the ionisation front (IF) where the partially ionised medium abruptly turns into a completely neutral one. It is as well sensitive to the variance introduced by using different random number generator seeds. Tests using different seeds have shown, that fluctuations of up to 0.4% between the various results can be seen. Therefore the small deviation in the ionisation front stems to some extent from the fact, that pCRASH2 uses a different set of random numbers. The parallel solver thus gives results which are in agreement with the serial version. Including recombination radiation however will change this picture, which we address with the next test.

Test 2: Realistic H II region expansion.

The test we are now focusing on is identical to Test 2 in the radiative transfer codes comparison paper (Iliev et al. 2006), except that we use a larger box size to better accommodate the region preceding the ionisation front where preheating due to higher energy photons is important. We use the same setup as in Test 0, but now we follow the temperature evolution, and self-consistently account for the reemission. Again the gas is initially fully neutral. Its initial temperature however is now set to $T = 100$ K. The source emits $N_p = 4 \times 10^9$ photon packets. Again we compare the solution obtained with CRASH2 to the one obtained with pCRASH2, which for the present test is run with $N_t = 10^7$ time steps on 8 CPUs.

In Fig. 4.8 we show the resulting neutral hydrogen fraction profiles and temperature profiles at three different time steps $t = 1 \times 10^7, 2 \times 10^8, 5 \times 10^8$ yrs. Plotted are the solutions obtained with CRASH2 and with pCRASH2 (solid black and red dotted lines

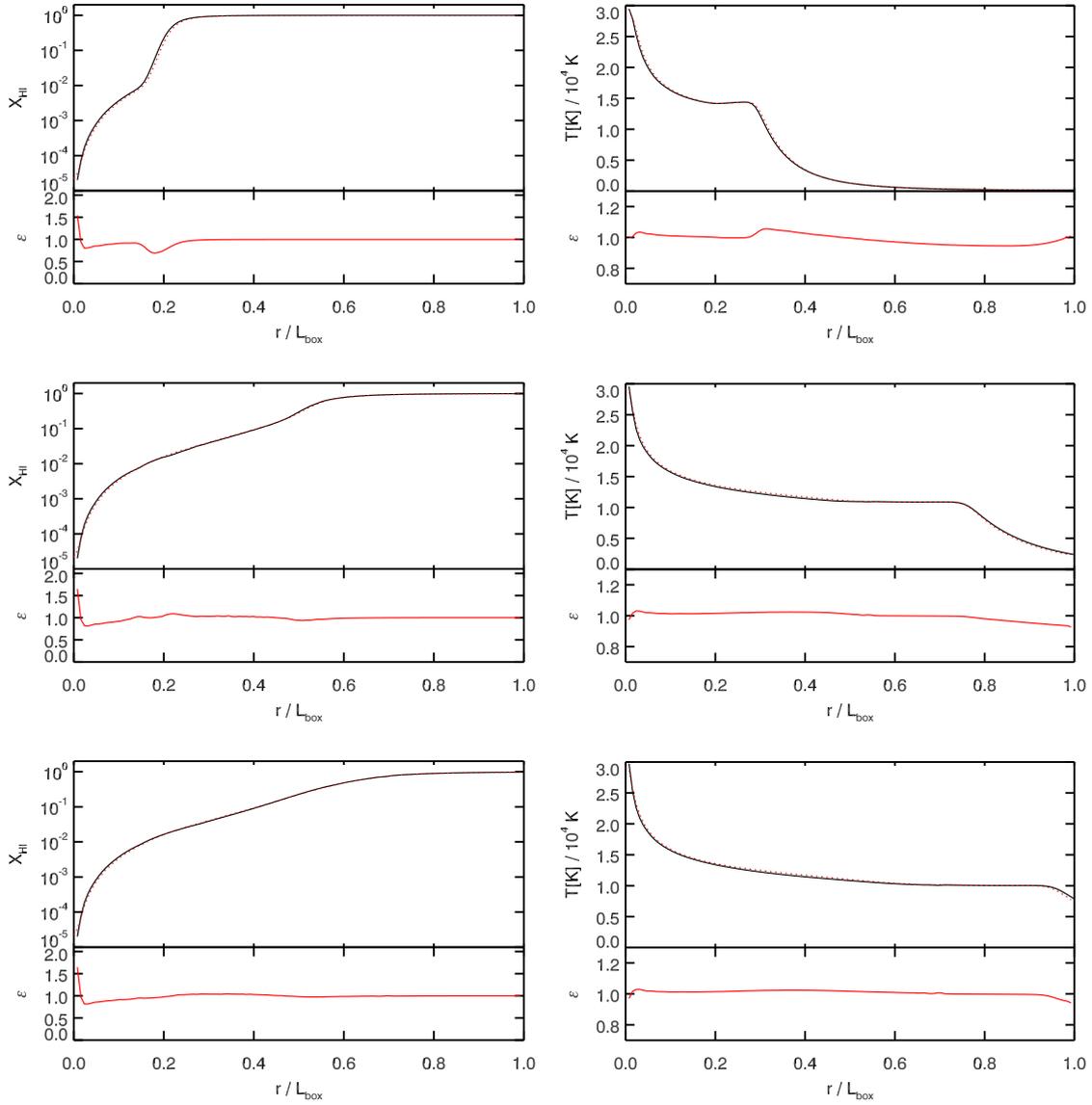


Figure 4.8.: Test 2: Comparing the results of pCRASH2 (red dotted line) with CRASH2 (black line). Shown are the spherically averaged neutral hydrogen fraction profiles of the sphere as a function of radius (left column) and the temperature profiles (right column) at $t = 1 \times 10^7, 2 \times 10^8, 5 \times 10^8$ yr (from top to bottom). The pCRASH2 results are obtained on 8 CPUs. The small bottom panels show the relative deviation from the CRASH2 reference solution.

4. Enabling parallel computing in CRASH

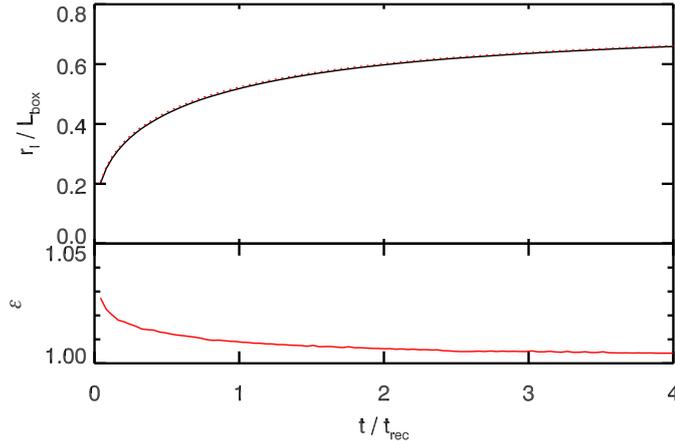


Figure 4.9.: Test 2: Comparing the time evolution of the H II region’s radius from the pCRASH2 run (red dotted line) with CRASH2 (black line). The pCRASH2 results are obtained on 8 CPUs. The small bottom panel show the relative deviation from the CRASH2 reference solution.

respectively). At $t = 1 \times 10^7$ yrs, the results of CRASH2 and pCRASH2 slightly differ at the ionisation front. The structure of the ionisation front appears slightly sharper in the pCRASH2 run, than in the CRASH2 run. In the front, recombination radiation is an important process (Ritzerveld 2005). Since pCRASH2 continuously emits recombination photons whenever a cell reaches the recombination threshold and not only at fixed time steps, the diffuse field is better sampled and the ionisation front becomes sharper. For this test, CRASH2 sampled the diffuse field 10^3 times through the whole simulation. In pCRASH2 the number of cell crossings determines the sampling resolution of the recombination radiation. Therefore recombination radiation originating in the cells of the ionisation front is evaluated at least 5×10^4 times for this specific run. A higher sampling rate in the recombination radiation reproduces its spectral energy distribution which is strongly peaked towards the H I photo-ionisation threshold (see the Milne spectrum) more accurately, resulting in a sharper ionisation front. Better sampling of the Milne spectrum reproduces its shape more accurately and reduces the possibility that too much energy is deposited in the high energy tail of the Milne spectrum due to the poor sampling resolution of the spectrum.

At later times the two codes give similar results for the hydrogen ionisation fraction, with almost negligible differences close to the source and across the ionisation front. This is also seen in the temporal evolution of the H II region’s radius given in Fig. 4.9, which is determined as described in Test 0. In the beginning the differences between the codes are at most 3%, decreasing steadily up to one recombination time. This clearly shows the effect of the smooth emission of recombination radiation in pCRASH2. After one recombination time, the differences between the two codes are small and convergence is achieved. A similar picture emerges for the temperature profile. As for the ionisation front, the preceding temperature front is as well slightly sharper with pCRASH2.

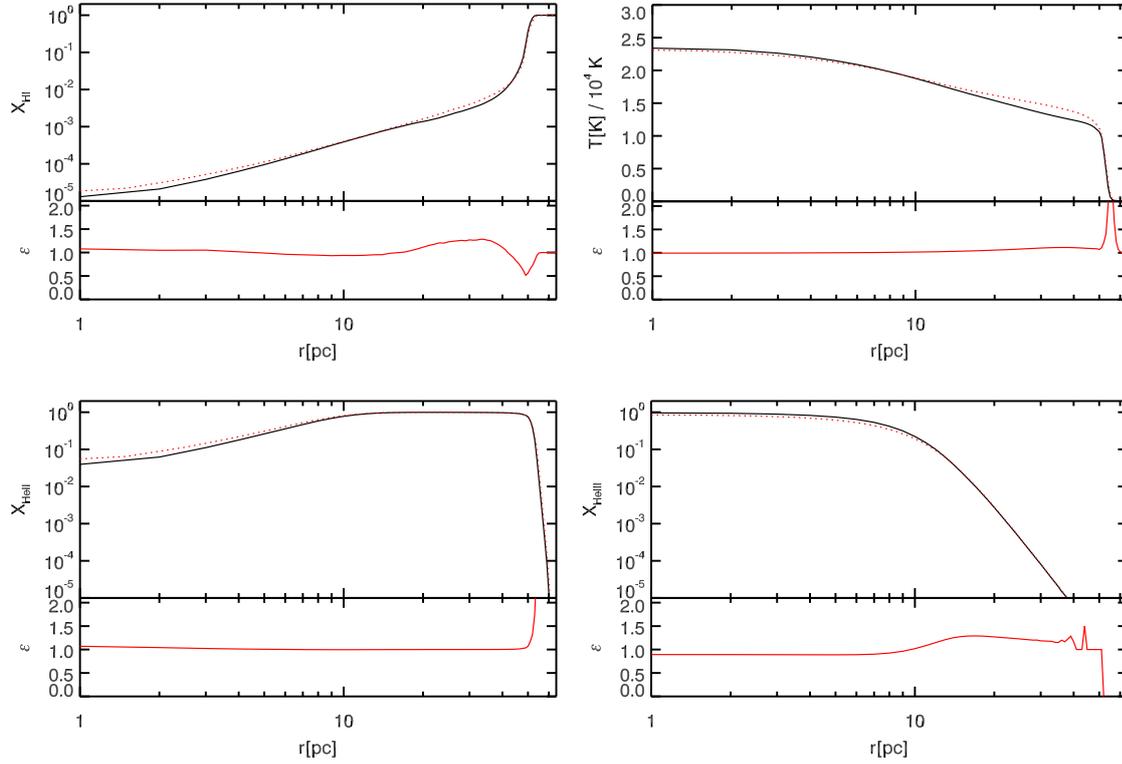


Figure 4.10.: Test 3: Realistic H II region extending in a hydrogen + helium medium. Compared are equilibrium results obtained with CRASH2 (black solid line) to ones obtained with pCRASH2 (red dotted line). The upper left panel gives the profile of neutral hydrogen fractions, lower left panel gives singly ionised helium fractions, and the lower right panel gives the double ionised helium fractions. The upper right panel gives the spherically averaged temperature profile as a function of radius. The small bottom panels show the relative deviation of the pCRASH2 run from the CRASH2 reference solution.

4. Enabling parallel computing in CRASH

Test 3: Realistic H II expansion in a H+He medium

In this test, we study the expansion of an H II region in a medium composed of hydrogen and helium. This corresponds to the CLOUDY94 test in Maselli et al. (2009). We consider the H II region produced by a $T = 6 \times 10^4$ K blackbody radiator with a luminosity of $L = 10^{38}$ erg s⁻¹. The point source ionises a uniform medium with a density of $n = 1$ cm⁻³ in a gas composed of 90% hydrogen number density and the rest helium. The temperature is initially set at $T = 10^2$ K and its evolution is solved for. The dimension of the box is $L_{\text{box}} = 128$ pc. The test is run for $t_s = 6 \times 10^5$ yrs after which ionisation equilibrium is reached with $N_p = 2 \times 10^8$. pCRASH2 was run on 8 CPUs with $N_t = 10^8$.

In Fig. 4.10 we compare the resulting temperature, hydrogen and helium density profiles from pCRASH2 with the ones obtained with CRASH2. Overall the resulting pCRASH2 profiles are in good agreement with CRASH2. Again only slight deviations from the CRASH2 solution are present in the outer parts of the H II region due to higher sampling resolution of the ionising radiation re-emitted in recombinations. As in the previous test, the pCRASH2 solution shows somewhat sharper fronts in hydrogen and temperature. The hydrogen fronts are shifted slightly further away from the source in pCRASH2, giving rise to the large relative error. Otherwise pCRASH2's performance is equivalent to that of CRASH2.

Test 4: Multiple sources in a cosmological density field

The setup of this test is identical to Test 4 in the radiative transfer codes comparison paper (Iliev et al. 2006). Here the formation of H II regions from multiple sources is followed in a static cosmological density field at redshift $z = 9$, including photo-heating. The initial temperature is set to $T = 100$ K. The positions of the 16 most massive halos are chosen to host 10^5 K black-body radiating sources. Their luminosity is set to be proportional to the corresponding halo mass and all sources are assumed to switch on at the same time. No periodic boundary conditions are used. The ionisation fronts are evolved for $t_s = 4 \times 10^7$ yr. Each source produces $N_p = 1 \times 10^7$ photon packets. pCRASH2 is run with $N_t = 10^6$ time steps on 16 CPUs.

A comparison between slices of the neutral hydrogen fraction and the temperatures obtained with pCRASH2 and CRASH2 are given in Fig. 4.11 for $t = 0.05$ Myr and Fig. 4.12 for $t = 0.2$ Myr. For both time steps the pCRASH2 results produce qualitatively similar structures compared to CRASH2 in the neutral fraction and temperature fields. Slight differences however are present, mainly in the vicinity of the ionisation fronts, due to the differences in how recombination is treated, as already discussed. In the lower panels of the same figures we show also the probability distribution functions for the hydrogen neutral fractions and the temperatures. Here the differences are more evident. The neutral fractions obtained with pCRASH2 do not show strong deviations from the CRASH2 solution. In the distribution of temperatures however, pCRASH2 tends to heat up the initially cold regions somewhat faster than CRASH2, as can be seen by the 20% drop in probability for the $t = 0.05$ Myr time step at low temperatures. This can be explained by the fact that the dominant component to the ionising radiation field in the outer parts of H II regions is given by the radiation emitted in recombinations. The better resolution adopted in pCRASH2 for the diffuse field is then responsible for slightly larger

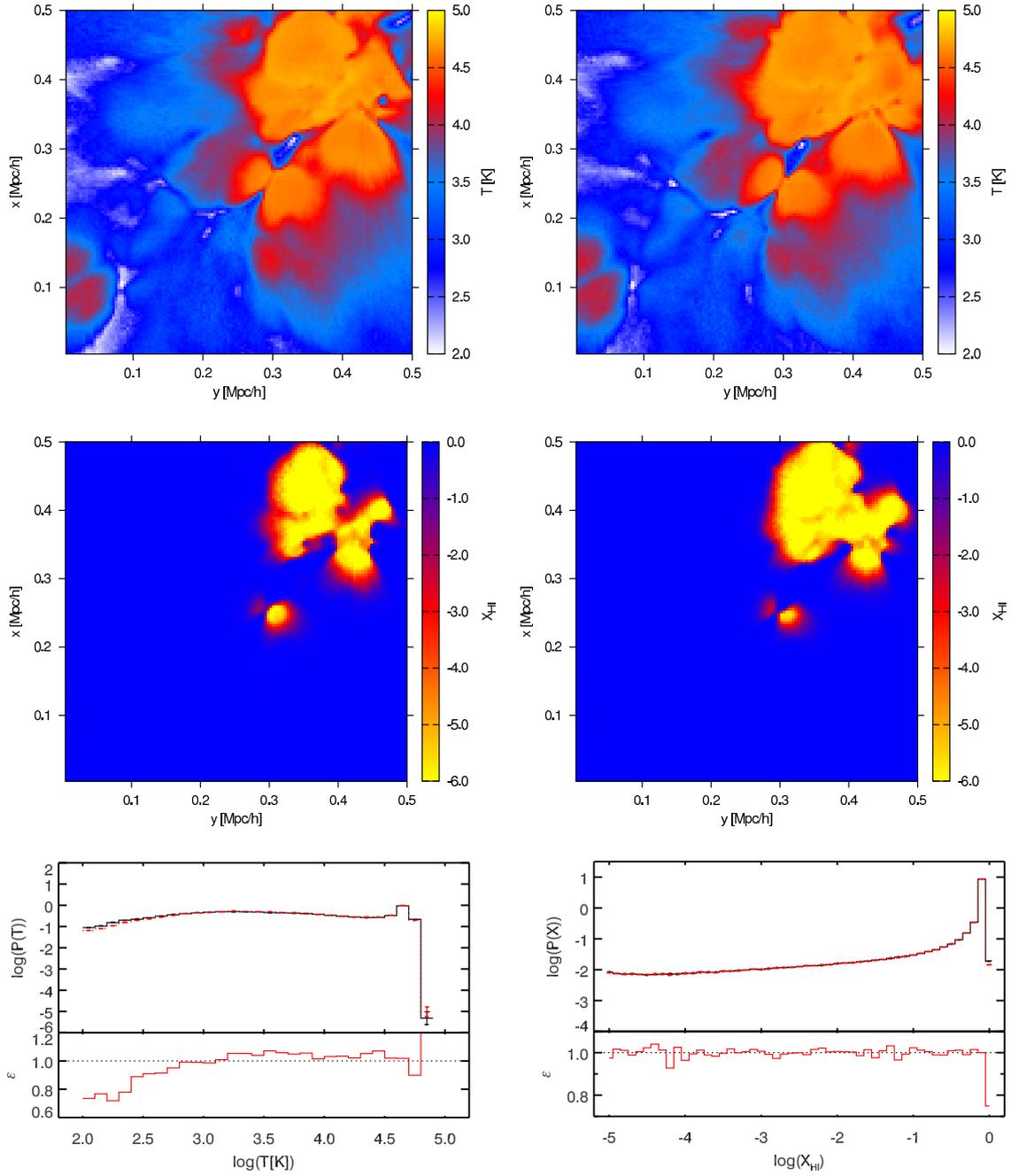


Figure 4.11.: Test 4: Cut through the temperature fields (upper panels) and neutral hydrogen fractions (middle panels) at time-step $t = 0.05$ Myr. Left panels show the solution obtained with CRASH2, right panels show the pCRASH2 solution. The lower panels give volume weighted temperature and neutral fraction probability distribution functions obtained with CRASH2 (black) and pCRASH2 (red dotted), where the error bars give Poissonian errors. The small bottom panels show the relative deviation of the pCRASH2 run from the CRASH2 reference solution.

4. Enabling parallel computing in CRASH

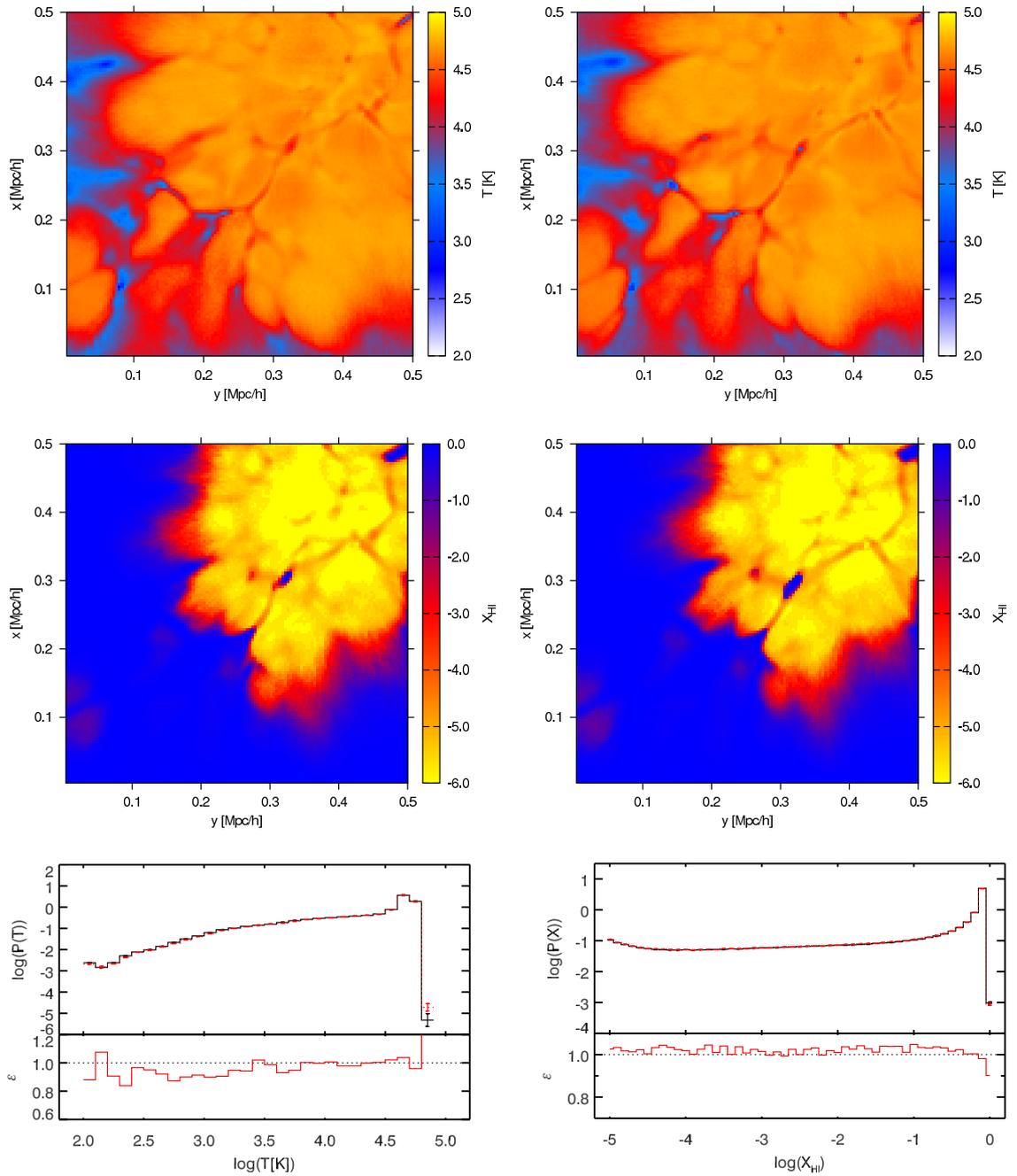


Figure 4.12.: Test 4: Same as in Fig. 4.11 at time-step $t = 0.2$ Myr.

H II regions. Since a larger volume is ionised and thus hot, the fraction of cold gas is smaller in the pCRASH2 run, than in the reference run. At higher temperatures however, the distribution functions match each other well. Only in the highest bin a discrepancy between the two code’s solutions can be noticed, but given that there are at most four cells contributing to that bin, this is consistent with Poissonian fluctuations.

4.4.3. Scaling properties

By studying the weak and strong scaling properties of a code, it is possible to assess how well a problem maps to a distributed computing environment and how much speed increase is to be expected from the parallelisation. Strong scaling describes the scalability by keeping the overall problem size fixed. Weak scaling on the other hand refers to how the scalability behaves when only the problem size per core is kept constant (i.e. the fraction between the total problem size and the number of CPUs stays constant).

We will use the test cases described above, to study the impact of the various input parameters such as grid size, number of photon packets, number of sources, and type of physics used, on the scalability. Scalability is measured with the speed up, which is defined as the execution time on a single core divided by the execution time on multiple cores. Due to the long simulation times of our test cases on a single core (for some test cases of the order of weeks), the execution time on a single core was only determined once. This certainly makes our normalisation point subject to variance, which has to be taken into account in the following discussion. All results presented in this section have been obtained on a cluster at the Leibniz-Institut für Astrophysik Potsdam (AIP) with computational nodes consisting of two Intel Xeon Quad Core 2.33GHz processors each. Since our parallelisation scheme has shortly spaced communication points, very good communication latency is required, which on our system is guaranteed through an InfiniBand network.

Test 2

Achieving good scaling for Test 2 is a challenge, since only one source sits in a corner of the box and a strong load imbalance between the sub-domains near to the source and the ones further away exists. The fact that the gas is initially neutral even intensifies the imbalance, since at first the domains far away from the source will be idle. Only as the H II region grows, more and more sub-domains will be involved in the calculation. This test is therefore a good proxy for how the code scales in extreme situations and shows the strong scaling relation for one source. In Fig. 4.13 we show the results obtained with the setup described above. In principle it is expected, that the more rays each domain has to process in one time-step, the better the scaling becomes. Therefore we run the test once without recombination emission and once with. Further we increased the grid size to 256^3 and 512^3 cells, which increases the amount of calculations needed per sub-domain. This yields the weak scaling properties for one source as a function of grid size.

The scaling relation is mainly determined by Ahmdal’s law (Amdahl 1967), which relates the parallelised parts of a code (in our case the propagation of photons) with the unparallelised parts (such as communication between sub-domains). It states that the maximum possible speed up is solely limited by the fraction of time spent in serial parts

4. Enabling parallel computing in CRASH

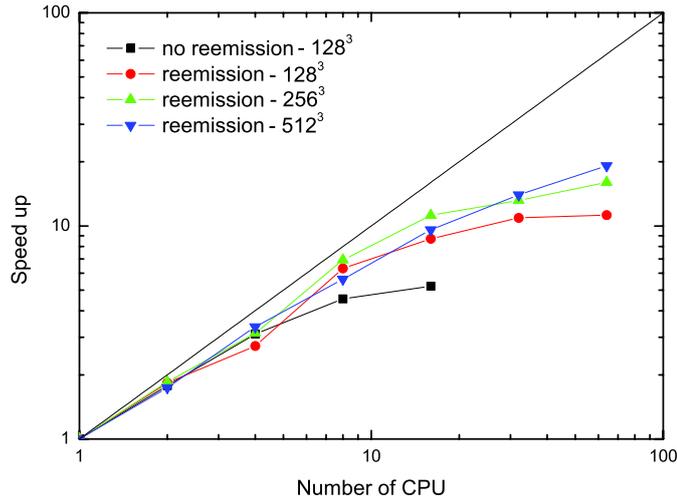


Figure 4.13.: Speed up achieved on our local cluster for Test 2 as a function of grid-size and number of CPUs. The straight black line indicates a linear ideal scaling property.

of the code. The scaling relation behaves for all runs similarly. An almost linear speedup is achieved up to 8 CPUs per node, and a deviation from linear scaling according to Amdal’s law arises when communication over the network is needed. The deviation is dependent on the amount of work to be processed per time-step.

In the worst case of a 128^3 celled grid without reemission, the deviation is the largest. Following recombination emission and thus increasing the workload yields better scaling up to the point where one computational node is saturated. For one source the best scaling is achieved with larger grid sizes. For the 512^3 grid it is even reasonable to use 16 CPUs on two nodes on our cluster and allow for communication of rays over the network. However with 32 CPUs, 50% of the computational resources remain unused.

Test 3

Up to now, we have only considered the gas to be made up solely of hydrogen. We now study the scaling properties of one source embedded in gas made of a mixture of hydrogen and helium. Now each domain needs to solve a more complicated set of equations and more time is spent in solving the chemical-thermal equations network than in propagating photons. The ionisation network needs to track the two additional species introduced with helium, plus the related contribution to heating and cooling which enters the temperature calculation. Since solving the ionisation-thermal network is the most computationally expensive part of our algorithm, the additional work results in an approximate increase of overall execution time by a factor of three or more. Furthermore recombining helium will produce additional photons that need to be followed, even further increasing the computational load. As the communication demand stays the same as in Test 2, the scaling relation is expected to be better than in the hydrogen only case. In Fig. 4.14 we present the scaling performance obtained with the Cloudy test case described above using $N_p = 2 \times 10^9$ photon packets and $N_t = 10^7$ time-steps. Comparing the scaling with the one achieved in Test 2, it is evident, that solving more detailed physics improves

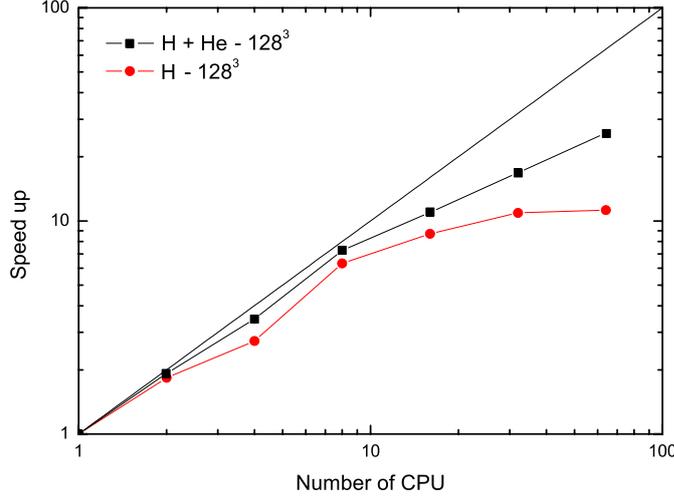


Figure 4.14.: Test3: Scaling relation of the Cloudy test case as a function of number of CPUs. For comparison, the results of the hydrogen only Test 2 are given as red line.

scaling. Now for the Cloudy test, ideal scaling is achieved as long as the problem is confined to one node (in our case 8 CPU) and no information needs to be communicated over the network. However the fact that more time is spend in solving the rate equations, communication latency across the network does not affect scaling as strongly as in Test 2. It now even makes sense to use 32 CPUs for one source, yielding a speed up of around 16. As a comparison Test 2 achieved a speed up of around 10 for 32 CPUs.

As a final check we have also run a set of simulations for Test 3 with an increased gas density, $n_H = 10 \text{ cm}^{-3}$, all the other quantities being the same. The aim of this experiment is to test the scaling under conditions in which diffuse radiation from recombinations becomes important. Somewhat to our surprise, we do not find any appreciable differences in the scaling relation between the low- and high-density cases. However, this can be easily understood as follows: The simulation time is set at five times the recombination time. Hence the same amount of recombinations per time step occur (when normalised to the density of the gas) in both runs. Since recombination photons are produced when a certain fraction of the gas has recombined, the number of emitted recombination photons is similar in the two runs. Therefore the amount of CPU time spent in the diffuse component is similar and the scaling does not change.

Test 4

By studying the scaling properties of Test 4, we can infer how the code scales with increasing number of sources. Since Test 4 uses an output of a cosmological simulation, the sources are not distributed homogeneously in the domain. Therefore large portions of the grid remain neutral as is seen in Fig. 4.12. This poses a challenge to our domain decomposition strategy and might deteriorate scalability.

First we study the scaling of the original Test 4, with a sample of 10^8 photon packets emitted by each of the 16 sources. Then we increase the number of photons per source to 10^9 . Since large volumes remain neutral, we further study an idealised case, where the

4. Enabling parallel computing in CRASH

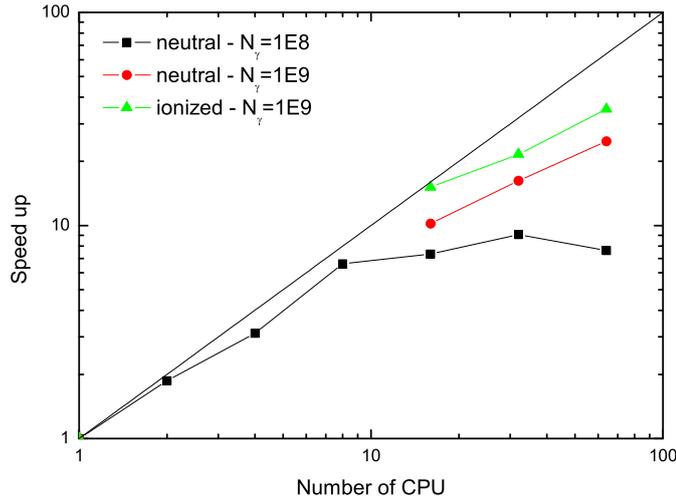


Figure 4.15.: Scaling achieved with Test 4 as a function of number of photon, optically thickness and number of CPUs.

whole box is kept highly ionised by initialising every cell to be 99% ionised. The results of these experiments are shown in Fig. 4.15. In the case of only 10^8 photon packets per source, good scaling can only be reached up to 8 CPUs (i.e. a single node), as was the case with only one source. Communication over the network cannot be saturated with such a small number of photons and its overhead is larger than the time spent in computations. However by increasing the number of samples per time-step improves scaling. In the idealised optically thin case, perfect scaling is reached even up to 16 CPUs and starts to degrade for higher numbers of CPU. The discrepancy between the neutral and fully ionised case shows the influence of load imbalance due to the large volumes of remaining neutral gas.

Since good scaling was found in the original Test 4, we now increase the number of sources. For this we duplicate the sources in Test 4, and mirror their position at an arbitrary axis. This preserves the clustered nature of the sources, while distributing them throughout the computational volume. With this prescription we increase the number of sources to 32 and 64. Again each source emits 10^9 photon packets and the idealised case of an optically thin box is studied. The results of the speed up function are given in Fig. 4.16. As expected, the more computational intensive the problem becomes, the better the scaling. In the figure, the speed up for the 64 sources lies below the others cases because we have linearly extrapolated the normalisation point of the 32 sources run to the 64 sources case, since the time to run the simulation on one CPU was too long. This of course is just a rough estimate, and in fact it is responsible for the lower speed up values. Despite of it, the trend still gives a measure of the improved scaling with respect to the cases with 32 and 16 sources. While a difference in the speed up between the three test-cases is seen when increasing from 1 to 16 CPUs, it is difficult to tell whether this is a genuine feature of our parallelisation strategy or just a manifestation of the variance in the normalisation.

Up to now we have only considered the case of a 128^3 box. We now re-map the density field of Test 4 to 256^3 and 512^3 grids. The resulting scaling properties for the ideal

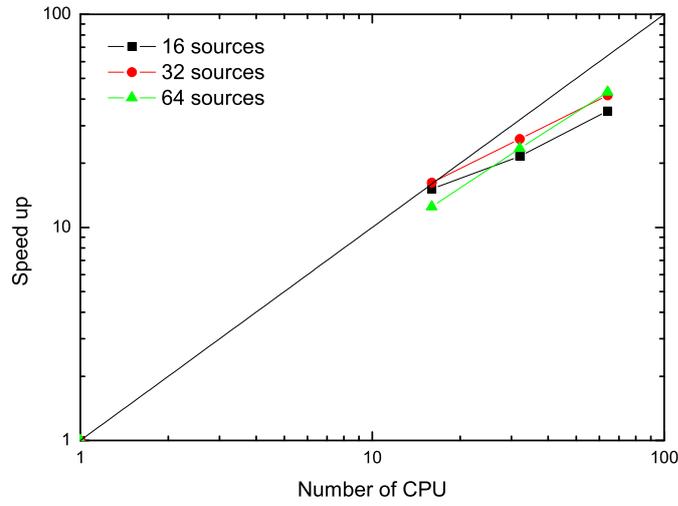


Figure 4.16.: Scaling achieved with Test 4 as a function of number of sources and number of CPUs for an ideal optically thin case.

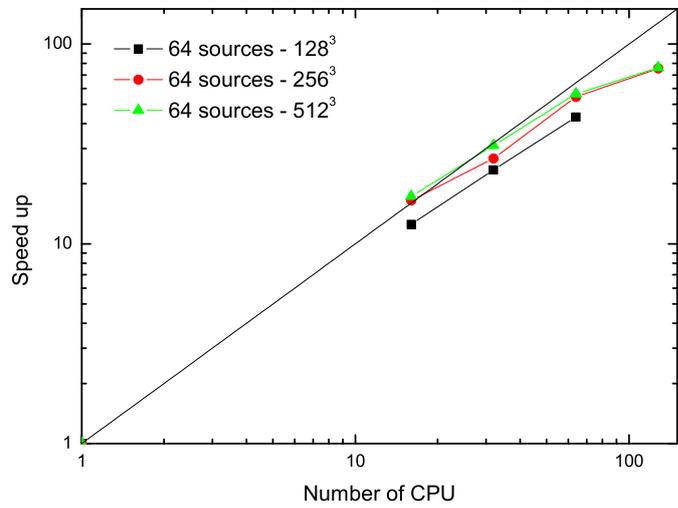


Figure 4.17.: Scaling relation of Test 4 as a function of box size with 64 sources and number of CPUs for an ideal optically thin case.

4. Enabling parallel computing in CRASH

optically thin case with 64 sources are presented in Fig. 4.17. The scaling properties are not dependent on the size of the grid, and weak scaling only exists as long as the number of cores does not exceed the number of sources.

We can thus conclude that our parallelisation strategy shows perfect weak scaling properties when increasing the numbers of sources (compare Fig. 4.16). Increasing the size of the grid does not affect scaling. However weak scaling in terms of the grid size only works as long as the number of CPUs does not exceed the number of sources (see Fig. 4.17 going from 64 to 128 CPUs).

From these tests, we can formulate an optimal choice for the simulation setup. We have seen that better scaling is achieved with large grids. Further linear scaling can be achieved in the optically thin case by using up to as many cores as there are sources. For the optically thick case however, deterioration of the scaling properties needs to be taken into account.

4.4.4. Dependence of the solution on the number of cores

Since each core has its own set of random numbers, the solutions of the same problem obtained with different numbers of cores will not be identical. They will vary according to the variance introduced by the Monte Carlo sampling. To illustrate the effect, we revisit the results of Test 2 and study how the number of cores used to solve the problem affects the solution.

The results of this experiment are shown in Fig. 4.18, where we compare the different solutions obtained with various numbers of CPUs. Solely by looking at the profiles, no obvious difference between the runs can be seen. Variations can only be seen in the relative differences of the various runs which are compared with the single CPU pCRASH2 run.

At $t = 10^7$ yr the runs do not show any differences. In Sec. 4.4.2 we have seen, that recombination radiation has not yet started to be important. This is exactly the reason why, at this stage of the simulation, no variance has developed between the different runs. Since the source is always handled by the first core and the set of random number is always the same on this core no matter how many CPUs are used, the results are always identical. However as soon as the Monte Carlo sampling process start to occur on multiple nodes, the set of random numbers starts to deviate from the single CPU run and variance in the sampling is introduced. At the end of the simulation at $t = 5 \times 10^8$ yr, large parts of the H II region are affected by the diffuse recombination field and variance between the different runs is expected. By looking at Fig. 4.18, the Monte Carlo variance for the neutral hydrogen fraction profile lies between 1% and 2%. The temperature profile is not as sensitive to variance as the neutral hydrogen fraction profile. For the temperature the variance lies at around 0.1%.

4.4.5. Thousand sources in a large cosmological density field

Up to now we have discussed pCRASH2's performance with controlled test cases. We now demonstrate pCRASH2's ability to handle large highly resolved cosmological density fields embedding thousands of sources. We utilise the $z = 8.3$ output of the *MareNostrum High-z Universe* (Forero-Romero et al. 2010) which is a 50 Mpc h^{-1} SPH simulation

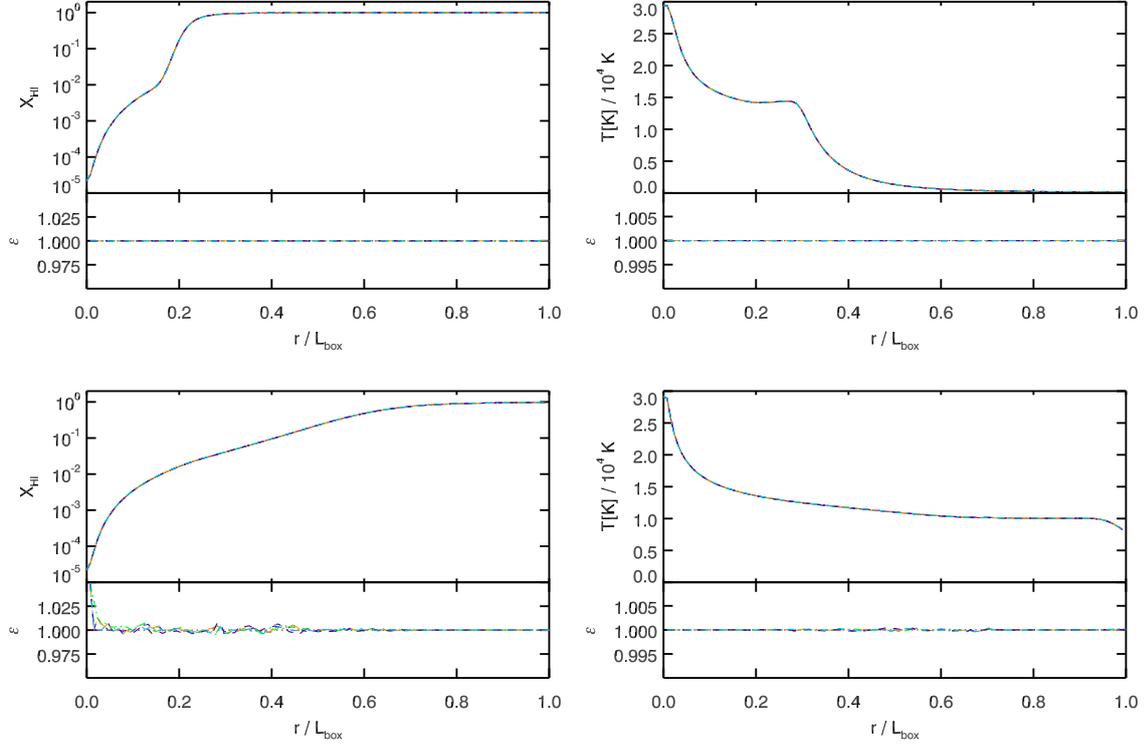


Figure 4.18.: Test 2: Dependence of the solution on the number of cores. The large upper panels show the results of pCRASH2 obtained on 1 CPU (red dotted line), 2 CPUs (blue dashed line), 4 CPUs (green dash dotted line), 8 CPUs (orange dash dot dotted line), 16 CPUs (violet dashed line), and 32 CPUs (light blue dash dotted line). The results are compared with the CRASH2 solution (black solid line). Shown are the spherically averaged neutral hydrogen fraction profiles of the sphere as a function of radius (left column) and the temperature profiles (right column) at $t = 1 \times 10^7$ (top), and 5×10^8 yr (bottom). The small bottom panels show the relative deviation from the single CPU pCRASH2 solution.

4. Enabling parallel computing in CRASH

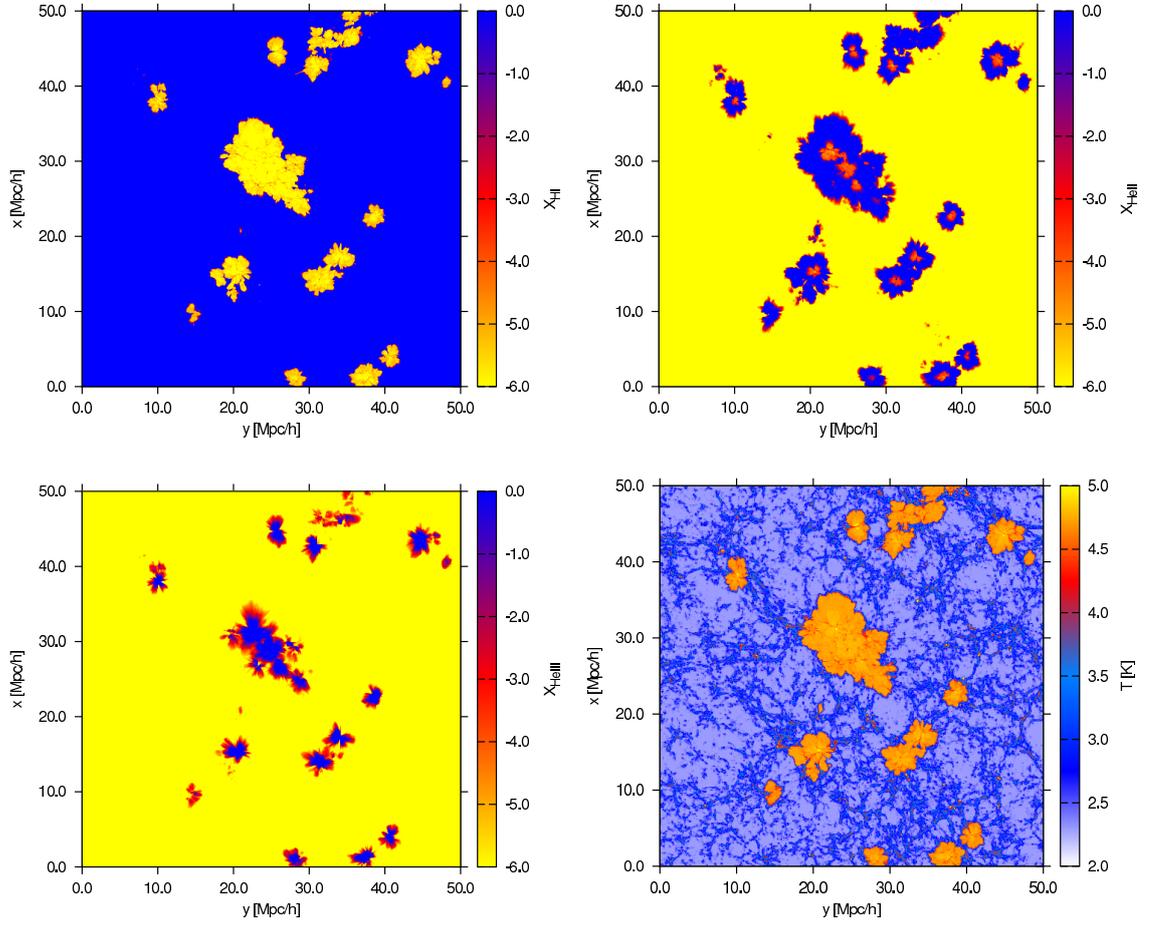


Figure 4.19.: Cut through a cosmological simulation at $z = 8.3$ with the neutral hydrogen fraction field (upper left panel), the singly ionised helium fraction (upper right panel), the doubly ionised helium fraction (lower left panel), and the temperature field (lower right panel) at time-step $t = 1 \times 10^6$ yr using 1000 sources.

using 2×1024^3 particles equally divided into dark matter and gas particles. The gas density and internal energy are assigned to a 512^3 grid using the SPH smoothing kernel. A hydrogen mass fraction of 76% and a helium mass fraction of 24% are assumed.

The UV emitting sources are determined similarly to the procedure used in Test 4 of the comparison project (Iliev et al. 2006) by using the 1000 most massive haloes in the simulation. We evolve the radiation transport simulation for 10^6 yr and follow the hydrogen and helium ionisation, as well as photo-heating. Recombination emission is included. Each source emits 10^8 photon packets in 10^7 time steps. In total 10^{11} photon packets are evaluated, not counting recombination events. The simulation was run on a cluster using 16 nodes, with a total of 128 cores. The walltime for the run as reported by the queuing system was 143.5 hours; the serial version would have needed over two years to finish this simulation.

In Fig. 4.19 cuts through the resulting ionisation fraction fields and the temperature field are shown at $t = 10^6$ yr. It can be clearly seen that the H II regions produced by the different sources already overlap each other. Further the ionised regions strongly deviate from spherical symmetry. This is caused by the fact that the ionisation fronts propagate faster in underdense regions than in dense filaments.

The distribution of He II follows the distribution of ionised hydrogen, except in the centre of the ionisation regions, where helium becomes doubly ionised and holes start to emerge in the He II maps. Photo-heating increases the temperature in the ionised regions to temperatures of around 30000 K.

These results are presented here as an example of highly interesting problems in cosmological studies which can be easily addressed with `pCRASH2`, but that would be impossible to handle with the serial version of the code `CRASH2`.

4.5. Summary

We have developed and presented `pCRASH2`, a new parallel version of the `CRASH2` radiative transfer scheme code, whose description can be found in Maselli et al. (2003), Maselli et al. (2009) and references therein. The parallelisation strategy was developed to map the `CRASH2` algorithm to distributed memory machines, using the MPI library.

In order to obtain an evenly load balanced parallel algorithm, we statically estimate the computational load in each cell by calculating the expected ray number density assuming an optically thin medium. The ray density in a cell is then inversely proportional to the distance to the source squared. Using the Peano-Hilbert space filling curve, the domain is cut into sub-domains; the integrated ray number density in the box is determined and equally divided by the number of processors. Then the expected ray number density is integrated along the curve and cut, whenever this fraction of the total ray number density is reached. The result of this is a well balanced domain decomposition, that even performs adequately in optically thick regimes.

For the parallelisation of the ray tracing itself, we have segmented the propagation of rays over multiple time steps. In the original `CRASH2` implementation it was assumed that photons instantly propagate through the whole box in one time step. In `pCRASH2` however rays are only propagated through one sub-domain per time step. After rays have propagated to the border of the sub-domain, they are then passed on to the neighbouring

4. Enabling parallel computing in CRASH

domain for further processing in the following time step. With this segmentation strategy we keep communication between the domains low and locally confined.

Since pCRASH2 is able to handle a larger amount of photon packets than the serial version, we have additionally increased the sampling resolution of the diffuse recombination radiation. With this parallelisation strategy, we obtain a high-performing algorithm, that scales well for large problem sizes. We have extensively tested pCRASH2 against a standardised set of test cases to validate the parallelisation scheme.

With pCRASH2 it is now possible to address a number of problems that require the tracking of the UV field produced by a large number of sources, that so far were not attainable for CRASH2 because of the high computational cost. A natural application is the evolution of the UV flux field, especially during the era of reionisation. The phase transition of the initially neutral H+He intergalactic to its fully ionised state can now be accurately followed in the large volumes required to achieve a good representation of the process on cosmological scales ($> 300\text{Mpc}$).

On the other hand, pCRASH2 has a broader application range. In fact, it is not limited to tackle cosmological configurations, and can be applied to study a wealth of astrophysical problems, e.g. the propagation and impact of UV flux field in the vicinity of star forming galaxies, to mention one example. Finally, due to pCRASH2's distributed memory approach, new physics such as extending the ionisation network to species other than hydrogen and helium can be now easily implemented.

4.6. Generating the Peano-Hilbert curve

Space filling curves present an easy method to systematically reach every point in space and are usually fractal in nature. Various space filling curves such as the Peano curve, the Z-order, or the Peano-Hilbert curve exist. The most optimal in terms of clustering (i.e. that all points on the curve are spatially near to each other) is the Peano-Hilbert curve. We will now sketch the fast Peano-Hilbert algorithm used by the domain decomposition algorithm for projecting cells on the grid which is a subset of the natural numbers including zero \mathbb{N}_0^3 onto an array in \mathbb{N}_0^1 . A detailed description of its implementation is found in Chenyang et al. (2008).

Let H_m^N , ($m \geq 1, N \geq 2$) describe an N -dimensional Peano-Hilbert curve in its m th-generation. H_m^N thus maps \mathbb{N}_0^N to \mathbb{N}_0^1 , where we call the mapped value in \mathbb{N}_0^1 a Hilbert-key. A m th-generation Peano-Hilbert curve of N -dimension is a curve that passes through a hypercube of $2^m \times \dots \times 2^m = 2^{mN}$ in \mathbb{N}_0^N . For our purpose we only consider $N \leq 3$.

Let the 1st-generation Peano-Hilbert curve be called a N -dimensional Hilbert cell C^N (see Fig. 4.20 for $N = 2$). In binary digits, the coordinates of the C^2 [Hilbert-key] Hilbert cell can be expressed as $C^2[0] = 00$, $C^2[1] = 01$, $C^2[2] = 11$, and $C^2[3] = 10$, where each binary digit represents one coordinate X_N in \mathbb{N}_0^N with the least significant bit at the end, i.e. $C^2[i] = X_2X_1$. For $N = 3$ the Hilbert cell becomes $C^3[0] = 000$, $C^3[1] = 001$, $C^3[2] = 011$, $C^3[3] = 010$, $C^3[4] = 110$, $C^3[5] = 111$, $C^3[6] = 101$, $C^3[7] = 100$.

The basic idea in constructing an algorithm that maps \mathbb{N}_0^N onto the m th-generation Peano-Hilbert curve is the following. Starting from the basic Hilbert cell, a set of coordinate transformations which we call Hilbert genes is applied m -times to the Hilbert cell and the final Hilbert-key is obtained. An illustration of this method is shown in

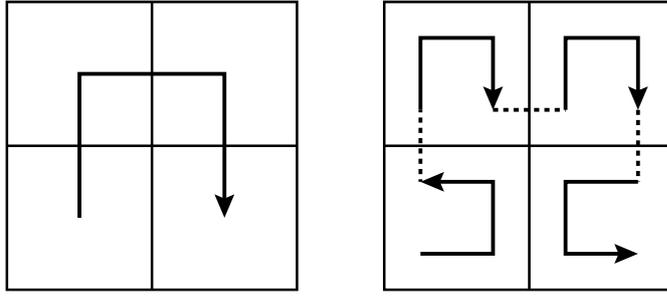


Figure 4.20.: Left panel: The Hilbert cell $C^2 \equiv H_1^2$ for $N = 2$. Right panel: Applying the $N = 2$ Hilbert genes on the Hilbert cell produces the 2nd-generation Peano-Hilbert curve.

Fig. 4.20. Here the method of extending the 1st-generation 2-dimensional Peano-Hilbert curve to the 2nd-generation is shown. Analysing the properties of the Peano-Hilbert curve reveals that two types of coordinate transformations operating on the Hilbert cell are needed for its construction. The exchange of coordinates and the reverse operation. The exchange operation $X_2 \leftrightarrow X_3$ on 011 would result in 101. The reverse operation denotes a bit-by-bit reverse (e.g. reverse X_1, X_3 on 010 results in 111). Using these transformations and the fact, that the m th-generation Hilbert-key can be extended or reduced to an $(m + 1)$ - or $(m - 1)$ th-generation Hilbert-key by bit-shifting the key by N bits up or down, the Hilbert-key can be constructed solely using binary operations.

For the $N = 2$ case illustrated in Fig. 4.20 the Hilbert genes $G^N[\text{Hilbert-key}]$ are the following: $G^2[0] = \text{exchange } X_1 \text{ and } X_2$, $G^2[1] = \text{no transformation}$, $G^2[2] = \text{no transformation}$, and $G^2[3] = \text{exchange } X_1 \text{ and } X_2 \text{ plus reverse } X_1 \text{ and } X_2$. Through repeated application of these transformations on the m th-generation curve, the $(m + 1)$ th-generation can be found.

The Hilbert genes for $N = 3$ are: $G^3[0] = \text{exchange } X_1 \text{ and } X_3$, $G^3[1] = \text{exchange } X_2 \text{ and } X_3$, $G^3[2] = \text{no transformation}$, $G^3[3] = \text{exchange } X_1 \text{ and } X_3 \text{ plus reverse } X_1 \text{ and } X_3$, $G^3[4] = \text{exchange } X_1 \text{ and } X_3$, $G^3[5] = \text{no transformation}$, $G^3[6] = \text{exchange } X_2 \text{ and } X_3 \text{ plus reverse } X_2 \text{ and } X_3$, and $G^3[7] = \text{exchange } X_1 \text{ and } X_3 \text{ plus reverse } X_1 \text{ and } X_3$.

Probleme kann man niemals mit derselben Methode lösen,
durch die sie entstanden sind.

Albert Einstein

5

Cosmological radiative transfer for the line-of-sight proximity effect¹

Aims. We study the proximity effect in the Ly α forest around high redshift quasars as a function of redshift and environment employing a set of 3D continuum radiative transfer simulations.

Methods. The analysis is based on dark-matter-only simulations at redshifts 3, 4, and 4.9 and, adopting an effective equation of state for the baryonic matter, we infer the HI densities and temperatures in the cosmological box. The UV background (UVB) and additional QSO radiations with Lyman limit flux of $L_{\text{LWL}} = 10^{31}$ and $10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ are implemented with a Monte Carlo continuum radiative transfer code until an equilibrium configuration is reached. We analyse 500 mock spectra originating at the position of the QSO in the most massive halo, in a random filament, and in a void. The proximity effect is studied using flux transmission statistics, in particular with the normalised optical depth $\xi = \tau_{\text{eff, QSO}}/\tau_{\text{eff, Ly}\alpha}$, which is the ratio of the effective optical depth in the spectrum near the quasar to that in the average Ly α forest.

Results. Beyond a radius of $r > 1 \text{ Mpc } h^{-1}$ from the quasar, we measure a transmission profile consistent with geometric dilution of the QSO ionising radiation. A departure from geometric dilution is only seen when strong absorbers transverse the line-of-sight. The cosmic density distribution around the QSO causes a large scatter in the normalised optical depth. The scatter decreases with increasing redshift and increasing QSO luminosity. The mean proximity effect identified in the average signal over 500 lines of sight provides an average signal that is biased by random large-scale density enhancements on scales up to $r \approx 15 \text{ Mpc } h^{-1}$. The distribution of the proximity effect strength, a parameter that describes a shift in the transmission profile with respect to a fiducial profile, provides a measurement of the proximity effect along individual lines of sight. It shows a clear maximum almost without any environmental bias. This maximum can therefore be used as an unbiased estimate of the UVB. Different spectral energy distributions between the QSO and the UVB modify the profile but this can be reasonably well corrected analytically. A few Lyman limit systems have been identified that prevent the detection of the proximity effect because of shadowing.

5.1. Introduction

The baryon content of intergalactic space is responsible for the large number of absorption lines observed in the spectra of high redshift quasars. This phenomenon, also known as the Ly α forest (Sargent et al. 1980; Rauch 1998), is mainly attributed to intervening HI clouds along the line of sight (LOS) towards a QSO. The majority of the HI Ly α

¹Accepted in A&A for publication as Part I, A.M.; Dall’Aglio, A.; Müller, V.; and Hensler, G. A&A 2010.

5. Cosmological radiative transfer for the LOS proximity effect

systems are optically thin to ionising radiation. The gas is kept in a high ionisation state by the integrated emission of ultra violet photons originating in the overall populations of quasars and star-forming galaxies: the ultra violet background field (UVB: Haardt & Madau 1996; Fardal et al. 1998; Haardt & Madau 2001). Accurate estimates of the UVB intensity at the Lyman limit are crucial for the understanding of the relative contribution of stars and quasars to the UVB and also to ensure realistic inputs for numerical simulations of structure formation (Davé et al. 1999; Hoeft et al. 2006).

The intensity of the UVB can be strongly affected by energetic sources such as bright QSOs leading to an enhancement of UV photons in their vicinity. The neutral fraction of H I consequently drops yielding a significant lack of absorption around QSOs within a few Mpc. Knowing the luminosity of the QSO at the Lyman limit, this “proximity effect” has been widely used in estimating the intensity of the UVB (Carswell et al. 1987; Bajtlik et al. 1988) at various redshifts on large samples (Bajtlik et al. 1988; Lu et al. 1991; Scott et al. 2000; Liske & Williger 2001; Scott et al. 2002, but compare Kirkman & Tytler 2008) and also towards individual lines of sight (Lu et al. 1996; Dall’Aglia et al. 2008b).

Principal obstacles in the analysis of the proximity effect are due to several poorly understood effects that in the end might lead to biased estimates of the UVB. The most debated influence is the gravitational clustering of matter around QSOs. Clustering enhances the number of absorption lines on scales of a few proper Mpc, which if not taken into account might lead to an overestimation of the UVB by a factor of a few. Alternatively, star-forming galaxies in the QSO-environment may contribute to an enhancement, or overionisation, of the local ionised hydrogen fraction, i.e., we would underestimate the UVB from the measured QSO luminosity. In the original formalism of the proximity effect theory introduced by Bajtlik et al. (1988) (hereafter BDO), the possibility of density enhancements near the QSO redshift was neglected. However, Loeb & Eisenstein (1995) and Faucher-Giguère et al. (2008b) found that the UVB may be overestimated by a factor of about 3 because of the high environmental density. Comparing observations and simulations, D’Odorico et al. (2008) found large-scale overdensities over about 4 proper Mpc. The disagreement between the UVB obtained via the proximity effect and from flux transmission statistics has been used to estimate the average density profile around the QSO (Rollinde et al. 2005; Guimarães et al. 2007; Kim & Croft 2008). Dall’Aglia et al. (2008a) showed that a major reason for overestimating the proximity effect is not a general cosmic overdensity around QSOs, but the methodological approach to estimate the UVB intensity by combining the proximity effect signal over several sight lines. Providing a definition of the proximity effect strength for a given LOS, they propose instead to investigate the strength distribution for the QSO sample, with which consistency with the theoretical estimates of the UVB is obtained.

Numerical simulations of structure formation have been a crucial tool in understanding the nature and evolution of the Ly α forest. One of the major challenges in the current development of 3D simulations is the implementation of radiative transfer into the formalism of hierarchical structure formation (Gnedin & Abel 2001; Maselli et al. 2003; Razoumov & Cardall 2005; Rijkhorst et al. 2006; Mellema et al. 2006; Pawlik & Schaye 2008). In particular, the coupling of radiative transfer with the density evolution requires a large amount of computational resources and remains up to now a largely unexplored field. Therefore, most available codes apply radiative transfer as a post-processing step.

Only a couple of radiative transfer studies exist for the Ly α forest (Nakamoto et al. 2001; Maselli & Ferrara 2005). They discuss the influence of radiative transfer on the widely used semi-analytical model of the forest by Hui et al. (1997). Maselli & Ferrara (2005) find that because of the self-shielding of the UVB flux in overdense regions the hydrogen photoionisation rate fluctuates by up to 20 per cent and the helium rate by up to 60 per cent. Since radiative transfer is not negligible for a background field, this should also be true for point sources. Until now numerical studies have dealt with ionisation bubbles in the pre-reionisation era or right at the end of reionisation (Iliev et al. 2007; Kohler et al. 2007; Maselli et al. 2007; Trac & Cen 2007; Zahn et al. 2007). These studies aim to determine the sizes of H II regions as observed in transmission gaps of Gunn-Peterson troughs (Gunn & Peterson 1965) in high-redshift sources, or in 21cm emission or absorption signals. By comparing radiative transfer simulations with synthetic spectra, Maselli et al. (2007) showed that the observationally deduced size of the H II regions from the Gunn-Peterson trough is underestimated by up to 30 per cent.

In the study of the proximity effect at redshifts lower than that of reionisation, the influence of radiative transfer has been neglected up to now, as the universe is optically thin to ionising radiation. We expect that absorption of QSO photons by intervening dense regions might reduce the QSO flux. These dense regions can shield themselves from the QSO radiation field (Maselli & Ferrara 2005) and would not experience the same increase in the ionisation fraction as low density regions. This leads to a dependence of the proximity effect on the QSO environment. Furthermore, the amount of hard photons in the QSO spectral energy distribution (SED) affects the proximity effect profile as suggested by Dall’Aglio et al. (2008b). As in the study of high redshift H II regions where ionisation fronts are broadened by the ionising flux’s shape of the spectral energy distribution (Shapiro et al. 2004; Qiu et al. 2007), we expect the size of the proximity effect zone to be a function of the spectral hardness. Harder UV photons have smaller ionisation cross-sections and cannot ionise hydrogen as effectively as softer UV photons.

In this study, we employ a three-dimensional radiative transfer simulation to study H II regions expanding in a pre-ionised intergalactic medium (IGM) at redshifts $z = 3, 4,$ and 4.9 . We intend to quantify the influence of the above-mentioned effects. To this end, we consider two realistic QSO luminosities. Furthermore, we study the influence of environment by placing the QSO either inside a massive halo, in a random filament, or in an underdense region (void).

The chapter is structured as follows. In Sect. 5.2, we discuss the dark-matter simulation used in this study, and we show that a realistic model of the Ly α forest is found from a semi-analytical model of the IGM. In Sect. 5.3, we describe the method to solve the radiative transfer equation, and discuss in Sects. 5.3.2 and 5.3.3 how the different UV sources are implemented. In Sect. 5.4, we review the standard approach used to characterise the proximity effect (Bajtlik et al. 1988). In Sect. 5.5 we describe the different radiative transfer effects on the overionisation profile. Then in Sect. 5.6, we introduce the proximity effect strength parameter and develop additional models to distinguish the various biases in the signal. In Sect. 5.7, we present the results for the mean line-of-sight proximity effect as determined from synthetic spectra. In Sect. 5.8, we discuss the proximity effect strength distribution of our spectra. We finally summarise our findings in Sect. 5.9.

5.2. Simulations

5.2.1. Initial realisation (IGM model)

We employ a DM simulation in a periodic box ² of 50 Mpc h^{-1} with 512^3 DM particles (von Benda-Beckmann et al. 2008). Using the PM-Tree code GADGET2 (Springel 2005) our simulations yield a force resolution of 2 kpc h^{-1} and a mass resolution of $m_p = 7.75 \times 10^7 M_\odot h^{-1}$. We fix the cosmological parameters to agree with the 3rd year WMAP measurements (Spergel et al. 2007): The total mass density parameter is $\Omega_{m,0} = 0.3$, while the baryon mass density is $\Omega_{b,0} = 0.04$ and the vacuum energy is $\Omega_\Lambda = 0.7$. The dimensionless Hubble constant is $h = 0.7$, and the power spectrum is normalised by the square root of the linear mass variance at 8 Mpc h^{-1} , $\sigma_8 = 0.9$.

Simulating the Ly α forest is a challenging task, since a large box is required to capture the largest modes that still influence the mean flux and its distribution (Tytler et al. 2009; Lidz et al. 2010). Furthermore, high resolution is required to capture the properties of single absorbers correctly. Our simulations yield a reasonable representation of the average statistics of the Lyman alpha forest (Theuns et al. 1998b; Zhang et al. 1998; McDonald et al. 2001; Viel et al. 2002) for low redshifts. However, it has been shown by Bolton & Becker (2009) and Lidz et al. (2010) that to resolve the Ly α forest properly at high redshifts, the mass resolution should be higher by two orders of magnitudes for $z = 5$ and one order of magnitude for $z = 4$. Given the large box sizes needed, these resolution constraints cannot be fully met with today's simulations. According to Bolton & Becker (2009), the error in our estimate of the mean flux at $z = 4$ is around 10%. For $z = 3$, the simulations will be marginally converged with mass resolution.

We record the state of the simulation at three different redshifts $z = 4.9, 4$, and 3. Using cloud in cell assignment we convert the DM particle distribution into a density and velocity field on a 400^3 regularly spaced grid.

We then select three different environments from the highest redshift snapshot: The most massive halo with a mass of $M_{\text{halo}} = (2.6 \times 10^{12}, 4.0 \times 10^{12}, 7.9 \times 10^{12}, 1.0 \times 10^{15}) M_\odot h^{-1}$ at redshifts $z = (4.9, 4, 3, 0)$, a random filament, and a random void. The last two environments are selected by visual inspection of the particle distribution and their position is tracked down to $z = 3$. These locations are assumed to host a QSO.

To characterise these environments more precisely, we estimate the volume-averaged overdensities δ_5 in $r = 5$ Mpc h^{-1} spheres and find that $\delta_5 = (1.7, 1.8, 2.1)$ at redshifts $z = (4.9, 4.0, 3.0)$ for the halo, $\delta_5 = (1.0, 0.8, 1.1)$ for the filament, and $\delta_5 = (0.8, 0.7, 0.7)$ for the void. Thus the halo resembles a locally overdense region, the void an locally underdense region, and the filament an average environment. Figure 5.1 illustrates a snapshot of the DM density field at redshift $z = 4$.

To describe the baryonic component of the IGM, we assume a universe containing hydrogen only, whose density and velocity fields are proportional to those of the dark-matter (Petitjean et al. 1995). To account for pressure effects on baryons, Hui et al. (1997) proposed convolving the density field with a window function cutting off power below the Jeans length. In Viel et al. (2002), differences in the gas and DM densities between hydrodynamical and DM-only simulations are studied. For $\delta \approx 3$, baryons follow the DM distribution quite well, i.e. over most of the densities relevant to the

²Distances are given as comoving distances unless otherwise stated.

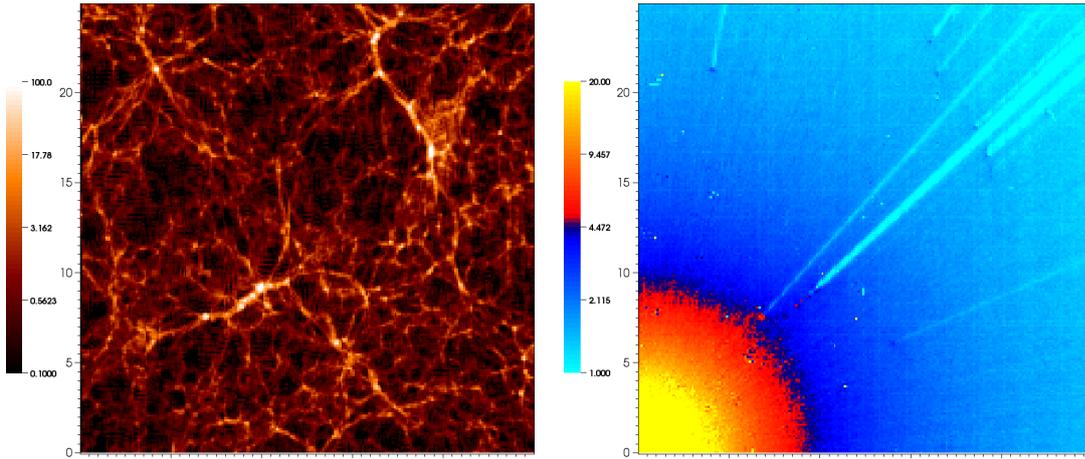


Figure 5.1.: Cuts through the 25 Mpc h^{-1} simulation box at redshift $z = 4$ for a QSO sitting in the void at the lower left corner. In the left panel, the overdensity field $1 + \delta$ is shown. The right panel provides the inverse of the overionisation fraction $n_{\text{HI,Ly-}\alpha}/n_{\text{HI,prox}}$ due to a $L_{\nu_{\text{Ly}\alpha}} = 10^{31}$ erg Hz $^{-1}$ s $^{-1}$ QSO. Here $n_{\text{HI,Ly-}\alpha}$ and $n_{\text{HI,prox}}$ are the neutral hydrogen fractions without and with the influence of the QSO respectively. Clearly visible are the extended overionisation zone and shadowing effects. The scattered white dots arise from numerical noise. The axes are in comoving h^{-1} Mpc.

Lyman- α forest. However, at higher DM densities, the corresponding gas density is lower because of the smoothing induced by gas pressure. We implicitly smooth our density field with the cloud-in-cell density assignment scheme and cell sizes of 125 h^{-1} kpc. This is comparable to the Jeans length of $\approx 150 h^{-1}$ kpc at $z = 3$ and mean density, which scales as $\delta^{-1/2}(1+z)^{-1/2}$.

The hydrogen density at position x is then given by

$$n_{\text{H}}(x) = \frac{3 H_0^2 \Omega_{b,0}}{8 \pi G m_{\text{p}}} (1+z)^3 (1+\delta(x)), \quad (5.1)$$

where G is the gravitational constant, H_0 the Hubble constant, and m_{p} the proton mass. The DM overdensity is $1 + \delta(x) = \rho(x)/\bar{\rho}$, where the DM density is $\rho(x)$. Finally the hydrogen velocity field $v(x)$ is assumed to be equal to the DM one.

5.2.2. Model and calibration of the intergalactic medium

The thermal evolution of the IGM is determined mainly by the equilibrium between photoionisation heating and adiabatic cooling, resulting in a tight relation between the density and temperature of the cosmic gas. This relation is known as the effective equation of state (Hui & Gnedin 1997) and is typically expressed by $T = T_0(1+\delta)^{\gamma-1}$. We apply the effective equation of state as an estimate of the baryonic density for overdensities $0.1 < \delta < 10$ relevant to the Lyman alpha forest. Higher densities correspond to collapsed regions that we approximate by assuming a cut-off temperature $T_{\text{cut-off}} = T(\delta = 10)$. Following the formalism of Hui et al. (1997), we can compute an HI Ly α absorption

5. Cosmological radiative transfer for the LOS proximity effect

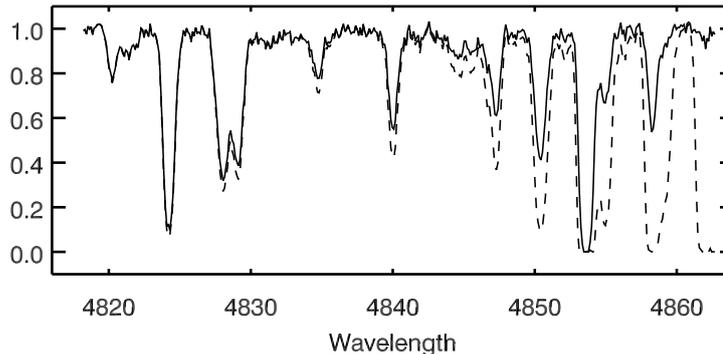


Figure 5.2.: Mock spectra synthesised from the $z = 3$ snapshot with a QSO in a filament. The QSO sits on the right hand side of each spectrum, and the wavelength is in Ångström. The dashed line gives the Ly α forest spectrum without the influence of the QSO, the solid line the spectrum from the radiative transfer simulation of the proximity effect for a QSO with a Lyman limit luminosity of $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$. The spectra have a resolution of $\Delta v = 6.7 \text{ km s}^{-1}$ and a signal-to-noise ratio of 70.

spectrum from the density and velocity fields, once T_0 , γ , and the UV background photoionisation rate (Γ_{UVB}) are fixed. Thus, by relating simulated to observed H I absorption spectra properties, we can constrain these free parameters. This calibration is crucial in the following to provide realistic representations of the IGM for the radiative transfer calculations.

To calibrate our spectra, we employ three observational constraints sorted by increasing importance: (i) The observed equation of state (Ricotti et al. 2000; Schaye et al. 2000b), (ii) the evolution of the UV background photoionisation rate (Haardt & Madau 2001; Bianchi et al. 2001), and (iii) the observed evolution of the effective optical depth in the Ly α forest (Schaye et al. 2003; Kim et al. 2007).

Our catalogues of synthetic spectra consist of 500 lines of sight at each redshift randomly drawn from the cosmological box. The spectra are binned to the typical resolution of UVES spectra of 6.7 km s^{-1} and convolved with the UVES instrument profile. Random noise with a signal-to-noise ratio (S/N) of 70 is added, the mean S/N in the QSO sample used by Dall’Aglia et al. (2008a). An example of these mock spectra is presented in Fig. 5.2.

The model parameters of the effective equation of state T_0 and γ were chosen according to observations of Ricotti et al. (2000) and Schaye et al. (2000b). To obtain values for $z > 4$, the observed T_0 and γ parameters were extrapolated to higher redshifts (see Fig. 6.3). We constrained our model to yield an observed average effective optical depth $\tau_{\text{eff}}(z) = -\ln \langle F(z) \rangle$, where F is the transmitted flux, and the averaging is performed over the whole line of sight. For this, we used observations from Kim et al. (2007). With these constraints, we determined the UVB photoionisation rate Γ_{UVB} for our models.

Our model parameters are presented in Table 6.1, and are plotted in Fig. 6.2 in comparison with different literature results. Both the inferred evolution of the UV background and the effective optical depth agree closely with the results of Haardt & Madau (2001) and Bolton et al. (2005), and high resolution observations by Schaye et al. (2003) and Kim et al. (2007), respectively.

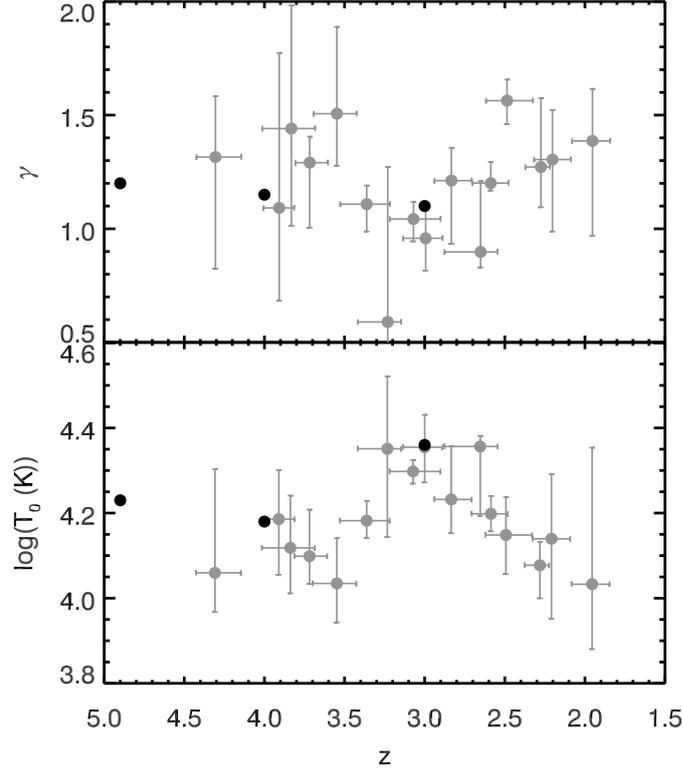


Figure 5.3.: *Upper panel:* Comparison of our choices for γ (black points) with observationally derived results by Schaye et al. (2000b) (grey points). *Lower panel:* Comparison of our model T_0 (black points) with observations by Schaye et al. (2000b) (grey points).

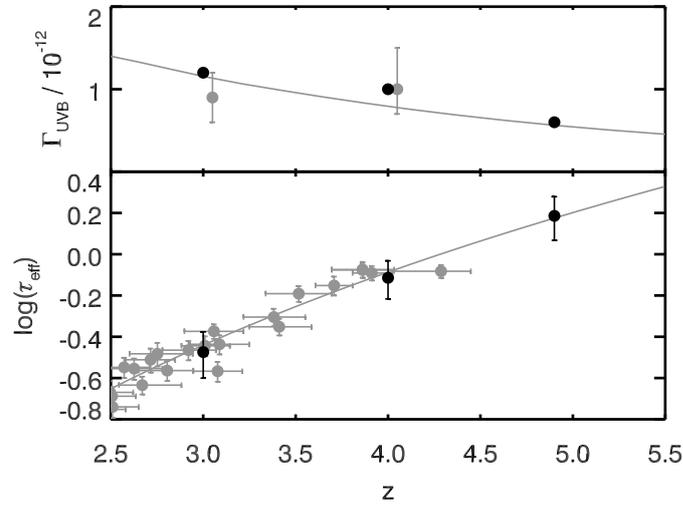


Figure 5.4.: *Upper panel:* The evolution of the UV background photoionisation rate in our three snapshots (black points) compared to Bolton et al. (2005) (grey points shifted by $z = 0.05$ for better visibility) and predictions by Haardt & Madau (2001) (grey line). *Lower panel:* The effective optical depth of our models (black points) in comparison to measurements by Schaye et al. (2003) (grey points). The continuous line shows the fit by Kim et al. (2007).

5. Cosmological radiative transfer for the LOS proximity effect

Table 5.1.: Model parameters of the semi-analytical model.

z	$\log T_0$	γ	Γ_{UVB}	$\log \tau_{\text{eff}}$
4.9	4.23	1.20	0.6×10^{-12}	0.186
4.0	4.18	1.15	1.0×10^{-12}	-0.114
3.0	4.36	1.10	1.2×10^{-12}	-0.474

5.3. Radiative transfer in cosmological simulations

5.3.1. Method

We now focus on the method implemented to solve the radiative transfer equation. In developing our Monte-Carlo radiative transfer code (A-CRASH), we closely follow the approach first introduced by Ciardi et al. (2001) and extended by Maselli et al. (2003). While their latest development includes multi-frequency photon packages (Maselli et al. 2009), we only implement the earlier formulation, which uses monochromatic packages³. The advantage of the Monte Carlo scheme is that it accounts for the emission of diffuse recombination photons, which is typically neglected in codes relying on the “on-the-spot” approximation (see Sect. 5.5.3 for further details).

The idea behind a Monte Carlo radiative transfer scheme is to bundle the radiation flux field into a discrete amount of radiation energy in the form of photon packages of energy content

$$\Delta E_i = \int_{t_{i-1}}^{t_i} L_{\text{UV}}(t) dt \quad (5.2)$$

emitted by a source of UV luminosity L_{UV} between two time steps t_{i-1} and t_i . At each time step t_i , we allow the corresponding content of energy ΔE_i to be evenly distributed among the number of photon packages emitted by the source. In our implementation of the Monte Carlo scheme, this number can be larger than one. In addition, each photon is characterised by a given frequency ν following a given SED. Thus, knowing the source SED and its UV luminosity we are able to infer the number of photons $N_{\gamma, i}$ per package. To guarantee a proper angular and spatial sampling of the radiative transfer equation, the photon packages are emitted in random directions.

Once all the photon packages are produced by the sources, they are propagated through the computational domain. Each time a package crosses a cell, a certain amount of photons is absorbed. The absorption probability in the l -th cell is

$$P(\tau_l) = 1 - e^{-\tau_l}, \quad (5.3)$$

where τ_l is the optical depth in the l -th cell

$$\tau_l = \sigma_{\text{HI}}(\nu) n_{\text{HI}, l} f_l \Delta x \quad (5.4)$$

and σ_{HI} is the hydrogen photoionisation cross-section, $n_{\text{HI}, l}$ the neutral hydrogen number density, and $f_l \Delta x$ the crossing path length of a ray through a cell of size Δx . We calculate

³Our A-CRASH code is OpenMP parallel and publicly available under the GPL license at <http://sourceforge.net/projects/acrash/>

the exact crossing length using the fast voxel traversal algorithm by Amanatides & Woo (1987).

The number of absorbed photons in the cell is $N_{A,l} = N_{\gamma,l} P(\tau_l)$ where $N_{\gamma,l}$ denotes the remaining number of photons in the package $N_{\gamma,l} = N_{\gamma,l-1} - N_{A,l-1}$ arriving at the l -th cell. In our simple case of pure hydrogen gas, we can write

$$n_{\text{H}} \frac{dx_{\text{HII}}}{dt} = \gamma_{\text{HI}}(T) n_{\text{HI}} n_{\text{e}} - \alpha_{\text{HII}}(T) n_{\text{HII}} n_{\text{e}} + \Gamma_{\text{HI}} n_{\text{HI}}, \quad (5.5)$$

where $x_{\text{HII}} = n_{\text{HII}}/n_{\text{H}}$ is the ionisation fraction, n_{H} is the total hydrogen density, $\gamma_{\text{HI}}(T)$ is the collisional ionisation rate, $\alpha_{\text{HII}}(T)$ the recombination rate, and Γ_{HI} is the photoionisation rate derived from the number of absorbed photons $N_{A,l}$. We adopt $\gamma_{\text{HI}}(T)$ and $\alpha_{\text{HII}}(T)$ as in Maselli et al. (2003). We solve this stiff differential equation using a 4th order Runge-Kutta scheme.

The extent of a time step is defined as the total simulation time t_s divided by the total number of photon packages N_p emitted by each source. The numerical resolution of the simulation is determined by calculating the mean number of packages crossing a cell

$$N_{\text{cr}} = \frac{N_s N_p}{N_c^2} \gg \frac{t_s}{t_{\text{min}}}, \quad (5.6)$$

where N_s is the number of sources, N_c the number of cells in one box dimension, and t_{min} the smallest characteristic timescale of all the processes involved. To efficiently parallelise the scheme, the original rule of producing one photon per source per time step is dropped; a source is allowed to produce more than one package per time step as long as $\Delta t \ll t_{\text{min}}$.

We confirmed that our implementation passes the simple tests described by Maselli et al. (2003) (see Partl (2007) for further details).

5.3.2. UV background field

The UV background photoionisation rates given in Table 6.1 need to be modelled in the framework of the radiative transfer scheme. We assume the spectral shape of the UVB to be a power law ν^{α_b} with $\alpha_b = -1.3$ (Hui et al. 1997), which is a bit lower than measurements by Fechner et al. (2006) yielding $\alpha_b = -1.99 \pm 0.34$. From the spectral energy distribution of the UV background field and its intensity, photon packages need to be constructed. To obtain the number of photons in a background package, the total energy content of the background field in the box is mapped to single photon packages; we follow the method of Maselli & Ferrara (2005), but for the hydrogen only case.

To derive the energy content of a single background field photon package N_γ , the total amount of energy carried by the background field in the whole simulation box needs to be considered. The change in the mean ionised hydrogen density n_{HII} by the UVB is

$$\Delta n_{\text{HII}} = \Gamma_{\text{UVB}} n_{\text{HI}} \Delta t, \quad (5.7)$$

where Γ_{UVB} is the photoionisation rate of the background field. This is compared to the total amount of absorbed photons in the box $N_\gamma (1 - \exp(-\Delta\tau))$, where $\Delta\tau = \sigma_{\text{HI}} \Delta n_{\text{HI}} d_{\text{B}}$ is the total optical depth in the box of length d_{B} . This corresponds to a

5. Cosmological radiative transfer for the LOS proximity effect

mean change in the ionised hydrogen number density of

$$\Delta n_{\text{HII}} = \frac{N_\gamma (1 - e^{-\Delta\tau})}{d_{\text{B}}^3}. \quad (5.8)$$

By equating Eqs. 5.7 and 5.8 and assuming an optically thin medium $\Delta\tau \ll 1$ with $\Delta\tau \approx (1 - e^{-\Delta\tau})$, we derive the photon number content of a background photon package

$$N_\gamma = \frac{\Gamma_{\text{UVB}} \Delta t (\Delta x \cdot N_c)^2}{\sigma_{\text{H1}}}. \quad (5.9)$$

We note that this is only true, if the photon package is propagated over the distance of exactly one box length.

Background photons are emitted isotropically from random cells in the box. Dense regions are allowed to shield themselves from the background flux. Thus we emit background photons only from cells below a certain density threshold δ_{UV} . In Maselli & Ferrara (2005) a threshold of $\delta_{\text{UV}} = 60$ was used, corresponding to the density at the virial radius of collapsed haloes. We chose a lower threshold $\delta_{\text{UV}} = 1$ to ensure that mildly overdense regions have the possibility of shielding themselves from the UV background.

5.3.3. QSO radiation

As with the UV background, discrete point sources are characterised by a SED and luminosity. We include only one source of radiation other than the UVB. The point source representing the QSO was chosen to follow the composite QSO spectra obtained by Trammell et al. (2007). The mean SED was constructed from over 3000 spectra available in both the Galaxy Evolution Explorer (GALEX) Data Release 1 and the Sloan Digital Sky Survey (SDSS) Data Release 3. It covers a broad wavelength range of about $9000 > \lambda > 300\text{\AA}$.

If a power law ν^{α_q} is assumed for the UV part of the QSO spectrum, the data of Trammell et al. (2007) leads to $\alpha_q \approx -2.5$ for $\lambda < 912\text{\AA}$. Since the scatter in the data of Trammell et al. (2007) is large in the wavelength interval we are interested in, and other authors find harder spectra (Telfer et al. 2002; Scott et al. 2004), we consider in addition a power law SED with a shallower slope of $\alpha_q = -1.5$. The upper energy limit at $\lambda = 300\text{\AA}$ results in a underestimation of the hydrogen photoionisation rate of 0.6% for $\alpha = -1.5$ and only 0.2% for $\alpha = -2.5$ with respect to spectra extending to higher energies.

The two QSO Lyman limit luminosities $L_{\nu_{\text{LL}}}$ studied are chosen to bracket the luminosity range of observed QSOs (Scott et al. 2000; Rollinde et al. 2005; Guimarães et al. 2007; Dall’Aglia et al. 2008a). We therefore chose the QSOs to have Lyman limit luminosities of $10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ and $10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$. To obtain the total UV luminosity for use with Eq. 5.2, the high energy part of the SED below 912\AA is scaled to the given Lyman limit luminosity and integrated from $912 > \lambda > 300\text{\AA}$. For the Trammell et al. (2007) spectral template and $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$, we obtain a total UV luminosity of $L_{\text{UV}} = 1.15 \times 10^{46} \text{ erg s}^{-1}$, and for $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ a luminosity

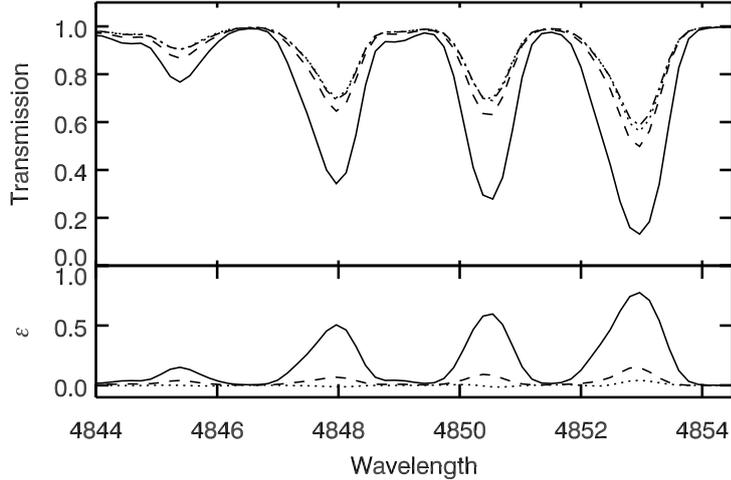


Figure 5.5.: Ly α absorption along a sight line near a $z = 3$ $L_{\nu_{\text{LL}}} = 10^{32}$ erg Hz $^{-1}$ s $^{-1}$ QSO simulated with photon numbers $N_{p,QSO} = 1 \times 10^8$ (solid line), 1×10^9 (dashed line), 2×10^9 (dotted line), and 4×10^9 (dash dotted line). The wavelength is in Ångström and no noise is added to the spectra for better visibility. The lower panel shows the relative deviation from the highest resolution run.

of $L_{UV} = 1.15 \times 10^{47}$ erg s $^{-1}$.

5.3.4. Simulation setup and convergence

Owing to the high precision required to resolve the slight changes in the ionisation fractions governing the proximity effect, a large amount of photon packages had to be tracked. To achieve this, sub-boxes of 25 Mpc h^{-1} with 200^3 cells were selected from the whole box. To ensure that the resulting proximity effect region could be traced to as large a radius as possible, the QSO source was located at the box origin at the cost of analysing only 1/8 of the full sphere around the QSO.

A convergence study was carried out for one of our models with resulting mock spectra shown in Fig. 5.5. We tested QSO samplings of $N_{p,QSO} = 10^8$, 10^9 , 2×10^9 , and 4×10^9 for a $L_{\nu_{\text{LL}}} = 10^{32}$ erg Hz $^{-1}$ s $^{-1}$ QSO residing in a filament at $z = 3$. For this test, the $\alpha = -2.5$ spectrum was used. Simulating the bright QSO at $z = 3$ provides the most challenging convergence test. The IGM is already highly ionised, and the influence of the quasar result in very small neutral hydrogen fractions and we thus need high numerical resolution. The lower luminosity QSO will not alter the IGM as strongly and therefore resolving the ionisation fractions numerically is easier. The same is true at higher redshift, where the IGM is not yet as strongly ionised as at the low redshifts, and the resulting neutral fraction are larger because of the larger optical depth. The $\alpha = -2.5$ spectrum is used for this test, since it is more challenging to properly sample the high energy tail than for the harder $\alpha = -1.5$ model.

It is obvious that 10^8 photons are insufficient to model the QSO, even if the mean number of photons crossing each cell is around 10. This number has been considered sufficient to resolve Strömberg spheres in a homogeneous medium (Maselli et al. 2003). The solutions for $N_{p,QSO} = 2 \times 10^9$ and 4×10^9 are similar, with relative differences from

5. Cosmological radiative transfer for the LOS proximity effect

the highest resolution run of at most 5 per cent. To achieve good angular sampling of the QSO environment and sufficiently large cell-crossing numbers, we consider the solution for $N_{p,\text{QSO}} = 2 \times 10^9$ as converged. To correctly sample the whole sphere, 1.6×10^{10} packages would be required, which is at the moment beyond the capabilities of our code.

The Monte Carlo radiative transfer method is able to follow photo ionisation heating, although since this effect is already included in the equation of state, we keep the temperatures derived from the equation of state fixed throughout the simulation. This approach neglects any temperature fluctuations caused by He II photo-heating during He II reionisation around $z \approx 3$ (McQuinn et al. 2009; Meiksin et al. 2010). The simulations are run up to $t_s = 2.5 \times 10^8$ yr and the ionisation fractions were averaged over different outputs at different time steps to reduce the Monte Carlo simulation noise.

5.4. Line-of-sight proximity effect

In addition to the radiative transfer method for simulating the proximity effect, we also exploit a standard approach to account for the QSO radiation field. This method builds on the work of Bajtlik et al. (1988, also BDO) and Liske & Williger (2001) and we briefly summarise it here. On the basis of the assumption that the IGM is in photoionisation equilibrium with the global UV radiation field, the number density of UV photons produced by bright QSOs is higher in their vicinity than far away and dominate over the UVB. This leads to a reduction in neutral hydrogen, and with it to an opacity deficit of the IGM within a few Mpc from the QSO.

The influence of the QSO onto the surrounding medium leads to an effective optical depth in the Ly α forest

$$\tau_{\text{eff,QSO}}(z) = \tau_{\text{eff,Ly}\alpha}(z) (1 + \omega(z))^{1-\beta} \quad (5.10)$$

(Liske & Williger 2001) where $\tau_{\text{eff,Ly}\alpha}(z)$ represents the evolution of the effective optical depth in the Ly α forest with redshift, and $\tau_{\text{eff,QSO}}(z)$ is the effective optical depth including the alterations by the QSO radiation. Further β is the slope in the column density distribution and $\omega(z)$ is the ratio between the QSO and the UVB photoionisation rates.

Then following the assumption of pure geometrical dilution of the QSO radiation as proposed by Bajtlik et al. (1988), we obtain

$$\omega_{\text{BDO}}(z) = \frac{L_{\nu_{\text{LL}}}}{16\pi^2 J_{\nu_{\text{LL}}} (1+z) d_L(z_q, z)^2} \quad (5.11)$$

with z as the redshift of absorbers along the LOS such that $z < z_q$, and $d_L(z_q, z)$ is the luminosity distance from the absorber to the QSO. $L_{\nu_{\text{LL}}}$ is the Lyman limit luminosity of the QSO and $J_{\nu_{\text{LL}}}$ the UVB flux at the Lyman limit. In observations, the emission redshift of the QSO z_q is subject to uncertainties which propagates through to the ω scale and increase the uncertainties in the determined UVB photoionisation rate (Scott et al. 2000). In this work we consider the emission redshift to be perfectly known.

We note that this formula is derived assuming identical spectral indices of the QSO and the UVB. In our simulation we consider two different SEDs for the UVB and the

QSO. Therefore in using Eq. 5.11 to estimate the proximity effect, we would introduce a bias. In $\omega(z)$ the two different spectral shapes can be accounted for by the ratio of the photo-ionisation rates of the QSO as a function of radius $\Gamma_{\text{QSO}}(r)$ and the photo-ionisation rate of the UVB, Γ_{UVB}

$$\omega(z) = \frac{\Gamma_{\text{QSO}}(r)}{\Gamma_{\text{UVB}}} = \omega_{\text{BDO}}(z) \frac{3 - \alpha_{\text{b}}}{3 - \alpha_{\text{q}}}, \quad (5.12)$$

where α_{b} and α_{q} describe the UVB and the QSO SED, respectively, by power laws $f_{\nu} \propto \nu^{\alpha}$. Further it is assumed that $\sigma_{\text{HI}} \propto \nu^{-3}$. Finally, the ω scale is uniquely defined once the QSO Lyman limit flux, redshift, and the UVB flux are known. In the following analysis, we always use Eq. 5.12 unless otherwise stated.

This correction for differing SEDs has been used in measurements of the UVB photoionisation rate (Dall’Aglia et al. 2009). However, depending on the spectral shape of the QSO, omitting this correction can result in over- or underestimation of the background flux. As stated in Sect. 5.3.3, we use a QSO spectrum with a rather steep slope of $\alpha_{\text{q}} = -2.5$. Using the original BDO formulation alone would result in an overestimation of the UV background flux by almost 30% assuming $\alpha_{\text{b}} = -1.3$. We later check whether this correction is able to model the SED effect by comparing our simulations with one that uses a shallower QSO spectral slope of $\alpha_{\text{q}} = -1.5$.

As a side note, this spectral effect will alter the QSO’s sphere of influence. We assume as size of the proximity effect the radius $r_{\omega=1}$ where $\Gamma_{\text{QSO}} = \Gamma_{\text{UVB}}$. Then $r_{\omega=1} \propto r_{\text{BDO},\omega=1} \sqrt{(3 - \alpha_{\text{b}})/(3 - \alpha_{\text{q}})}$, and the radius is in the case of our steep QSO spectra 13% smaller for equal $L_{\nu_{\text{LL}}}$.

5.5. Radiative transfer effects

5.5.1. The overionisation profile

The overionisation profile of the proximity effect around the QSO decays radially because of to geometric dilution. In addition, differences in the spectral energy distribution between the UVB and the QSO influences the size of the proximity effect zone. Both factors are included in the ω scale. To determine by how much the QSO decreases its surrounding neutral hydrogen density in our simulations, we determine the overionisation profile $\Xi(r) = n_{\text{HI,Ly}-\alpha}/n_{\text{HI,Prox}}$, where $n_{\text{HI,Ly}-\alpha}$ gives the neutral hydrogen fraction unaffected by the QSO radiation. In addition $n_{\text{HI,Prox}}$ denotes the neutral hydrogen density when we include the additional QSO radiation. Assuming ionisation equilibrium, the overionisation profile is directly proportional to the ω scale.

In Fig. 5.6, we show the median overionisation profile determined along 200 lines of sight. These have been extracted from the simulation of a $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO residing in the most massive halo at $z = 3$ and $z = 4.9$. For both redshifts, the simulation data closely follows the analytical model described in Sect. 5.4 at radii $r \gtrsim 5 \text{ Mpc } h^{-1}$ at $z = 3$ and $r \gtrsim 2 \text{ Mpc } h^{-1}$ at $z = 4.9$. However, for the faint QSO, the simulated overionisation profile begins to deviate from the analytical solution at smaller radii, where the near host environment leaves an imprint on the profile. In contrast, for the bright QSO shown in the upper left panel of Fig. 5.8, the profile more closely follows the

5. Cosmological radiative transfer for the LOS proximity effect

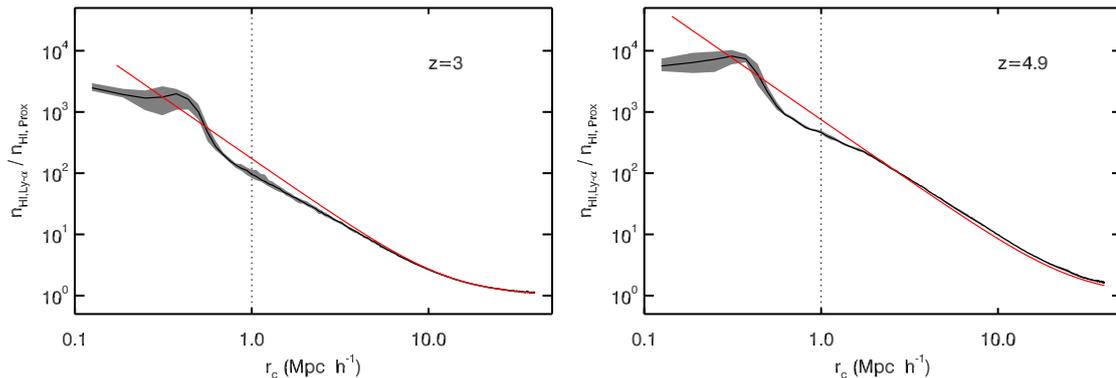


Figure 5.6.: Median overionisation profile as a function of the distance from the QSO (black line, grey area gives upper and lower quartiles) hosted by a halo at $z = 3.0$ (left) and $z = 4.9$ (right side) for a QSO luminosity $L_{\nu_{\text{LL}}} = 1 \times 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$. The analytical overionisation profile including the SED effects is given by the solid line without the grey area. The region affected by oversampling ($r < 1 \text{ Mpc}$) is marked with a dotted line.

analytical model. We need to emphasise that the spatial resolution of the simulations are limited to $r = 0.125 \text{ Mpc } h^{-1}$. Therefore the immediate vicinity of the QSO is poorly resolved. Furthermore, all cells up to a radius of 1 Mpc away from the QSO are subject to oversampling in all our cases. Only beyond a radius of 1 Mpc are cells at most sampled once every time step. Therefore, the solution in this part of the overionisation profile, which is marked in our plots, is unreliable.

5.5.2. Shadowing by Lyman limit systems

We now focus on the overionisation field and discuss how optically thick regions in the intergalactic medium affect the proximity effect. A slice of the overionisation field is shown in the right panel of Fig. 5.1 for the $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO residing in a void at $z = 4$. Again the smooth overionisation profile is seen as expected. However in this smooth transition zone, large shadow cones originating in dense regions are present. The optical depth of these clouds is large enough to absorb most of the ionising photons produced by the QSO. Hence, the ionisation state of the medium behind this absorber remains at the unaltered UVB level.

We determined the hydrogen column density of the shadowing regions in Fig. 5.1. They lie in the range of $1 \times 10^{17} < N_{\text{H I}} < 5 \times 10^{18} \text{ cm}^{-2}$. The regions causing the shadows therefore represent Lyman limit systems. Whether these systems are able to absorb a large amount of QSO photons depends on the one hand on their density and on the other on the QSO flux reaching them. Absorbers further away from the source will receive less flux from the QSO because of geometric dilution. Therefore, they are more likely to remain optically thick. For systems closer to the QSO, the flux field is more intense and a higher density is needed for them to remain optically thick. Thus the number of shadows increases with increasing distance to the source. Moreover, a highly luminous QSO will produce a larger amount of ionising photons and a higher number of absorbers are rendered optically thin. Hence fewer shadows are present in the overionisation field of a highly luminous QSO.

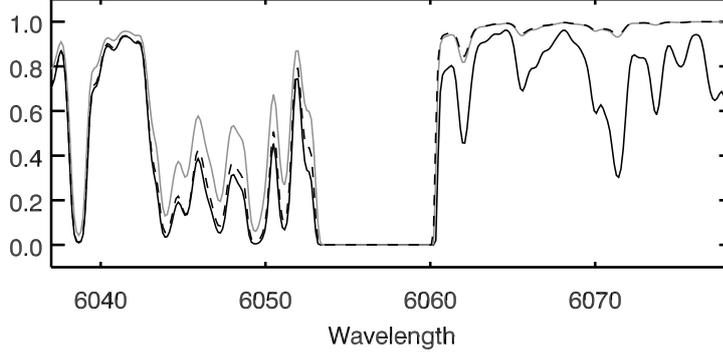


Figure 5.7.: Mock spectra synthesised from the $z = 4$ simulation with the QSO in a void. The solid line gives the Ly α forest spectrum without the influence of the QSO, the dashed line the spectrum from the radiative transfer simulation of the proximity effect for a QSO Lyman limit luminosity of $L_{\nu_{LL}} = 10^{31}$ erg Hz $^{-1}$ s $^{-1}$. The grey line provides the results of a semi-analytical model of the proximity effect that differ significantly from the radiative transfer results behind the strong absorber at the middle of the spectrum. No noise is added to the spectra for better visibility.

By extracting a spectrum through such an absorber (see Fig. 5.7), the influence of the Lyman limit system on the proximity effect signature can be seen. Behind the strong absorber on the left hand side, the spectrum follows the unaffected Ly α forest spectrum. Measurements of the proximity effect behind this system do not show any QSO influence. This shadowing is not described by a semi-analytical model where the optical depth in the forest is altered in proportion to $1 + \omega(z)$. We discuss and use this semi-analytical model in Sect. 5.6. By looking at the semi-analytic model spectrum shown in Fig. 5.7, it is clear that in this model the proximity effect extends beyond the Lyman limit system.

5.5.3. Diffuse recombination radiation

We now discuss the role played by photons produced by recombining electrons with the help of the overionisation profile, since our radiative transfer simulations provide the opportunity to study this diffuse component. With the Monte Carlo scheme we can directly model the radiation field produced by recombination, which in most other schemes is not treated self-consistently. Diffuse radiation produced by recombining electrons plays an important role in radiative transfer problems if the medium is optically thin. Whenever photons originating in recombination events are able to travel over large distances (i.e. larger than one cell size), energy is redistributed on the scale of the photon's mean free path and the ionisation state is altered over these distances. The importance of diffuse radiation in the simplest of radiative transfer problems, the Strömgren sphere, has been shown in Ritzerveld (2005) where the outer 30% of the Strömgren radius are found to be dominated by recombination photons. In addition Aubert & Teyssier (2008) used their radiative transfer code to discuss how much the on-the-spot approximation, a popular assumption that treats recombination photons as a local process, alters the results of widely used test cases.

Recombination photons to the hydrogen ground level possess enough energy to ionise

5. Cosmological radiative transfer for the LOS proximity effect

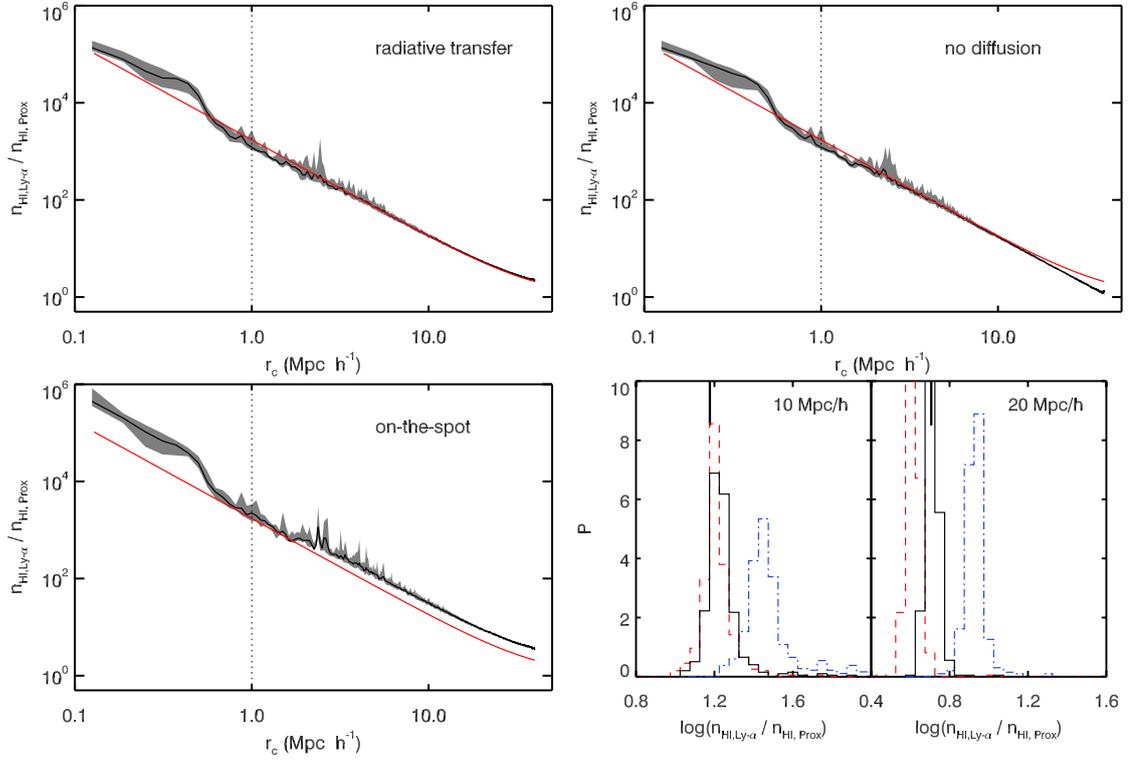


Figure 5.8.: Median overionisation profile as a function of the distance from the QSO hosted by a halo as in Fig. 5.6. We show the influence of the diffuse field on the H II region for $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ at $z = 3.0$. The upper left plot shows the full radiative transfer simulation, the right plot the one without diffusive radiation, and the lower left plot results using the on-the-spot approximation. The analytical overionisation profile including the SED effects is given by the solid line. The region affected by oversampling ($r < 1 \text{ Mpc}$) is marked with a dotted line. The lower right histograms gives the distribution of the overionisation fraction at radii $r = 10 \text{ Mpc h}^{-1}$ (left) and $r = 20 \text{ Mpc h}^{-1}$ (right). The solid lines give radiation transfer results, the dashed lines shows results omitting diffusion, and the dot-dashed distribution illustrates results obtained with the on-the-spot approximation. The solid tick indicates the analytical solution.

other hydrogen atoms. If this photon is absorbed by a nearby neutral atom, the on-the-spot approximation can be applied where it is assumed that all recombination photons are absorbed by neighbouring atoms. This is true if the medium is dense and the optical depth high enough for local equilibrium between recombination and absorption to be reached. If however the optical depth is small, the mean free path is longer than the simulation cell size and the recombination photons are absorbed far away from their origin. In this case, the on-the-spot approximation breaks down.

We expect recombination radiation to play an important role in the proximity effect, since the optical depths in the Ly α forest are low. To identify the region where recombination processes influence the solution, we run one simulation without emitting recombination photons. However, we still use case A recombination rates and call this the ‘no diffusion case’. Furthermore we implement the on-the-spot approximation using case B recombination rates. As before, we do not follow any recombination photons. We expect to see differences at lower redshift, where the mean free paths for ionising photons are large. Therefore, we consider the snapshot at redshift $z = 3$ with the $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1}\text{s}^{-1}$ QSO residing in a halo for this experiment.

The resulting median overionisation profiles are presented in Fig. 5.8. We also show the overionisation distribution at two radii to more accurately compare the different runs. At radii $r > 1 \text{ Mpc } h^{-1}$ we obtain profiles that closely reproduce the theoretical estimates. We see some noisy fluctuations in the profiles. Owing to the larger ionisation fractions for the bright QSO, the ionisation fractions themselves are more susceptible to Monte Carlo noise. Since the neutral hydrogen fractions in the direct vicinity of the QSO are 2 dex lower for the bright QSO than for the faint one, higher precision is needed to more precisely evaluate the extremely low neutral fractions. However, increasing the number of photon packages yet further is at the moment beyond the capabilities of our code.

Comparing the no diffusion run with the full radiative transfer solution shows that recombination greatly contributes to the outer parts of the proximity effect region. The no diffusion solution does not gradually go to unity, but continues to decay rapidly until the QSO’s influence vanishes. At $r = 20 \text{ Mpc } h^{-1}$, the difference between the result with and without the diffuse recombination field is 23%. This means that recombination radiation contributes to the photoionisation rate by 23% at this radius, if one assumes that the ionisation fraction is directly proportional to the photoionisation rate.

The on-the-spot solution on the other hand over predicts the extent of the proximity effect zone by about a factor of 2. This means that keeping the energy produced by recombining electrons locally confined over predicts the amount of ionised hydrogen. As a result, the QSO’s sphere of influence is overestimated. Furthermore, the on-the-spot solution exhibits a larger dispersion between the lines of sight than the radiative transfer simulation (see Fig. 5.8 lower right panel). We therefore conclude that the on-the-spot approximation is insufficient when modelling the proximity effect.

5.6. Radiative transfer on the line-of-sight proximity effect

5.6.1. Strength parameter

We discuss the imprint of radiative transfer on the line-of-sight proximity effect in Ly α spectra as analysed in observations. To this aim, we study a sample of 500 lines of sight originating in the QSO position with randomly drawn directions. We then measure the proximity effect signature.

First an appropriate ω -scale is constructed for each QSO by applying Eq. 5.12 including the SED correction term. In Sect. 5.7.2, we discuss how well the correction term accounts for the SED effect. Given the ω -scale, we then determine the mean transmission in bins of $\Delta \log \omega = 0.5$ for each line of sight. The mean proximity effect signal is calculated by averaging the mean transmission in each $\log \omega$ bin over the QSO lines of sight. We determine the mean optical depth in each bin. The relative change introduced in the optical depth by the QSO is evaluated by computing the normalised optical depth ξ defined as

$$\xi = \frac{\tau_{\text{eff, QSO}}}{\tau_{\text{eff, Ly}\alpha}}, \quad (5.13)$$

where $\tau_{\text{eff, QSO}}$ is the effective optical depth in a $\log \omega$ bin and $\tau_{\text{eff, Ly}\alpha}$ the effective optical depth in the Ly α forest without the influence of the QSO. For $\Delta \log \omega < 1$, we find the resulting normalised optical depths to be weakly dependent on the chosen bin size.

To quantify any difference in the proximity effect signature to the expected one given by Eq. 5.12, we adopt the ‘strength parameter’ as in Dall’Aglio et al. (2008b,a). The strength parameter a is defined as

$$\xi = \left(1 + \frac{\omega}{a}\right)^{1-\beta}, \quad (5.14)$$

where ω is defined by Eq. 5.12. As we know the UVB photoionisation rates in our models, we use the values given in Table 6.1 to define the reference ω . Then values $a > 1$ or $a < 1$ describe weaker or stronger proximity effects, respectively. The SED correction term in Eq. 5.12 can also be interpreted in the framework of the proximity strength parameter.

The strength parameter for each line-of-sight can also be measured individually. By obtaining the proximity effect strength for each line of sight, the strength parameter distribution function can be constructed. Studying its properties provides additional insight into the proximity effect signature and was used by Dall’Aglio et al. (2008b) to derive the UVB photoionisation rate. How well this method performs in the present study is discussed in Sect. 5.8.

5.6.2. Additional models

In addition to the lines of sight obtained from the radiative transfer simulation, we explored two other models allowing us to characterise the imprint of the cosmological density distribution and radiative transfer in the analysis of the proximity effect. These complementary analyses are used to obtain spectra that can be directly compared to our radiative transfer results.

Semi-Analytical Model (SAM) of the proximity effect: From the radiative transfer simulation of only the UV background, we know the ionisation state of hydrogen in the box without the influence of the QSO. We can then semi-analytically introduce the proximity effect along selected lines of sight from the QSO by decreasing the neutral fractions by the factor $(1 + \omega)$. We choose the same 500 lines of sight as in the full radiative transfer analysis.

This model includes the imprint of the fluctuating density field on the spectra. Since the semi-analytical model includes only geometric dilution and the influence of the SED, any differences in the results of the full radiative transfer simulations are indicative of additional radiative transfer effects.

Random Absorber Model (RAM): In our radiative transfer simulations, we can only cover the proximity effect signature up to $\log \omega \geq 0.5$ for the faint QSO and to $\log \omega \geq 1.0$ for the strong one, because of the finite size of our simulation box. To overcome this limitation, we employ a simple Monte Carlo method to generate Ly α mock spectra used in observations to study systematic effects in the data (Fechner et al. 2004; Worseck et al. 2007). The Ly α forest is randomly populated with absorption lines following observationally derived statistical properties. The constraints used are: 1) the line number density distribution approximated by a power law of the form $dn/dz \propto (1 + z)^\gamma$ where $\gamma = 2.65$ (Kim et al. 2007); 2) the column density distribution given by $f(N_{\text{HI}}) \propto N_{\text{HI}}^{-\beta}$, where the slope is $\beta \simeq 1.5$ (Kim et al. 2001); and 3) the Doppler parameter distribution given by $dn/db \propto b^{-5} \exp[-b_\sigma^4/b^4]$, where $b_\sigma \simeq 24$ km/s (Kim et al. 2001). For a detailed discussion of the method used, we refer to Dall’Aglia et al. (2008a,b).

The random absorber model does not take into account the effects of the clustering of absorption lines due to large-scale fluctuations in the density distribution. Therefore we use this model to determine any possible bias in the analysis method and to infer the effect of absorber clustering. Furthermore, since we can generate spectra of a broad wavelength coverage as in observed spectra, we can quantify the effect introduced by the truncated coverage of the ω scale in the spectra from the radiative transfer simulation. With this simple model, it is possible to study any systematics present in the analysis, such as the influence of the $\Delta \log \omega$ bin size on the resulting proximity effect.

5.7. Mean proximity effect

Now we discuss the results of the mean proximity effect as measured in the synthesised spectra.

5.7.1. Mean normalised optical depth

The main results are presented in Figs. 5.9 and 5.13. In both plots, we show the analytical model defined by to Eqs. 5.10 and 5.13 for comparison. The model predicts that at high ω values, the QSO shows the strongest influence and the normalised optical depth is very low (at the QSO where $\omega = \infty$, the normalised optical depth is $\xi = 0$). Then in the range $2 > \log \omega > -1$ the influence of the QSO gradually declines because of geometric dilution and reaches $\xi = 1$ for small ω values. There the Ly α forest is unaffected by the QSO.

5. Cosmological radiative transfer for the LOS proximity effect

The normalised optical depth profile derived from the transfer simulations closely resembles the functional form of the analytical expectations. However, a shift in the profile is present that we quantify using the proximity strength parameter. In observations, the strength parameter is used to measure the UVB photoionisation rate. We therefore faithfully apply the same method to the simulation results. However, since we know the UVB in our models, we can directly compare the photoionisation rate determined using the strength parameter with the model input. The strength parameter is determined by fitting Eq. 5.14 to the data points using the Levenberg-Marquardt algorithm and a least absolute deviation estimator weighted by the error in the mean. To obtain a reasonable fit, data points deviating strongly from the profile above $\log \omega > 2$ were excluded. This also excludes the unreliable region within 1 Mpc of the QSO in the radiation transfer simulation, which translates to an $\log \omega \geq 2.4$ for the faint QSO and $\log \omega \geq 3.4$ for the bright one. We provide the resulting strength parameters in Table 5.2 and include the fitted profiles in all plots of the mean proximity effect. The shift is due to the large-scale environment around the QSO hosts, which we discuss in more detail in Sect. 5.7.3.

For each data point, we show two error bars, the standard deviation in the sample and the 2σ error in the mean. The sample standard deviation indicates the variance between the 500 lines of sight. These variances are very large and extend to values above $\xi > 1$. Values of $\xi > 1$ arise when saturated absorption systems dominate the optical depth in a $\log \omega$ bin and increase the optical depth above the mean effective optical depth in the forest. A detailed illustration is given in the Appendix. Since the real space distance covered by the $\log \omega$ bins increases with decreasing ω , there the contribution of saturated systems and the variance is lower.

On the other hand, the error in the mean is small because of to our sample size of 500 lines of sight, thus the mean profiles are well determined. However, the mean profile does not strictly follow the smooth analytical model, but displays strong fluctuations. These fluctuations are enhanced by our being able to analyse the QSO only in one octant. We also focus our analysis on merely one object, making the signal sensitive to the surrounding density distribution. If multiple halos were combined to one mean profile, the fluctuations would diminish and the profile would follow the smooth analytical profile more closely. Since the data points deviate from the analytical profile, the determination of the strength parameter is quite uncertain. To estimate a formal error in the strength parameter, we applied the Jackknife method. The error values are provided in Table 5.2.

We verified whether the Monte Carlo noise of the radiation transport simulation contributes to the signal's variance by reducing the QSO's photon packages production by half. This does not contribute noticeably to the variance in the signal. The large scatter in the proximity effect signal therefore originates in the distribution of absorption systems along the line of sight. It is a direct imprint of the cosmic density inhomogeneities. The probability distribution of the normalised optical depths in each $\log \omega$ bin is approximately log-normal as shown in Fig. 5.16.

5.7.2. Performance of the SED correction term

With our radiative transfer simulations, we can test whether the SED correction term in Eq. 5.12 describes the effect of a QSO and UV background that have different spectral shapes. The SED adds to the mentioned shift in the normalised optical depth, ξ , already

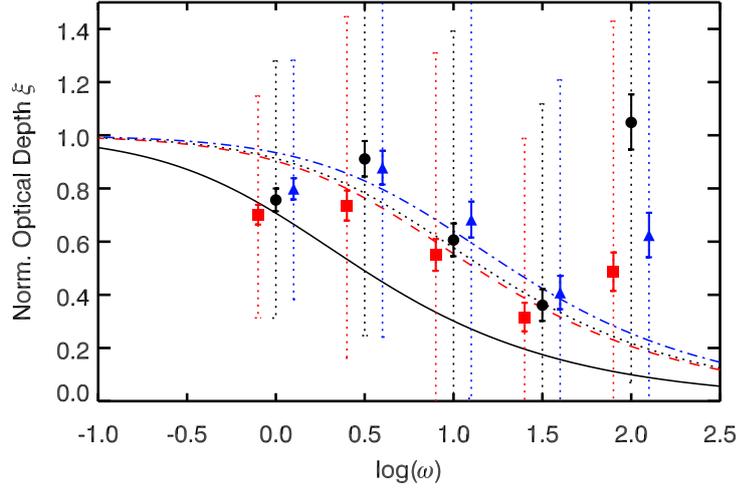


Figure 5.9.: The normalised optical depth for the $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO with the standard SED residing in a halo at $z = 3$. Continuous lines show the fitted profiles. The triangles and dash-dotted line indicate the radiation transfer proximity effect without the use of the SED correction term. The circles and dotted line provide the radiation transfer results including the SED correction. The QSO with the hard $\alpha_q = -1.5$ SED is given by the squares and dashed line. The original BDO model is shown as the solid line. The dotted error bars show the sample standard deviation. The solid errors show the 2σ error in the mean.

introduced by the large-scale density distribution around the QSO hosts. To identify the SED influence, we re-simulate the $z = 3$ faint QSO residing in a halo with a $\alpha_q = -1.5$ power-law SED. With the background's SED slope of $\alpha_b = -1.28$, we expect to detect only a slight shift caused by the different slopes of $\Delta \log a = 0.02$. The softer Trammell et al. (2007) SED used in the rest of the chapter yields a shift of $\Delta \log a = 0.11$. If the SED correction term performed as intended, the corrected proximity effect profile of the soft SED QSO would lie approximately on top of the one with the hard SED.

In Fig. 5.9, we present the results of this experiment, where we plot the proximity effect profile of the $\alpha_q = -1.5$ QSO, and SED-corrected and uncorrected profile for the QSO with $\alpha_q \approx -2.5$. The data points of the corrected $\alpha = -2.5$ profile lie almost on top of the $\alpha_q = -1.5$ profile as expected. A similar picture emerges by looking at the strength parameters. The uncorrected soft SED yields a strength parameter of $\log a = 0.84 \pm 0.06$. By including the correction term, we obtain $\log a = 0.70 \pm 0.12$. For the $\alpha_q = -1.5$ SED, we obtain a strength parameter of $\log a = 0.65 \pm 0.05$. The different values of the SED-corrected and the soft QSO results are consistent within the error bars.

5.7.3. Influence of the large-scale environment

The environmental density around the QSO has been shown to strongly influence the proximity effect. By studying random DM halos in a mass range of from $1.4 \times 10^{11} M_{\odot} h^{-1}$ to $9 \times 10^{12} M_{\odot} h^{-1}$, Faucher-Giguère et al. (2008b) detected an enhancement in the mean overdensity profiles up to a comoving radius of $(10 - 15) \text{ Mpc } h^{-1}$. Only at larger radii does the average overdensity profiles exceed the cosmic mean. The overdensity profiles fluctuate strongly between different halos. Since an enhanced density directly translates

5. Cosmological radiative transfer for the LOS proximity effect

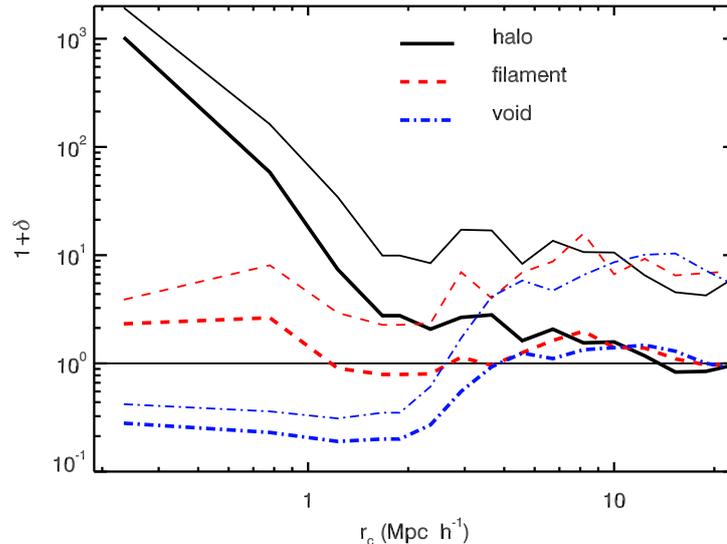


Figure 5.10.: The mean overdensity profiles of the halo (solid line), filament (dashed line), and void (dash-dotted line) environment at redshift $z = 3$ as a function of comoving radius. The mean profiles are shown by thick lines. The thin lines indicate the 1σ fluctuations in the density field. An excess in overdensity is seen for all environments in the range of $3 < r_c < 10 \text{ Mpc } h^{-1}$, where the profile stays above the average overdensity.

into an enhancement of the Ly α forest optical depths, the proximity effect measurement is biased accordingly.

We determined the mean overdensity profiles and the fluctuations around the mean for our three QSO host environments. The resulting mean profiles are shown in Fig. 5.10 for $z = 3$. Our density profiles show two different influences.

First, the local environment around the QSO host leaves an imprint on the profiles up to a radius of $r \approx 2 \text{ Mpc } h^{-1}$. This is seen in the halo case where the massive host halo is responsible for a strong overdensity that steadily declines up to $r \approx 2 \text{ Mpc } h^{-1}$. The contrary is seen in the void case, where up to $r \approx 3 \text{ Mpc } h^{-1}$ the underdense region of the void is seen. At larger radii, the density increases quite strongly.

All our environments display an overdensity between $r \approx 3 \text{ Mpc } h^{-1}$ and a radius of around $r \approx 15 \text{ Mpc } h^{-1}$. In this region, the density lies above the cosmic mean. We refer to this phenomenon as a large-scale overdensity. Further away from the QSO, the cosmic mean density is reached. The same behaviour is seen at higher redshifts. There, however, the amplitude of the large-scale overdensity is lower. The presence of the large-scale overdensity in the case of the most massive halo is a direct consequence of the most massive halo forming where there is a large overdensity. However, around the filament and the void, this a large-scale overdensity is a random selection effect.

An excess in density translates into an excess of Ly α optical depth in this region compared to the mean optical depth. Thus, the normalised optical depth ξ increases and leads to a biased proximity effect. The immediate environment around the quasar extends to $\log \omega \approx 1.5$ for the faint source and to $\log \omega \approx 2.5$ for the bright one. At smaller $\log \omega$, the large-scale overdensity influences the proximity effect signal up to $\log \omega \approx 0.3$ for the faint and to $\log \omega \approx 1.3$ for the bright QSO.

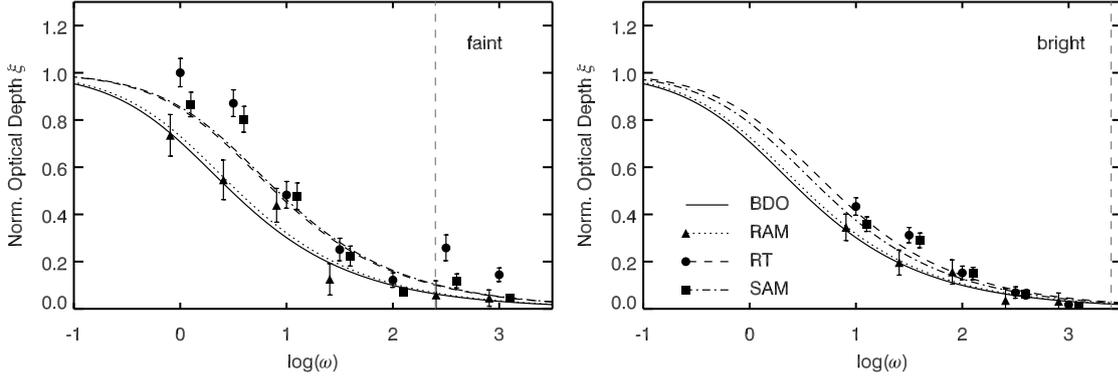


Figure 5.11.: Comparison of the proximity effect at redshift $z = 3$ of a) the RAM (triangles and dotted line), b) the semi-analytical model (circles and dash dotted line), and c) the results from the radiative transfer simulation (squares and dashed line). The solid line shows the BDO reference analytical model. The dashed vertical line marks the start of the unreliability region of $\log \omega \geq 2.4$ for the faint QSO and $\log \omega \geq 3.4$ for the bright QSO. Results are shown for a QSO sitting in a filament. The left panel is for the low luminosity source, and the right one for the high luminous one. We show only the 2σ errors in the mean.

Motivated by these considerations, we study the influence of the cosmic density fluctuations on the proximity effect profile. To understand whether radiative transfer effects or the large-scale overdensity are responsible for the shift in the normalised optical depth shown in Fig. 5.9, we use the reference models introduced in Sect. 5.6.2.

We show the resulting proximity effect profiles for the reference models of the faint and bright QSO residing in the filament at $z = 3$ in Fig. 5.11. We use the RAM to infer the bias introduced by the analysis method itself because of the finite size of our spectra. In addition, the SAM allows us to determine the influence of the fluctuating density field by omitting any possible radiative transfer effect. On top of that, we show the radiative transfer results.

The data points of the RAM follow the standard BDO profile well for both luminosities. However, a slight shift is seen in the fitted profiles. The strength parameters are $\log a = 0.06 \pm 0.05$ in the case of the weak QSO and $\log a = 0.05 \pm 0.10$ for the strong QSO. Within the error bars, this is consistent with the analytical predictions. A small shift in the signal was found to be intrinsic to the analysis method because of the asymmetry that it produced in the ξ -distribution (Dall’Aglio & Gnedin 2010).

As for the radiative transfer results, those for the SAM display a strong shift towards the QSO. This is a clear indication that the cosmological density distribution is responsible for the large shift. Comparing the SAM with the radiative transfer results reveals no strong difference. The ratio of the strength parameter of the two models $\log(a_{\text{RT}}/a_{\text{SAM}})$ for the faint QSO is -0.1 and for the bright one 0.1 . Considering the uncertainties in the strength parameters, the two models yield identical results. We provide the ratios of the strength parameters of the two models for all the studied environments and redshifts in Table 5.2. The ratios vary widely between the different models. However, owing to the uncertainties in the strength parameter, we attribute these differences to the inability of the analytical model to account for the large-scale density distribution.

To strengthen this conclusion, we performed one additional test, where we applied

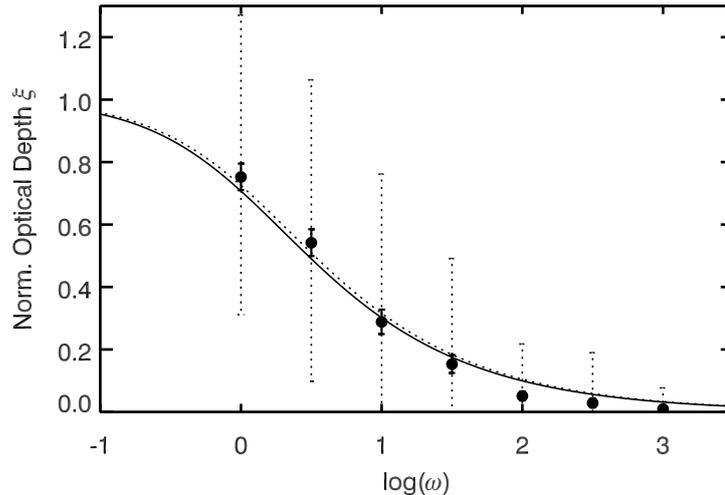


Figure 5.12.: The combined proximity effect signature of the SAM applied to 500 randomly selected lines of sight with random origin. A $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO is assumed. The continuous line shows the fit to the analytical profile to the data.

the SAM to 500 randomly selected lines of sight of random origins and directions in the box and not only one specific origin. The resulting proximity effect profile is shown in Fig. 5.12. The profile now smoothly follows the analytical model without any significant fluctuations around it. The strength parameter of $\log a = 0.11 \pm 0.10$ is also consistent with the SED-corrected BDO model. The analytical model is thus a good description of lines of sight with random origins and directions.

Thus the large shift in our simulations is solely due to large-scale overdensities around our sources. We cannot identify significant radiative transfer effects in addition to the one modelled with the SED correction term, since the SAM model gives similar results to the radiative transfer simulations.

5.7.4. Influence of the host environment

We present the resulting proximity effect profiles for the halo, filament, and void QSO host environments as a function of redshift in Fig. 5.13. The corresponding strength parameters are given in Table 5.2. In the analysis, the unreliability region of $\log \omega \geq 2.4$ for the faint QSO and $\log \omega \geq 3.4$ for the bright QSO were excluded. As we have seen in the previous section, the local host environment of the QSO extends to a comoving radius of $r \approx 2 \text{ Mpc } h^{-1}$. Hence, we expect the local environment to affect the proximity effect profiles up to $\log \omega \approx 1.5$ for the faint QSO and to $\log \omega \approx 2.5$ for the bright one.

The effect of the local environment is significant in the halo case. Except for the $z = 4.9$ bright QSO, the ξ -profile steepens strongly in the direct vicinity of the QSO. The ξ -values start to rise to such an extent, that they lie beyond the plotted region and deviate strongly from the analytical model. In addition, the fluctuations in the normalised optical depth strongly increase. This departure from the analytic profile turns out to pose a problem to the determination of the strength parameter. To obtain a reasonable fit, the region of the local environment needs to be excluded.

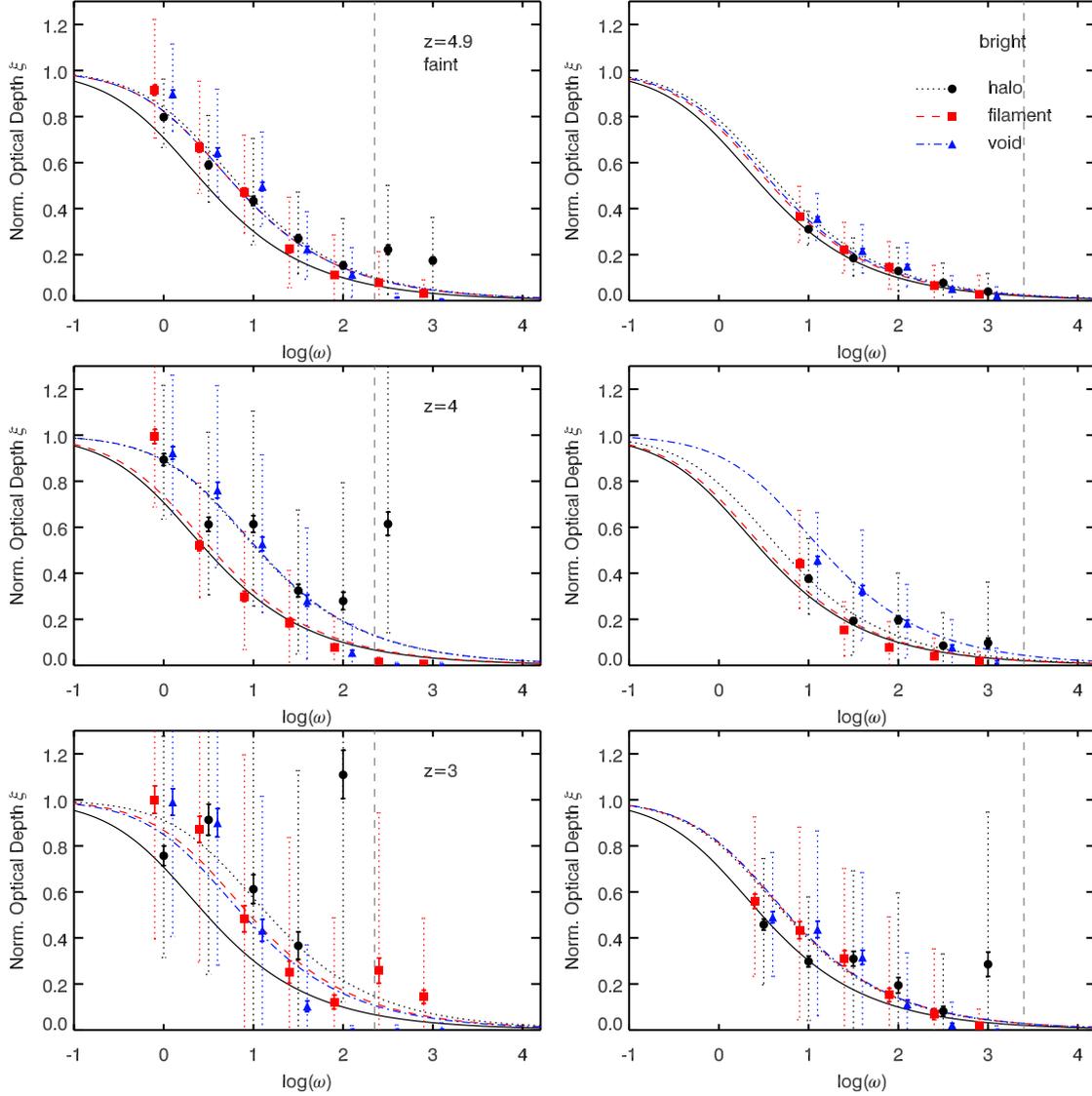


Figure 5.13.: The combined proximity effect signature of all 500 lines of sight analysed with the BDO model showing the dependence on environment, luminosity, and redshift. The left column shows the case of a faint $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO, the right one of a luminous $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO. From top to bottom, redshifts are 4.9, 4, and 3. The solid lines show the geometrical dilution model, the other lines are best fits to the measurements with parameters given in Table 5.2. The circles and dotted lines show the halo cases, the squares and dashed lines the filament case and triangles and dash-dotted lines the void case. Symbols are horizontally shifted by ± 0.05 for better visibility. The dotted error bars show the sample 1σ standard deviation. The solid errors give the 2σ error in the mean. The dashed vertical line marks the start of the unreliability region of $\log \omega \geq 2.4$ for the faint QSO and $\log \omega \geq 3.4$ for the bright QSO.

5. Cosmological radiative transfer for the LOS proximity effect

Table 5.2.: Parameters of the $\log a$ fits to the combined proximity profile for the semi-analytical model (SAM) and the radiative transfer simulation (RT). We indicate the 1σ errors.

z	$L_{\nu_{\text{LL}}}^{\dagger}$	$\log(a_{\text{RT}}/a_{\text{SAM}})$			$\log a_{\text{RT}}$		
		halo	fil.	void	halo	filament	void
4.9	10^{31}	0.1	0.1	0.2	0.4 ± 0.1	0.3 ± 0.1	0.3 ± 0.1
4.0	10^{31}	0.3	0.0	0.3	0.6 ± 0.1	0.1 ± 0.1	0.6 ± 0.1
3.0	10^{31}	0.3	0.0	0.1	0.7 ± 0.1	0.5 ± 0.1	0.4 ± 0.2
4.9	10^{32}	0.2	0.1	0.1	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 0.1
4.0	10^{32}	0.4	0.0	0.7	0.2 ± 0.2	0.1 ± 0.1	0.7 ± 0.2
3.0	10^{32}	0.5	0.1	0.1	0.3 ± 0.3	0.3 ± 0.1	0.3 ± 0.1

†: $\text{erg Hz}^{-1} \text{s}^{-1}$

For the filament and void environment however, this departure is not seen as expected. Only in the profile of the $z = 3$ faint QSO in the filament environment is there an enhancement at $\log \omega \approx 2.5$. In the filament and void environments, the scatter in the normalised optical depth decreases. The lack of density enhancements in the immediate surroundings of the host thus reduces the scatter. The influence of the local environment also decreases with increasing luminosity, which is not only because of the lower opacity, but also the larger spatial coverage of each $\log \omega$ bin.

Beyond the local environment, the ξ -values return to the smooth profile expected from the analytical model. At $z = 4.9$, the different environments have identical profiles. With decreasing redshift however, the profiles in the different environments start to deviate from the analytical profile without any clear trend. Similar deviations can also be seen in the corresponding strength parameters. For example, the results for the faint halo and void at $z = 4$ agree. In contrast, for the bright QSO, the strength parameters of the halo and the void disagree. The halo and the filament results agree within the error bars. Owing to the limitations of the analytical model to account for the large-scale overdensity present in the simulation data, the fitted strength parameters exhibit large unsystematic fluctuations. In Sect. 5.5, we have seen that the local environment does not cause deviations from the geometric dilution model beyond 2 Mpc h^{-1} . The overionisation profile was not found to be influenced by the halo, filament, or void environment at large radii. Therefore we conclude that the local environment does not influence the proximity effect profile at radii larger than the local host environment. Furthermore, the fluctuations in the strength parameter do not correlate with the source environment and solely arise from the analytical model not taking the cosmological density distribution into account.

5.7.5. Redshift evolution

Studying the strength parameters, no apparent evolution with redshift can be seen, if the fitting errors are taken into account. However, the uncertainties in the strength parameter decrease with increasing redshift. Only in the faint halo case is there a hint of a decrease in proximity strength with increasing redshift.

By looking at the ξ -profiles in Fig. 5.13, a clear decrease in the fluctuations around the mean profile can be seen with increasing redshift. This is because the density fluctuations of the dark-matter are not as well developed at high redshifts as at lower ones. A minor effect may result from the decreasing transmission of the Ly α forest at higher redshift, i.e. the narrow range of gas densities probed by the Ly α forest. The analysis of the large-scale overdensity discussed in Sect. 5.7.3 demonstrates however the dominance of the evolution of the density field. There is a slight tendency of the profile to be stronger shifted at low redshifts than at $z = 4.9$. This is especially prominent in the case of the $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO.

5.7.6. Luminosity dependence

The environmental bias of the proximity effect signature depends on the QSO luminosity. For all redshifts that we have studied, the shift in the profile is smaller for the bright QSO than for the weak one. The variance in the signal also decreases with increasing QSO luminosity. This can be most clearly seen for redshifts $z = 4.9$ and $z = 3$.

The cause of the luminosity dependence is that the ω scale is a function of the QSO luminosity. Therefore an increase in the QSO luminosity translates into larger distances covered by each $\log \omega$ bin. This dampens the influence of the fluctuating density field and thus the variance in ξ . Hence luminous QSOs help to constrain more tightly the proximity effect signature.

5.8. Proximity effect strength distribution

We now discuss our results for the distribution of the strength parameter determined on individual lines of sight. The asymmetry of the ξ -distribution and the large-scale overdensity biases the mean proximity effect profiles, as we have discussed in the previous Section. To avoid this bias, Dall’Aglio et al. (2008a) used the proximity strength distribution along individual lines of sight to measure the UVB flux, and proposed that the peak of the distribution provides a measure of the UVB. We test the proposed method with our simulations.

The proximity effect strength distribution is obtained by determining the ξ -profile for each individual line of sight and by fitting Eq. 5.14 to the profile. In contrast to the last section, we use all available ξ data points (except the ones inside the numerically unreliable region) in the fitting procedure. Whenever the simulated data deviates strongly from the analytical model because of the large variance in the ξ -values, it is not possible to determine a strength parameter. This problem is significant for the faint QSO residing in a halo, where the proximity effect profile deviates from the analytical one at $\log \omega \geq 1.5$. There the fitting procedure failed in most cases and we cannot derive a strength parameter distribution.

In Fig. 5.14, we present the proximity effect strength distribution from our simulations. Only those lines of sights contribute to the distribution, for which a strength parameter could be determined, although we always normalise the distributions $P(a)$ to unity. We employ the comparison models to investigate the various effects that alter the strength distribution. To compare with the results obtained from the mean proximity effect profile, we mark the strength parameters of the mean profiles as given in Table 5.2.

5. Cosmological radiative transfer for the LOS proximity effect

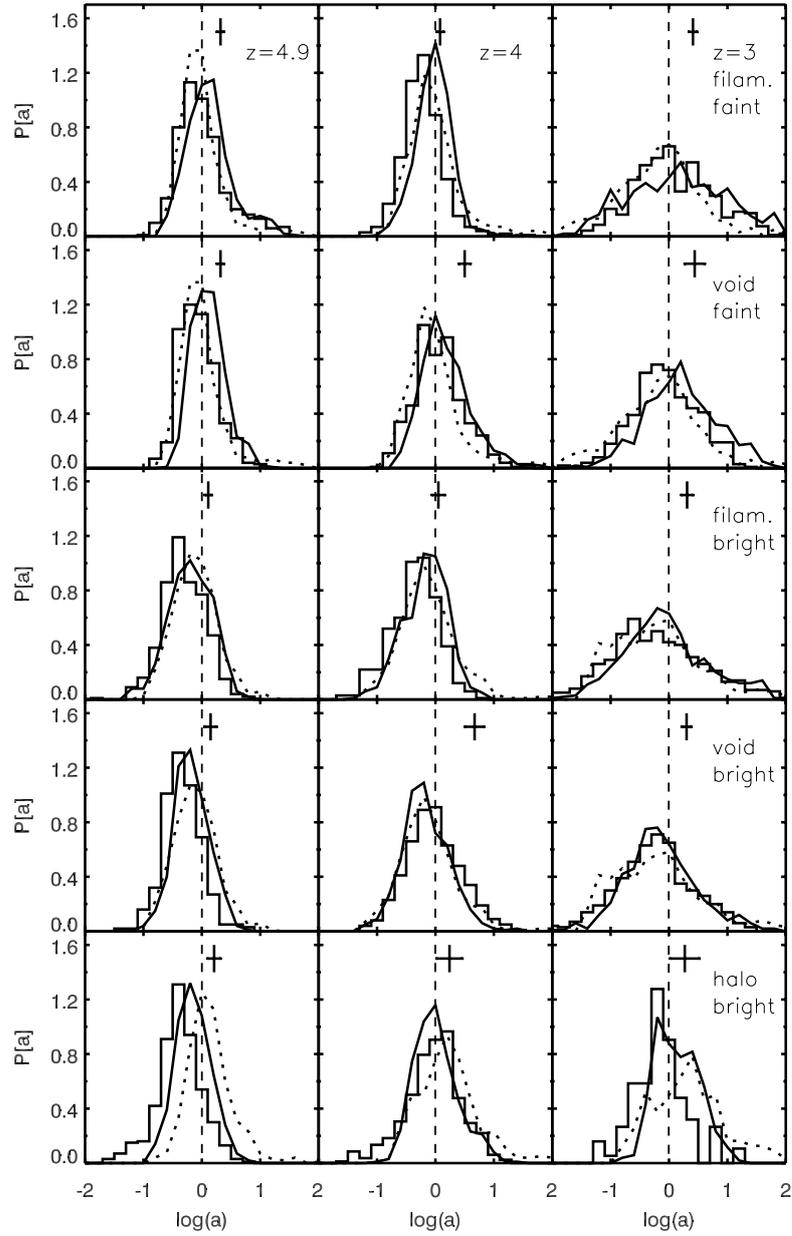


Figure 5.14.: The distribution of the strength parameter $\log a$ of the proximity effect along the different lines of sight for the radiative transfer simulations (solid histogram), compared with RAM (dotted line) and the SAM (solid line). The mean a -values from the radiative transfer simulations (Table 5.2) and their uncertainties are shown as vertical lines on top of each panel. The upper two panels show the low luminosity case. The lower three panels provide the results for the luminous QSO. From left to right, the redshifts are 4.9, 4, and 3. The vertical dashed line marks $\log a = 0.0$.

All our models display a strength distribution that is clearly peaked and has a large scatter of $\sigma(\log a) = 0.3, 0.4,$ and 0.7 at redshifts $z = 4.9, 4, 3$. This scatter is much larger than the variance in the $\log a$ values of the mean analysis. All the models show a similar strength distribution as the respective ones of the RAM. In the case of the faint QSO, they all peak around $\log a = 0$. For the bright QSO, the peak is somewhat shifted to negative $\log a = -0.1$. This shift is also present in the RAM results. We therefore suspect that it is due to the small ω scale coverage for the bright QSO and the corresponding uncertain fits. It follows that the peak of the strength distribution can be used to derive unbiased estimates of the proximity effect.

The strength distribution is thus surprisingly not sensitive to the large-scale overdensity environment discussed in the mean analysis. In the averaging process for deriving the mean profile, we are biased by the spherically averaged overdensities shown in Fig. 5.10. By fitting the strength parameter individually, less weight is given to single $\log \omega$ bins deviating from the expected profile. This confirms the claim of Dall’Aglia et al. (2008a) that using the proximity effect along individual lines of sight leads to a stable unbiased estimate of the proximity effect and thus of the UVB.

As in the mean analysis, we can investigate the influence of the local environment, the quasar redshift, and the quasar luminosity. The shape of the strength parameter distribution is not influenced by the QSO’s local environment. Whether the QSO resides in a halo, a filament, or a void, the maximum, median, and width of the distribution is not affected. Only for the faint quasar does the halo environment strongly influence the proximity effect profile so that in many cases no reasonable fit could be given. The width of the distribution is found to increase strongly with redshift. This is clearly a consequence of the increasing density inhomogeneities. Overall we find a large uncertainty in the proximity effect along individual sight lines. This is due to the limited ω -scale coverage of the simulated spectra, which are much smaller than in the observations analysed in Dall’Aglia et al. (2008a).

5.9. Conclusions

We have performed high resolution Monte Carlo radiative transfer simulations of the line-of-sight proximity effect. Using dark-matter-only simulations and a semi-analytic model of the IGM, we have applied our radiative transfer code in a post-processing step. The aim of this chapter was to identify radiative transfer influence on the proximity effect for QSOs with different Lyman limit luminosities, different host environments, and at different redshifts.

Owing to the low optical depths in the Ly α forest, we demonstrated that the on-the-spot approximation is insufficient in treating the proximity effect. Diffuse radiation from recombining electrons significantly contribute to the ionising photon budget over large distances and need to be taken into account in the simulations.

In the radiative transfer simulations, we identified Lyman limit systems that cast shadows behind them in the overionisation profile. Any Lyman limit system between the QSO and the observer will result in a suppression of the proximity effect. Behind such a system, the Ly α forest is not affected by the additional QSO radiation and a detection of the proximity effect is not possible. However, Lyman limit systems appear only along

5. Cosmological radiative transfer for the LOS proximity effect

rare lines of sight and influence only marginally the proximity effect statistics.

Mock spectra have been synthesised from the simulation results, and methods similar to those used in observations have been applied to extract the observable proximity effect signature. We compared these results with the widely used assumption of the BDO model that only geometric dilution governs the overionisation region around the QSO. We have also used a random absorber model applied extensively in the analysis of observational data. Furthermore, a semi-analytical model was employed to analytically model the proximity effect signal onto spectra drawn from the UVB-only simulations.

We have shown that differences in the shape of the UVB spectral energy distribution function and the spectrum of the QSO introduce a bias in the proximity effect strength. In the original formulation of the BDO model, the SED of the two quantities was assumed to be equal. Differences in the SED of the QSO and the UVB result in a weakening or strengthening of the proximity effect. This bias can be analytically modelled by introducing a correction term in the BDO formalism. We have tested the performance of the correction term. The correction term accounts mostly but not completely for the effect. However, the remaining difference lies within the precision of the determination of the strength parameter.

In all simulations and models, we have confirmed a clear proximity effect profile. However, we found a strong scatter between different lines of sight. Cosmological density inhomogeneities introduced strong fluctuations in the normalised optical depths of the proximity effect. They are caused by individual strong absorbers along the line of sight. The influence is significant near the QSO and can dominate over the standard proximity effect profile if the QSO resides in a massive halo. The fluctuations in the normalised optical depth are smaller, if more luminous QSOs are considered. There the real space coverage of each $\log \omega$ bin is larger and therefore the fluctuations decrease. In addition the fluctuations decrease with increasing redshift.

We have studied the influence of different QSO host environments on the proximity effect. The QSO was placed in the most massive halo, in a random filament, and in a random void. The local host environment up to a scale of $r \approx 2 \text{ Mpc } h^{-1}$ is responsible for deviations from the BDO model near the QSO. However, it does not affect the proximity effect at radii larger than $r \approx 2 \text{ Mpc } h^{-1}$, and the smooth proximity effect signature was regained.

In the three environments selected for this study, random density enhancement on scales up to $r \approx 15 \text{ Mpc } h^{-1}$ were present. These overdensities introduced a bias in the mean analysis, that decreases the proximity strength. It is a strong effect, clearly seen by averaging the profiles over the 500 lines of sight and then estimating the mean profile. This bias will affect any UVB measurement made using this method. The influence of the bias decreases with increasing QSO luminosity.

Our radiative transfer results complement the analysis of the proximity effect zone performed using a semi-analytical model by Faucher-Giguère et al. (2008b). There the authors find a significant dependence of the combined proximity effect on the QSO host halo both by the action of the absorber clustering and the associated gas inflow. Coinciding with our analysis, a stronger proximity effect signal is found at lower redshift because of the higher gas density inhomogeneities. The authors also discuss the nearly log-normal probability distribution of the flux decrement, and they find a strong dependence of the average overionisation in the near zone around the QSO host on the halo mass on

scales up to 1 Mpc proper radius.

The proximity effect strength distribution derived along single lines of sight does not show a similar dependence on the large-scale overdensities. The distributions were always consistent with the random absorber model. This is consistent with the results of the analysis of real QSO spectra and the resulting unbiased estimate of the UV-background in Dall’Aglio et al. (2008a). We found that the distributions are slightly shifted to negative $\log a$ values, since the ω range covered in the simulation was smaller for the high luminosity QSO. The fluctuations in the normalised optical depth on single lines of sight are responsible for a large scatter in the derived strength distribution. The scatter increases with decreasing redshift. Since the ω scale in our analysis is limited, the fitted strength parameters show large uncertainties. This contributes to the width of the strength parameter distribution. A complete coverage of the proximity effect region will reduce these uncertainties in the strength distributions.

To reliably study all the influences discussed here, the full proximity effect region needs to be simulated and we have to make our radiative transfer code more efficient. It should also be noted that helium was not included in our simulations, which could play an important role in softening the QSO spectra. Helium would absorb a fraction of the hard flux produced by the QSO and increase the influence of the QSO SED.

We finally conclude that the QSO host environment, i.e. whether it sits in a halo, a filament, or a void influences the proximity strength only locally. Apart from shadowing by Lyman limit systems, the proximity effect does not show any radiative transfer influence other than the SED bias.

5.10. Strong absorber contamination of the normalised optical depth

In Section 5.7 we have discussed the single line of sight variance of the normalised optical depth ξ and its dependency on luminosity and redshift. In this Appendix we illustrate where and how this variance arises.

The normalised optical depth ξ is the ratio of the effective optical depth in the QSO spectra measured in $\log \omega$ bins to the effective optical depth present in the unaffected Ly α forest. If the optical depth in a $\log \omega$ bin is larger than in the unaffected forest, ξ will be larger than unity. The range covered by the $\log \omega$ bins in the spectra decreases towards the QSO. Therefore any strong absorption system present in the small ranges covered by high $\log \omega$ bins will bias the optical depth in the bins and therefore increase ξ to values above unity. This bias gives rise to the large variance in ξ at large $\log \omega$. It is interesting to note, that an increase in the QSO Lyman luminosity by one dex also increases the proper distance covered in a $\log \omega$ bin by one dex. Thus the influence of strong absorbers is lower, the more luminous the QSO is.

In Fig. 5.15, we show five spectra taken from the $z = 3$ QSO situated in a void with a luminosity of $L_{\nu_{LL}} = 10^{31}$ erg Hz $^{-1}$ s $^{-1}$. The spectra are plotted in $\log \omega$ space and the normalised optical depth measured in the $\log \omega$ bins is overplotted by solid dots. The spectrum at the bottom panel shows an uncontaminated proximity effect profile. However, all the others panels illustrate the influence of high column density systems on the profile. In the second and third spectra, prominent absorption lines elevate ξ to

5. Cosmological radiative transfer for the LOS proximity effect

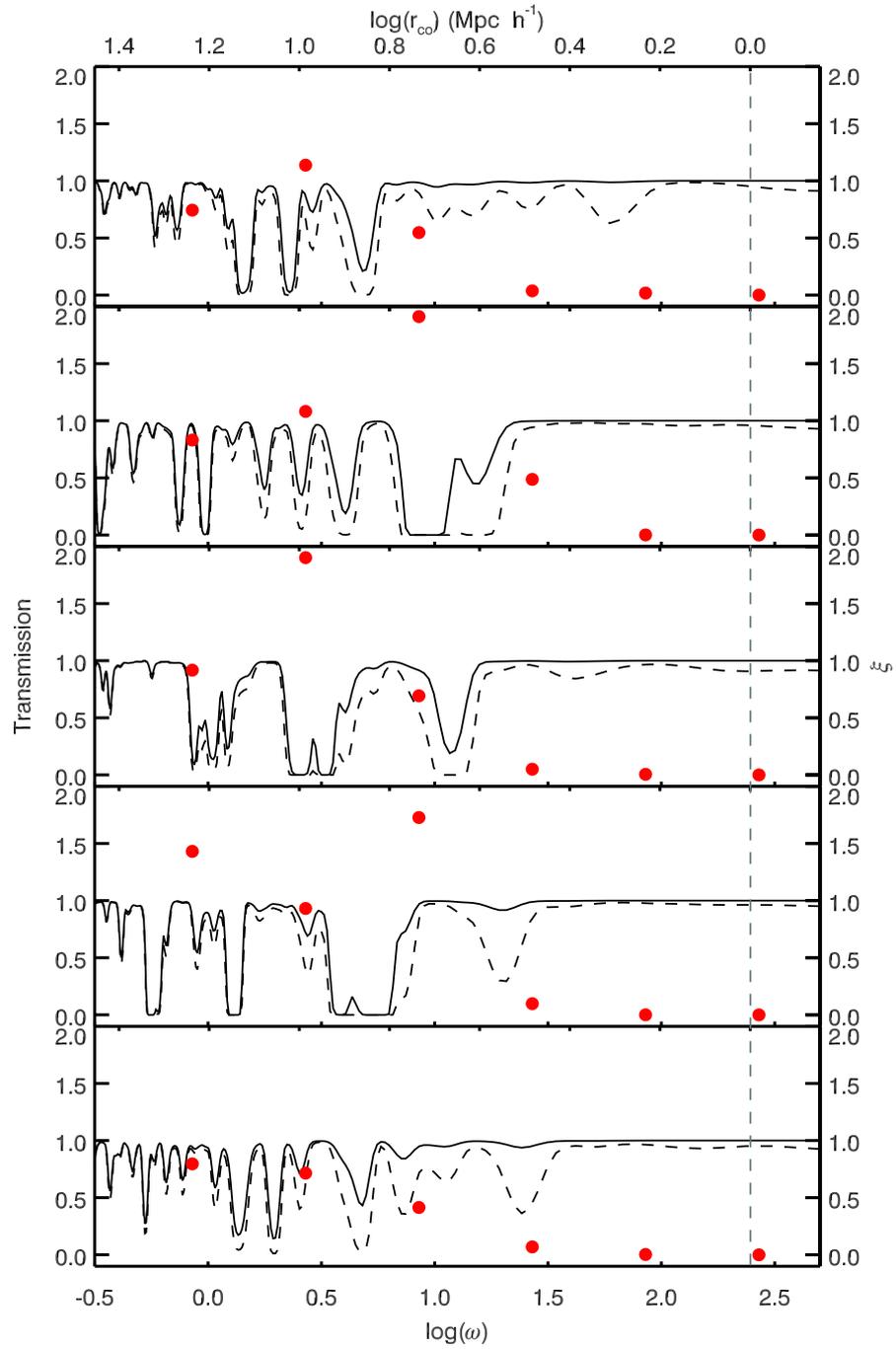


Figure 5.15.: Sample spectra taken from the $z = 3$ low luminosity void QSO showing the radiative transfer simulations including the proximity effect (solid line) and without it (dashed line). The spectra are given in $\log \omega$ scale and for better visibility, no noise has been added to the spectra. The normalised optical depths ξ at each $\log \omega$ bin is overplotted by solid dots to show the effect of absorption features on ξ . The dashed vertical line marks the start of the unreliability region of $\log \omega \geq 2.4$.

5.10. Strong absorber contamination of the normalised optical depth

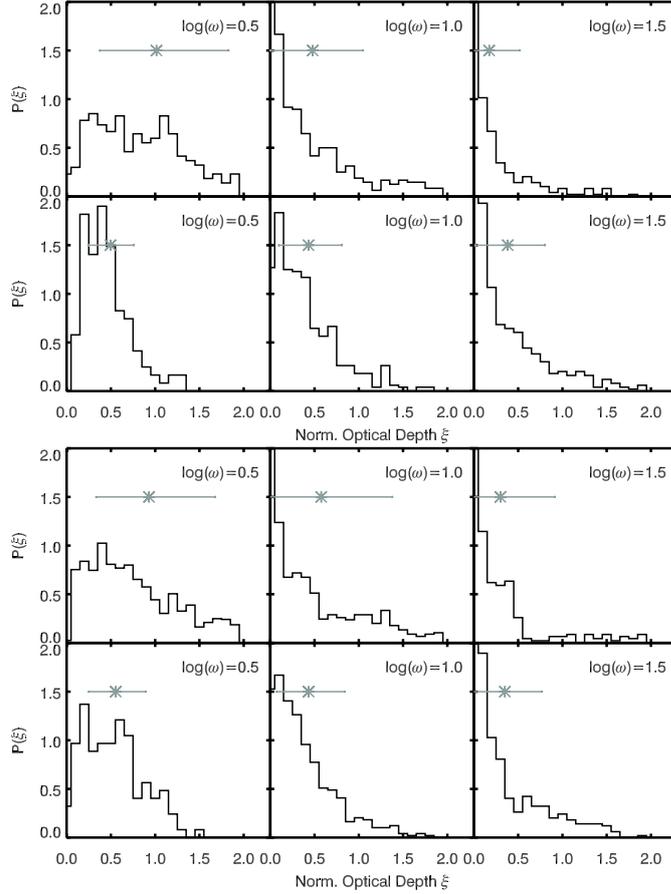


Figure 5.16.: Normalised optical depth ξ distribution of the void (left panel) and filament QSO (right panel) at redshift $z = 3$ for three $\log \omega$ bins. The upper panels show results for the $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSOs, the lower one for the $L_{\nu_{\text{LL}}} = 10^{32} \text{ erg Hz}^{-1} \text{ s}^{-1}$ QSO. The grey data points give the average normalised optical depths with their variance as shown in Fig. 5.13.

$\xi = 2$.

The distribution of ξ in a $\log \omega$ bin arising from the influence of strong absorbers varies from bin to bin. In Fig. 5.16, we show the ξ distribution in three bins for QSOs located in a void and a filament at redshift $z = 3$. Overplotted are the mean values and the corresponding standard deviations shown in Fig. 5.13. We note that for the bright QSO the $\log \omega = 0.5$ distribution is incomplete, since not all lines of sight reach the proper distance needed for this bin.

The distributions are anisotropic with an extended tail to large optical depth. This behaviour reflects the probability distribution of the simulated density field. The distribution of the high luminosity QSO is narrower than the one of the low luminosity QSO. This is again due to the larger distances covered by the $\log \omega$ bins in the high luminous case. The width of the distribution covers a large ξ range. The contamination of a single line of sight by strong absorption systems is thus an issue that can only be eliminated by analysing large number of QSO sight lines to reliably determine the mean proximity effect profile.

*Mancher klopft mit dem Hammer an der Wand herum
und glaubt, er treffe jedes mal den Nagel auf den Kopf.*

Johann Wolfgang von Goethe

6

Large scale environmental bias on the QSO line of sight proximity effect.¹

We analyse the overionisation or proximity zone of the intergalactic matter around high-redshift quasars in a cosmological environment. In a box of $64 h^{-1}$ Mpc base length we employ high-resolution dark matter only simulations with 1024^3 particles. For estimating the hydrogen temperature and density distribution we use the effective equation of state by Hui & Gnedin (1997). Hydrogen is assumed to be in photoionisation equilibrium with a model background flux which is fit to recent observations of the redshift dependence of the mean optical depth and the transmission flux statistics. At redshifts $z = 3, 4$, and 4.8 , we select model quasar positions at the centre of the 20 most massive halos and 100 less massive halos identified in the simulation box. From each assumed quasar position we cast 100 random lines of sight for two box length including the changes in the ionisation fractions by the QSO flux field and derive mock Ly α spectra. The proximity effect describes the dependence of the mean normalised optical depth $\xi = \tau_{\text{eff, QSO}}/\tau_{\text{eff, Ly}\alpha}$ as a function of the ratio of the ionisation rate by the QSO and the background field, $\omega = \Gamma_{\text{QSO}}/\Gamma_{\text{UVB}}$, i.e. the profile $\xi = (1 + \omega/a)^{-0.5}$, where a strength parameter a is introduced. The strength parameter measures the deviation from the theoretical background model and is used to quantify any influence of the environmental density field. We improve the statistical analysis of the profile fitting in employing a moving average to the profile. We reproduce an unbiased measurement of the proximity effect which is not affected by the host halo mass. The scatter between the different lines of sight and different quasar host positions increases with decreasing redshift, $\sigma_{\log a} \approx 0.08, 0.20$ and 0.36 for $z = 4.8, 4$, and 3 , respectively. Around the host halos, we find only a slight average overdensity in the proximity zone at comoving radii of $1 < r_c < 10 h^{-1}$ Mpc. However, a clear power-law correlation of the strength parameter with the average overdensity in r_c is found, showing an overestimation of the ionising background in overdense regions and an underestimation in underdense regions.

6.1. Introduction

The Lyman alpha (Ly α) forest in high-resolution quasar (QSO) spectra blueward of the QSO Ly α emission represents an excellent tracer of the high redshift matter distribution. According to detailed simulations and a large observational material (cp. e.g. Meiksin 2009) the forest of lines is generally ascribed to the absorption by a tiny fraction of remaining neutral hydrogen H I. This H I forms the cosmic web at low gas overdensities of a factor of a few and arises naturally due to gravitational instabilities (Petitjean et al. 1995; Miralda-Escudé et al. 1996; Hui et al. 1997) in a standard Λ CDM cosmological

¹This work is accepted by MNRAS for publication and has been prepared in collaboration with Müller, V.; Gottlöber, S.; Yepes, G.

6. Large scale environmental bias on the LOS proximity effect.

model. The intergalactic gas distribution and its ionisation state forms from a balance of the ionising intergalactic UV radiation stemming from galaxies and quasars and the developing inhomogeneous density field. Due to the high ionisation degree, the mean free path of UV photons in the Ly α forest is hundreds of Mpc long. Hence, many sources which are distributed over large distances contribute to the UV background (UVB), creating a quite homogeneous UVB.

This picture changes slightly in the vicinity of a QSO. There the QSO radiation dominates over the overall background, additionally increasing the ionisation state of the intergalactic gas. This manifests itself in a reduction of the number of observed absorption features towards the emission redshift of the QSO. This "proximity effect" was first observed in low resolution spectra by Carswell et al. (1982) and was later confirmed by Murdoch et al. (1986) and Tytler (1987). In a seminal paper, Bajtlik et al. (1988) used the proximity effect to estimate the intensity of the intergalactic UV radiation in comparison to the QSO luminosity, assuming that QSOs reside in random regions of the universe. Investigations of the proximity effect along these lines using absorption line counting statistics (Lu et al. 1991; Williger et al. 1994; Cristiani et al. 1995; Giallongo et al. 1996; Srianand & Khare 1996; Cooke et al. 1997; Scott et al. 2000) or using pixel statistics of the transmitted flux (Liske & Williger 2001; Dall'Aglio et al. 2008b,a, 2009; Calverley et al. 2010) reveal the proximity effect in many high-resolution QSO spectra. Recent results by Dall'Aglio et al. (2008a), Calverley et al. (2010), and Haardt & Madau (2011) indicate a steady drop in the ionising background radiation's photoionisation rate $\Gamma_{\text{UVB}} \sim 10^{-11.7} \text{ s}^{-1}$ at $z \sim 2$ to $\Gamma_{\text{UVB}} \sim 10^{-13.85} \text{ s}^{-1}$ for $z \gtrsim 5.5$. However results obtained from SDSS spectra indicate a constant ionising background level between $2.5 < z < 4.6$ (Dall'Aglio et al. 2009).

Independent measurements of the UV background radiation can be obtained with the flux decrement method, which requires cosmological simulations to model statistical properties of the Ly α forest (Rauch et al. 1997b; Theuns et al. 1998b; Songaila et al. 1999; McDonald & Miralda-Escudé 2001; Meiksin & White 2003; Tytler et al. 2004; Bolton et al. 2005; Kirkman et al. 2005; Jena et al. 2005; Faucher-Giguère et al. 2008a; Bolton & Haehnelt 2007). It becomes clear that measurements using the proximity effect seem to overestimate the UV background by a factor of a few. It is discussed whether this discrepancy arises from environmental effects such as clustering, gas infall, or the large scale density environment, which are neglected in the proximity effect modelling (Rollinde et al. 2005; Guimarães et al. 2007; Faucher-Giguère et al. 2008b; Partl et al. 2010). Such effects can result in an overestimation of the UV background by up to a factor of 3 (Loeb & Eisenstein 1995).

QSOs are thought to reside in very massive halos. Observational estimates on the QSO halo mass revealed values of a couple of $10^{12} M_{\odot}$ (da Ângela et al. 2008) up to $\sim 10^{14} M_{\odot}$ (Rollinde et al. 2005). From numerical simulations of structure formations it is known that such massive halos are not located at random positions in the universe, but form in dense environments where galaxies cluster. Observational determinations of the large scale environment around QSOs from Ly α forest spectra indicate that QSOs are embedded in large scale overdensities extending from proper $\sim 3 - 5$ Mpc (D'Odorico et al. 2008) to proper $\sim 10 - 15$ Mpc (Rollinde et al. 2005; Guimarães et al. 2007). Numerical simulations by Faucher-Giguère et al. (2008b) show large scale overdensities of proper $\sim 3 - 6$ Mpc for redshifts $4 > z > 2$, consistent with the results by D'Odorico

et al. (2008). Using detailed radiative transfer simulations of three QSOs residing in different cosmic environments Partl et al. (2010, hereafter P1) found indications that such large scale overdensities weaken the apparent proximity effect signal, resulting in an overestimation of the UV background. However, the low number of QSO hosts in the P1 study does not allow us to securely confirm the effect. We therefore extend this study using a large sample of different dark matter halos in various mass ranges and inquire how such large scale overdensities affect UV background measurements. It is also checked whether a possible dependence of the UV background with the host halo mass exists, as was suggested by Faucher-Giguère et al. (2008b).

This chapter is structured as follows. In Sect. 6.2 we present the dark-matter simulation used for this study and determine realistic models of the Ly α forest from a semi-analytical model of the intergalactic medium. In Sect. 6.3 we introduce the proximity effect as a measure of the intergalactic ionising background flux and characterise the halo sample used in this study. It is further discussed how Ly α forest mock spectra are generated. Subsequently in Sect. 6.4 we evaluate the effects of large scale overdensities and infall velocities on the proximity effect. A possible dependence of the UV background measurements on the halo mass and on the large scale mean density around the QSO are assessed. We summarise our results in Sect. 6.5.

6.2. Simulation

6.2.1. Distribution of baryons in the IGM

In order to obtain a model for the gas content in the IGM that is in agreement with observed properties of the Ly α forest, we employ the $64 h^{-1}$ Mpc² DM simulation of the CLUES project³ (Gottlöber et al. 2010) with 1024^3 DM particles. The GADGET2 (Springel 2005) simulation has a mass resolution of $m_{p,DM} = 1.86 \times 10^7 M_{\odot} h^{-1}$ and uses a WMAP5 (Hinshaw et al. 2009) cosmology.

The distribution of DM particles was obtained at four different redshifts $z = 4.8, 4, 3,$ and 2 . Using triangular shaped cloud (TSC) density assignment we obtain a 800^3 regularly spaced density and velocity grid from the DM particle distribution. Halos have been identified in the simulation using the hierarchical friends-of-friends (HFOF) algorithm (Klypin et al. 1999) with a linking length of 0.17.

To obtain an IGM gas density field from which we can construct H I Ly α forest spectra, it is assumed that the properties of the baryonic component, such as density and bulk velocity, are proportional to those of the dark matter (Petitjean et al. 1995; Meiksin & White 2001). We have checked this assumption with a comparable GADGET2 gas dynamical simulation using 2×256^3 particles with a size of $12.5 h^{-1}$ Mpc (Forero-Romero et al 2011, *in preparation*). This corresponds to a mass resolution of $m_{p,bar.} = 3.7 \times 10^5 M_{\odot} h^{-1}$ in gas. The gas dynamical SPH simulation includes also radiative and Compton cooling, star formation, and feedbacks through galactic winds using the model of Springel & Hernquist (2003). Furthermore, a UVB generated from QSOs and AGNs is included at $z < 6$ (Haardt & Madau 1996). The state of the simulation has been recorded

²Distances are given as comoving distances unless otherwise stated.

³<http://www.clues-project.org>

6. Large scale environmental bias on the LOS proximity effect.

at $z = 5$ and 4. The simulation was rerun with DM only using the same realisation of the initial conditions. We obtain density fields from the two simulations with equal spatial resolution as the larger $64 h^{-1}$ Mpc sized fields. The DM densities are then compared to the densities derived from the gas dynamical simulation and the resulting DM to gas density relation is shown in Fig. 6.1. The baryonic component follows the DM density closely, especially at $z = 5$. For lower redshifts a linear relation is still obtained, however the scatter around the line of equality, where the DM overdensity is equal to the gas overdensity, increases. A small bump in the DM to gas density relation towards higher gas densities develops with decreasing redshift at gas densities where a fraction of the gas becomes shock heated. The density in these shocks is slightly larger than the underlying DM, giving rise to the small bump in the DM to gas density relation. Studying the density probability distribution reveals that at $z = 5$ the two components are similar (see Fig. 6.1). At lower redshifts, the baryon distribution is slightly shifted to higher densities and decreases steeper than the DM one at high densities. For completeness we additionally show the effective equation of state derived from the density and temperature fields in Fig. 6.1.

From the gas dynamical simulation we find the assumption that baryons follow the DM distribution to be reasonable for the low density regions from which the Ly α forest arises. We therefore use the DM only simulations to derive the gas density and velocity fields as described in P1. By using a TSC mass assignment scheme, we implicitly smooth our density field on sizes of 1.5 cells (i.e. $120 h^{-1}$ kpc). This is comparable to a constant Jeans length smoothing of $\lambda_{\text{Jeans}} \propto \delta^{-1/2}(1+z)^{-1/2} \sim 150 h^{-1}$ kpc at $z = 3$, where δ is the density contrast.

6.2.2. Model and calibration of the intergalactic medium

The characteristics of the IGM gas are determined analogously to the method described in P1. We will therefore only briefly sketch the method with which a representative IGM model is obtained.

The temperature in the IGM can be approximated by the so called effective equation of state (Hui & Gnedin 1997), which is expressed as $T = T_0(1 + \delta)^{\gamma-1}$. The effective equation of state shows a linear trend for overdensities $0.1 \lesssim \delta \lesssim 10$ (compare Fig. 6.1). For higher densities its behaviour is approximated by assuming a constant value above a temperature cut-off $T_{\text{cut-off}} = T(\delta = 10)$. By employing the method developed in Hui et al. (1997), we compute HI Ly α absorption spectra from the DM density and velocity fields, once the parameters T_0 , γ , and the UV background photoionisation rate (Γ_{UVB}) are determined. These are constrained by matching statistical properties of the simulated HI absorption spectra with observed relations. Such a calibration will be important in obtaining realistic representations of the IGM that closely mimic the observed characteristics.

In order to calibrate our IGM model, we employ four observational constraints sorted by increasing importance: (i) The observed equation of state (Ricotti et al. 2000; Schaye et al. 2000b; Lidz et al. 2010), (ii) the evolution of the UV background photoionisation rate (Haardt & Madau 2001; Bianchi et al. 2001; Bolton et al. 2005; Dall’Aglio et al. 2008b, 2009), (iii) the observed evolution of the effective optical depth in the Ly α forest (Schaye et al. 2003; Kim et al. 2007), and (iv) the transmitted flux probability distribution

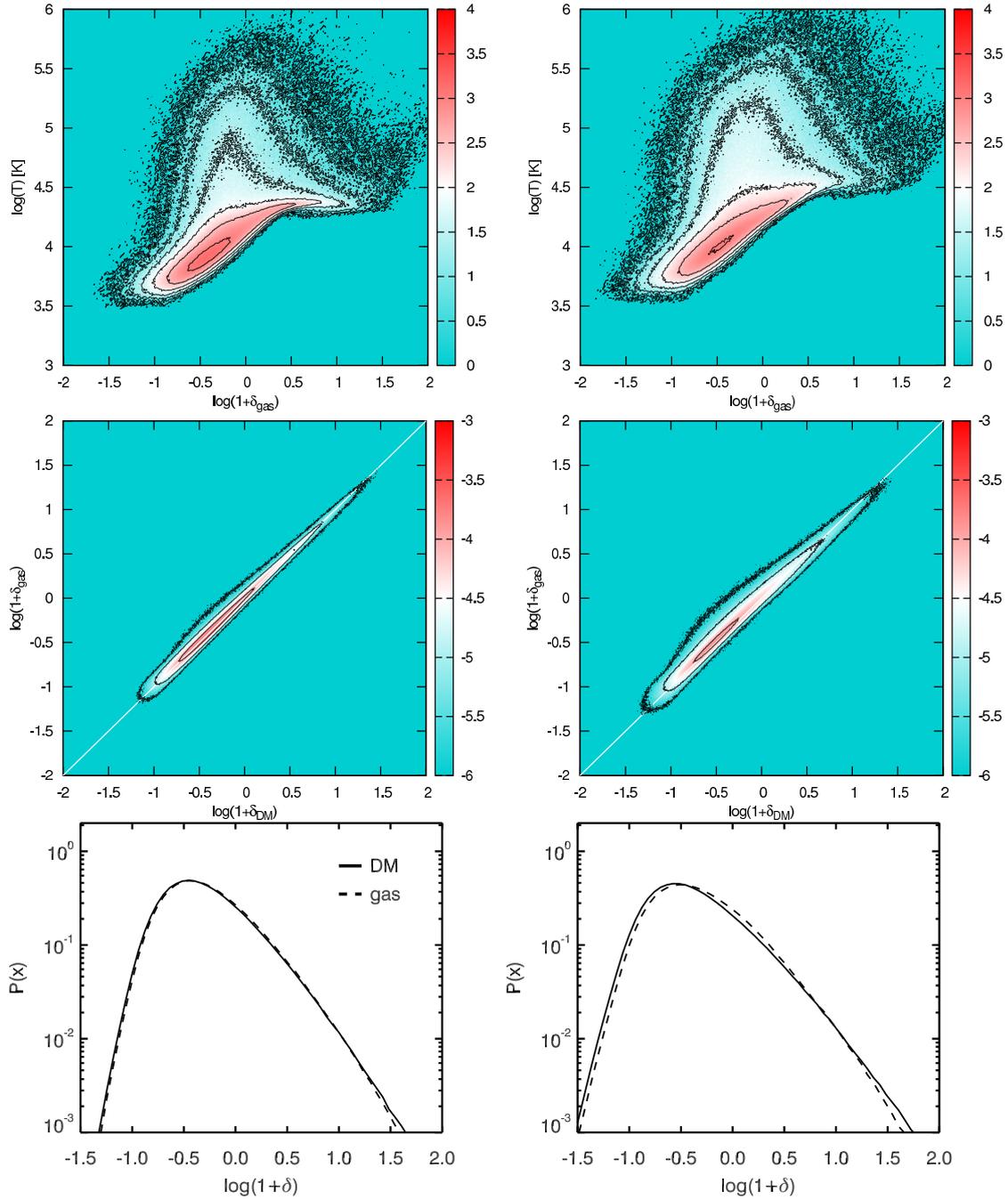


Figure 6.1.: The top row shows the effective equation of state of a $12.5 h^{-1}$ Mpc gas dynamic simulation binned to cells of $80 h^{-1}$ kpc at redshift $z = 5$ (left panels) and $z = 4$ (right panels). The middle row shows the gas overdensity $1 + \delta_{\text{gas}}$ as a function of dark matter density $1 + \delta_{\text{DM}}$. The white diagonal line marks the linear relation if gas strictly follows the dark matter. In both rows the colour coding refers to the logarithm of the respective number of cells. The lowest row shows the probability distribution of the dark matter (solid line) and gas (dashed line) overdensity.

6. Large scale environmental bias on the LOS proximity effect.

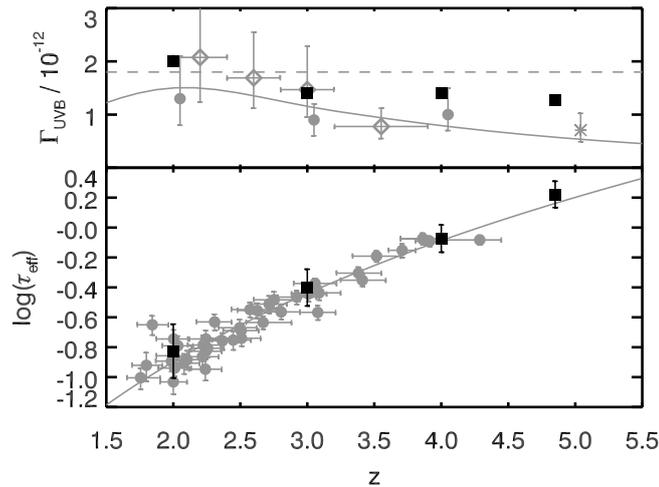


Figure 6.2.: *Upper panel:* The evolution of the UV background photoionisation rate in the four snapshots (black squares) compared to Bolton et al. (2005) (grey points shifted by $\Delta z = 0.05$ for better visibility), measurements by Dall’Aglio et al. (2008a) (grey open diamonds) and Calverley et al. (2010) (grey star), a fit to SDSS data measurements by Dall’Aglio et al. (2009) (grey dashed line) and predictions by Haardt & Madau (2001) (grey line). *Lower panel:* The effective optical depth of our models (black points) in comparison to measurements by Schaye et al. (2003) (grey points). The continuous line shows the fit to observational data by Kim et al. (2007).

(FPD, Becker et al. 2007).

The statistical quantities of the simulated spectra are derived from the simulation using 500 lines of sight randomly drawn through the cosmological box. Due to the availability of high resolution spectra with high signal-to-noise (S/N) ratios of up to 120, we approximate such high quality data by assuming noise free spectra. It has been noted by Calverley et al. (2010) that low S/N levels introduce a systematic shift in the proximity effect signal. Since we want to quantify the physical effect of the halo’s surrounding environment on UVB measurements, adding noise would only lead to degeneracies between the two effects. However we will briefly address the influence of noise in Section 6.4.2.

The spectra are convolved with the instrument profile of the UVES spectrograph and are then binned to the typical resolution of UVES spectra of 6.7 km s^{-1} . We further assume that the QSO continuum can be perfectly determined and we therefore consider no uncertainties in the QSO continuum. This procedure is adopted to precisely quantify physical effects without contamination of the signal with observational uncertainties. These will only increase the variance in the results discussed below.

To determine the model parameters for the IGM, we adopt an iterative method based on χ^2 minimisation between simulated and observed FPD, as the latter provides one of the strongest constrains on the Ly α forest properties. Our steps are as follows:

1. We choose a initial guess $(T_0, \gamma, \Gamma_{\text{UVB}})_0$ according to the measurements of the equation of state by Schaye et al. (2000b) and Lidz et al. (2010) and the evolution of the UVB by Haardt & Madau (2001) (see Figs. 6.2 and 6.3).

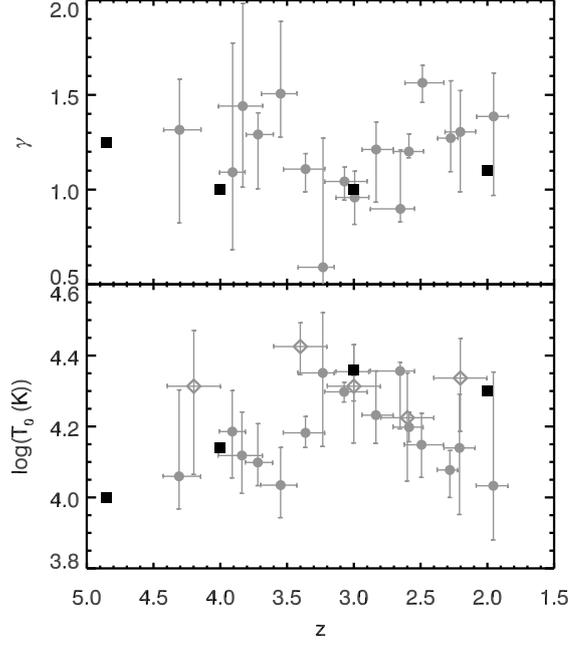


Figure 6.3.: *Upper panel:* Comparison of our choices for γ (black squares) with observationally derived results by Schaye et al. (2000b) (grey points). *Lower panel:* Comparison of our model T_0 (black points) with observations by Schaye et al. (2000b) (grey points) and Lidz et al. (2010) (grey open diamonds).

2. With the above initial guess and using 500 lines of sight, we determine the average effective optical depth $\tau_{\text{eff}}(z) = -\ln \langle F(z) \rangle$ with F being the transmitted flux and the averaging is performed over the whole line of sight. As T_0 and γ have a subdominant effect on $\tau_{\text{eff}}(z)$, we first tune Γ_{UVB} to obtain a match in the effective optical depths of the model spectra with recent observations from Kim et al. (2007). This results in a new set of parameters $(T_0, \gamma, \Gamma_{\text{UVB}})_1$.
3. Finally we construct our simulated FPD. As observational constraint we employ the log normal fits obtained by Becker et al. (2007) in a redshift interval of $\Delta z = \pm 0.25$ centred on the snapshot redshift z (see Fig. 6.4). By using their fits to the FPD, our model spectra can be compared with the observations without considering the effect of detector noise. However for consistency we also compare the simulated FPD with the raw combined observational results by Becker et al. (2007), folding the simulation result with the global noise function of the combined observed sample (see Fig. 6.5). The best fit parameters $(T_0, \gamma, \Gamma_{\text{UVB}})_*$ are then iteratively determined with a Simplex optimisation procedure.

Our best fit parameters are presented in Table 6.1, and plotted in Figs. 6.2, 6.3, 6.4, and 6.5 in comparison with different literature results.

The inferred evolution of the UV background in Fig. 6.2 closely follows recent results by Haardt & Madau (2001), Bolton et al. (2005), and Dall’Aglia et al. (2008b). The values for redshift $z = 4.8$ is rather high when compared with Calverley et al. (2010), however it is still in agreement within the 2σ limits. Furthermore the effective optical

6. Large scale environmental bias on the LOS proximity effect.

Table 6.1.: Model parameters of the semi-analytical model.

z	$\log(T_0[K])$	γ	$\Gamma_{\text{UVB}}[\text{s}^{-1}]$	$\log \tau_{\text{eff}}$
4.8	4.00	1.25	1.16×10^{-12}	0.222
4.0	4.14	1.00	1.4×10^{-12}	-0.074
3.0	4.36	1.00	1.4×10^{-12}	-0.401
2.0	4.30	1.10	2.0×10^{-12}	-0.826

depth, also shown in Fig. 6.2, is consistent with high resolution observations by Schaye et al. (2003) and Kim et al. (2007). Additionally the final parameters of the equation of state T_0 and γ lie within the measurement uncertainties of Schaye et al. (2000b) and Lidz et al. (2010), which is evident from Fig. 6.3.

The flux probability distributions estimated from the simulated sight lines agree reasonably well with the noise-corrected observed profiles estimated by Becker et al. (2007) who assumed a log-normal distribution of the optical depth in the Ly α forest (see Fig. 6.4). At redshifts $z = 2$ and 3 the match between the mean observed profiles and our simulations is good. However at higher redshifts, the agreement marginally decreases, even though the distributions are consistent within the variation between different lines of sight. This discrepancy manifests itself clearer when comparing the simulations with the raw observational data in Fig. 6.5 which include the effects of noise. The discrepancy is especially strong at the high transmission end of the distribution. This is most certainly caused by our crude noise modelling of a sample with inhomogeneous S/N, which we obtained by stacking the noise functions of the various lines of sight. On the other hand we assume a perfect knowledge of the continuum level in the spectra, which is a challenge to determine in observed spectra, especially at high redshifts.

6.3. Method

6.3.1. Line-of-sight proximity effect

Bright UV sources, such as QSOs, alter the ionisation state in their vicinity strongly up to proper distances of a couple of Mpc. The resulting change in the ionisation state directly translates into a decrease in optical depth. This decrease manifests itself in a reduction of the absorption line density in the Ly α forest when approaching the redshift of the source. Using line-counting statistics, Bajtlik et al. (1988) measured the line-of-sight proximity effect in QSO spectra for the first time. They assumed that QSOs are situated in a mean IGM environment, neglecting any redshift distortions from peculiar velocities of the absorption clouds. They further assumed that the QSO's radiation field is radially decreasing from the source proportional to r^{-2} , i.e. geometric dilution. Using detailed radiative transfer simulations, P1 showed that radiative transfer effects only play a marginal role in the line-of-sight proximity effect and that the assumption of geometric dilution holds.

Assuming photoionisation equilibrium, the change in the optical depth as a function of

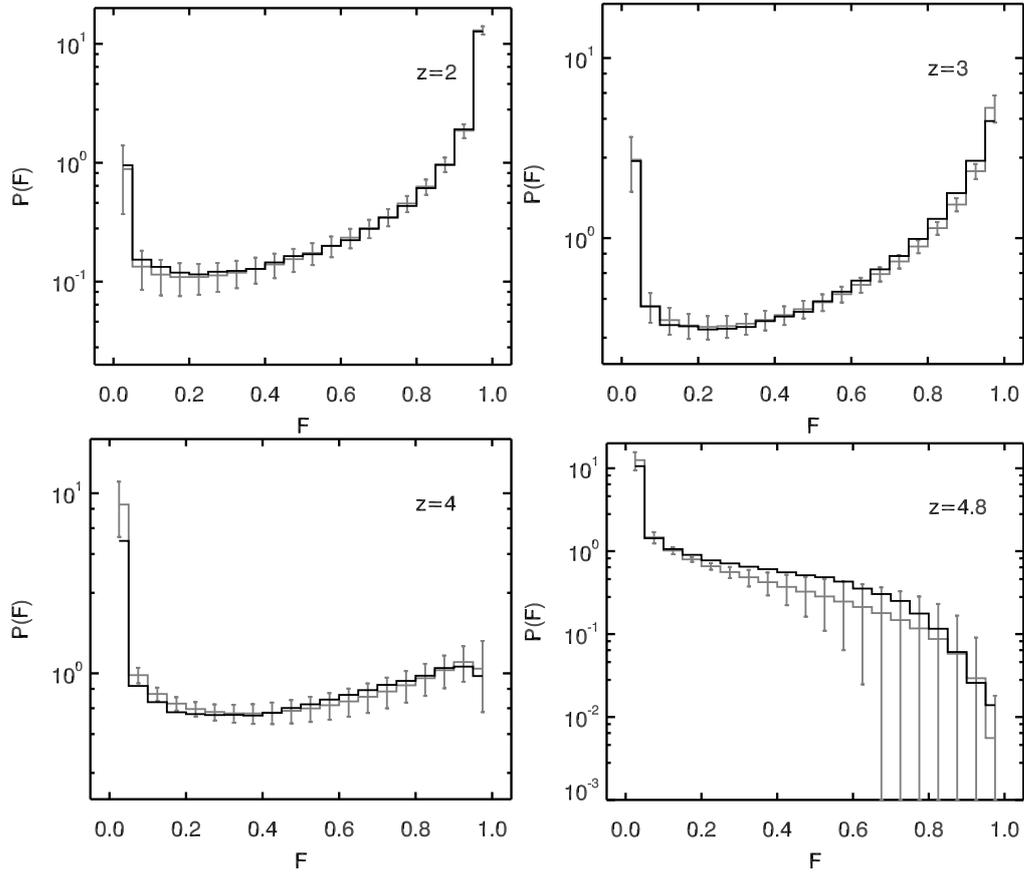


Figure 6.4.: Probability distribution function of transmitted flux F at redshifts 2, 3, 4, and 4.8. Grey lines give the mean observed PDF obtained from the fits excluding noise given in (Becker et al. 2007) with 1σ error bars denoting the variation between different lines of sight. The mean is derived in redshifts bins of ± 0.25 centred on the model redshift. The black full line shows the results from our models.

6. Large scale environmental bias on the LOS proximity effect.

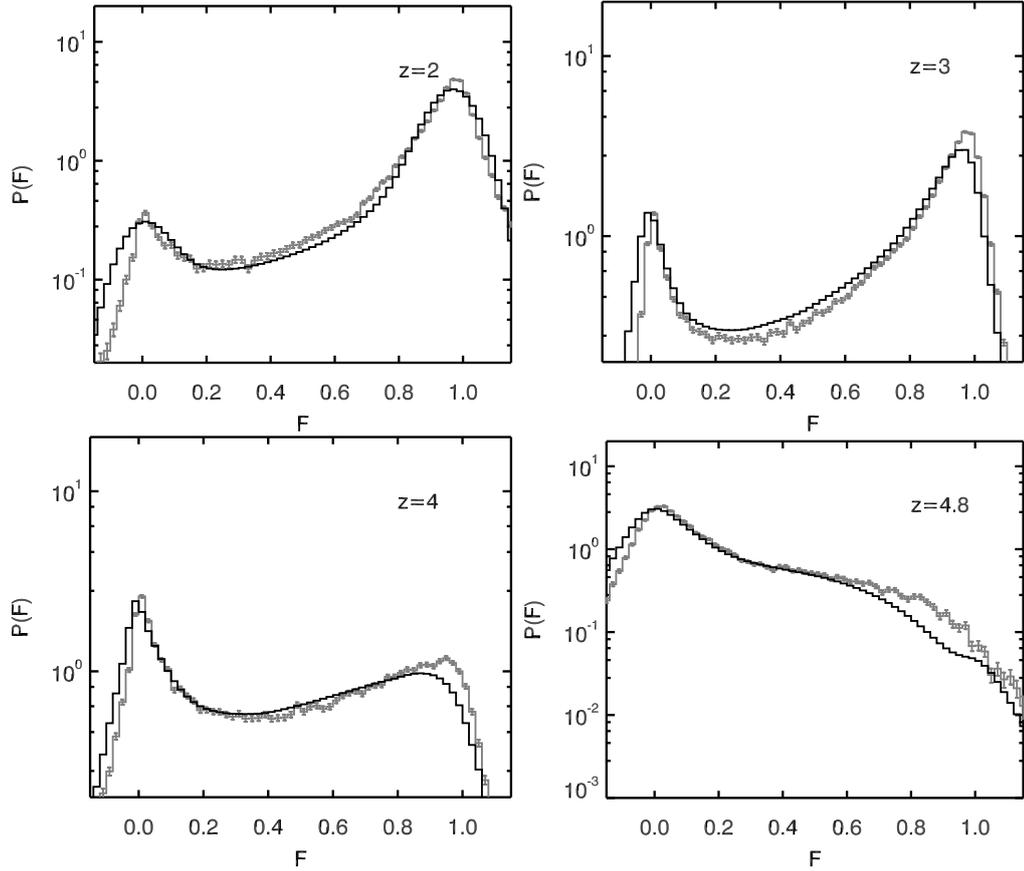


Figure 6.5.: Probability distribution function of transmitted flux F at redshifts 2, 3, 4, and 4.8. Grey lines with 1σ error bars give the observed PDF (Becker et al. 2007) and the black full line show the results obtained from our models folded with the signal to noise function of the combined observations. The combined observational data is derived in redshifts bins of ± 0.25 centred on the model redshift.

distance r to the QSO can be expressed as

$$\tau_{\text{PE}}(r) = \tau_{\text{Ly}\alpha}(r) (1 + \omega(r))^{-1} \quad (6.1)$$

(Liske & Williger 2001), where τ_{PE} is the observed optical depth, $\tau_{\text{Ly}\alpha}$ is the optical depth in the absence of a QSO, and

$$\omega(r) = \frac{\Gamma_{\text{QSO}}(r)}{\Gamma_{\text{UVB}}} = \frac{1}{16\pi^2} \frac{1}{r^2} \frac{L_{\nu_{\text{LL}}}}{J_{\nu_{\text{LL}},\text{UVB}}} \frac{3 - \alpha_{\text{b}}}{3 - \alpha_{\text{q}}} \quad (6.2)$$

acts as a normalised distance to the QSO (Dall’Aglio et al. 2008b). Here $\Gamma_{\text{QSO}}(r)$ and Γ_{UVB} are the photoionisation rates of the QSO and the UV background respectively, $L_{\nu_{\text{LL}}}$ is the QSO’s luminosity at the Lyman limit, while $J_{\nu_{\text{LL}},\text{UVB}}$ is the UVB’s Lyman limit flux, α_{b} is the slope of the UVB’s spectral energy distribution $\propto \nu^{\alpha_{\text{b}}}$, and α_{q} denotes the spectral slope of the QSO’s emission $\propto \nu^{\alpha_{\text{q}}}$. In this work we assume for simplicity the UVB to be dominated by QSOs. We therefore follow Haardt & Madau (1996) and chose the slope of the UVB to be $\alpha_{\text{b}} = -1.5$. This is consistent with observations by Telfer et al. (2002) who find $\alpha_{\text{q}} = -1.57 \pm 0.17$. The effect of differing spectral shapes between the UVB and the QSO has been studied in Dall’Aglio et al. (2008b) and P1. The proximity effect is introduced into lines of sight drawn from our simulation boxes by modifying the neutral hydrogen fraction in real space using $n_{\text{HI,PE}}(r) = n_{\text{HI,Ly}\alpha}(r) (1 + \omega(r))^{-1}$.

6.3.2. Environments

To study the impact of the large scale environment on the proximity effect measurement as a function of QSO host halo mass, we pick halos in given mass ranges from the simulation to serve as QSO hosts. Around each halo, 100 lines of sight with a length of two comoving box sizes ($128 h^{-1}$ Mpc) are randomly drawn from the box assuming periodic boundary conditions. Our sample of QSO host halos cover a wide range of masses. At each redshift we use the 20 most massive halos, 50 halos with a mass around $10^{12} M_{\odot}$, and 50 halos with a mass around $10^{11} M_{\odot}$. For the 10^{12} and $10^{11} M_{\odot}$ mass objects we choose 50 halos from the mass sorted halos with a mass larger than the cutoff. For the more massive bin, less than 50 halos with mass higher than the cutoff are present in the simulation at $z \geq 4$. In this case we choose all the halos above $10^{12} M_{\odot}$ and extend the range to lower masses until the sample consists of 50 halos. For these redshifts, the $10^{12} M_{\odot}$ mass range overlaps with the 20 most massive halos. The covered mass intervals are given in Table 6.2 as a function of redshift. We further note that the halos in the $10^{11} M_{\odot}$ mass bin are certainly not massive enough to host QSOs, since da Ângela et al. (2008) for instance estimated a QSO host halo mass of around $3 \times 10^{12} M_{\odot}$ independent of redshift and luminosity from the 2dF-SDSS survey. However to establish any dependency of the proximity effect signal with the host’s mass, they are included for the purpose of an extreme low mass range.

In order to test the influence of the large scale environment around QSO host halos we additionally construct a sample of lines of sight for each redshift where the origins are randomly selected. This sample provides a null hypothesis since any influence of large scale density fluctuations averages out if the origins of the lines of sight are randomly distributed in the box.

6. Large scale environmental bias on the LOS proximity effect.

Table 6.2.: Mass range of the halo samples as a function of redshift. The $10^{12} M_{\odot}$ and $10^{11} M_{\odot}$ samples contain 50 halos each. The 20 most massive halos are given in units of $10^{12} M_{\odot}$, the other mass bins in units corresponding to the mass bin.

z	20 most massive	$10^{12} M_{\odot}$	$10^{11} M_{\odot}$
4.8	2.07 – 0.92	2.07 – 0.53	1.04 – 1.00
4	4.84 – 1.51	4.84 – 0.88	1.03 – 1.00
3	15.8 – 3.41	1.38 – 1.00	1.01 – 1.00
2	40.2 – 6.19	1.14 – 1.00	1.01 – 1.00

6.3.3. Measuring the proximity effect in spectra

The proximity effect signature is measured in Ly α forest spectra by first constructing an appropriate ω -scale for the observed QSO. In this study we assume the QSO to have a Lyman limit luminosity of $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$, and the photoionisation rates of our model UVB field are used. We have shown in P1 that for higher QSO luminosities the proximity effect is more pronounced and less affected by the density distribution. Therefore we use a low luminosity QSO to obtain upper limits of the environmental bias. Given the ω -scale, the transmission spectra is then evaluated for each line of sight in bins of $\Delta \log \omega$ (Dall’Aglio et al. 2008b). Using the mean transmission per bin, the effective optical depth in the bin $\tau_{\text{eff,QSO}}(\Delta \log \omega)$ is calculated and normalised to the effective optical depth in the Ly α forest unaffected by the QSO’s radiation $\tau_{\text{eff,Ly}\alpha}$. The normalised optical depth ξ is thus

$$\xi(\Delta \log \omega) = \frac{\tau_{\text{eff,QSO}}(\Delta \log \omega)}{\tau_{\text{eff,Ly}\alpha}}. \quad (6.3)$$

The imprint of the proximity effect onto the normalised optical depth for a given ω -scale becomes

$$\xi(\omega) = (1 + \omega(r))^{1-\beta} \quad (6.4)$$

(Liske & Williger 2001) where β is the slope of the Ly α absorber’s column density distribution. Throughout this work we assume $\beta = 1.5$ (Kim et al. 2001). The UVB photoionisation rate Γ_{UVB} can then be determined using the proximity effect strength parameter a

$$\xi = \left(1 + \frac{\omega}{a}\right)^{1-\beta}. \quad (6.5)$$

This parametrisation was introduced by (Dall’Aglio et al. 2008b,a) in analyses of observed spectra with an assumed UVB as reference. Values of $a > 1$ or $a < 1$ indicate a weaker or stronger proximity effect than the model, respectively. The measured photoionisation rate of the UVB is then determined using the reference value $\Gamma_{\text{UVB,ref}}$ multiplied by the strength parameter a . In our case the strength parameter a indicates any deviation of the measured UVB photoionisation rate from the input value. The strength parameter is determined by fitting Eq. 6.5 to the binned normalised optical depth $\xi(\Delta \log \omega)$. In order to exclude any direct impact of the host halo on the strength parameter fit, only data with $\log \omega < 2$ have been used.

In P1 we found the observed ξ to fluctuate strongly around the analytical proximity

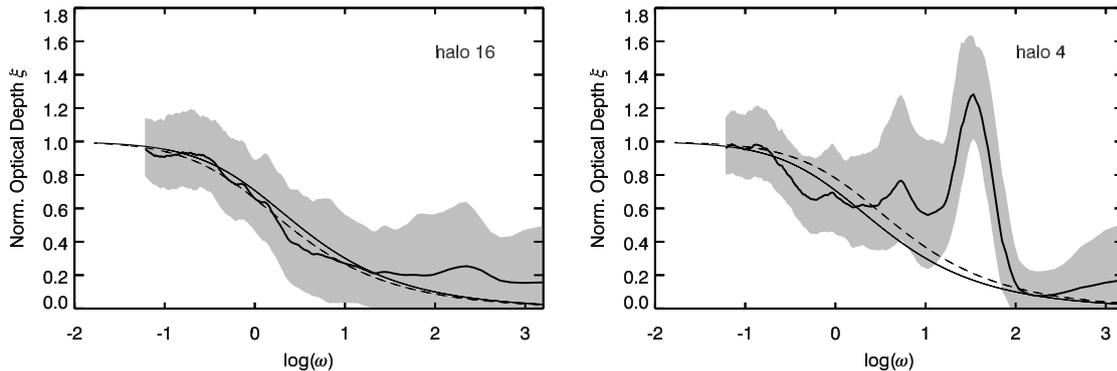


Figure 6.6.: Mean normalised optical depth profiles obtained using a moving average smoothing with a kernel size of $\Delta \log(\omega) = 0.5$ for two halos taken from the 20 most massive halo sample at redshift $z = 4$ (black solid line with grey shading). The mean halo profiles are determined using 100 lines of sight and the shaded area gives the 1σ standard deviation. The analytical proximity effect model is given as the smooth black line, while the fitted profile is indicated by the dashed line. The left panel shows a halo without strong intervening density features and closely following the analytic model. The right panel illustrates a halo having strong density features in its vicinity, causing strong deviations from the analytical form.

effect profile. Any fit of Eq. 6.5 is thus biased by these large fluctuations which arise from the presence of strong absorbers along the line of sight.

In order to obtain a smoother ξ -profile, the wavelength scale of the spectrum is transformed into the $\log \omega$ -scale and a boxcar smoothing (also known as moving average) with the size of $\Delta \log \omega$ is applied to the transmission spectrum. Analogously to the method given above, we then calculate the effective optical depth in each smoothed pixel and determine the normalised optical depth ξ . This results in smooth ξ -profiles allowing the identification of areas dominated by strong absorption systems.

In Fig. 6.6 two examples of mean ξ -profiles taken out of the 20 most massive halos sample at $z = 4$ are given. A $\Delta \log \omega = 0.5$ is found to yield a good balance between smoothing and retaining structure in the ξ -profile. For the mean ξ -profiles shown in Fig. 6.6, 100 lines of sight have been constructed randomly around the DM halos. The two examples illustrate a halo without any signs of intervening strong absorption systems (left panel) and one where a strong density feature is located near the host halo at $\log \omega = 1.5$ (right panel). From the halo without strong nearby systems it becomes evident that the profile closely follows the analytical one. A fit to this smooth ξ -profile is now very robust, since the functional form of the profile is well constrained and the fit of Eq. 6.5 is not solely determined by a small number of data points. The same is true for cases with strong intervening absorption systems, as long as the underlying smooth profile of Eq. 6.4 is still visible. Strong deviations from the analytical form however, as presented in our second example, are large enough to prevent a clear identification of the smooth analytical profile and the obtained strength parameters have to be treated with caution.

6.4. Results

6.4.1. Null hypothesis: Random locations

In order to assess whether large scale over-densities affect the proximity effect profile and the associated strength parameter, we first establish results that are unaffected by such large scale density features. A sample of 500 randomly selected lines of sight originating at random points in the simulation will serve as a null hypothesis. No noise is added to the spectra. However we will use this sample in the next section to discuss the effect of detector noise on our results.

In Fig. 6.7 we discuss results obtained for redshifts $z = 4.8$, and 3. For each line of sight, the normalised optical depth ξ was calculated using the method described in Sect. 6.3.3. For each redshift we determine the mean and median ξ profile, as well as the ξ probability distribution as a function of ω . The ξ probability distribution is shown in colour coding in Fig. 6.7.

At $z = 4.8$ the mean profile follows the analytic proximity effect model very well, with just a slight increase in its slope towards higher ω values. Note however that at large ω values of $\log \omega > 2$, ξ can only be poorly determined due to a very small number of pixels contributing to the $\Delta\omega$ bins. The median profile as well follows the input model up to $\log \omega \sim 0.5$. However at $\log \omega > 0.5$ the median profile steepens strongly and starts to deviate from the input model and the mean profile. This indicates a growing asymmetry in the ξ distribution with increasing ω (approaching the QSO), resulting in the growing discrepancy between the mean and median profiles. Considering the ξ distribution function, the increasing skewness of the distribution becomes apparent. For $\log \omega < 1$, the width of the distribution stays constant and appears symmetrical and normally distributed in the logarithmic scale. This indicates a log normal distribution of ξ values. However at larger ω values the distribution starts to widen up and the peak in probability shifts towards lower values of $\log \xi$, moving away from the expectations of the analytical formalism.

For redshifts $z = 4$ and $z = 3$ a similar picture emerges. The ξ distribution widens as a function of redshift, with an increase in its variance with decreasing redshift. The mean profile however always regains the input model well. However the discrepancy between the median profile and the mean at high ω values increases with decreasing redshift. The median profile becomes steeper and steeper, indicating a growing skewness of the ξ distribution at high ω . However at low ω , the ξ distribution stays symmetric and continues resembling a log normal distribution up to $\log \omega \sim 0$. At $z = 2$, the ξ -distribution shows a strongly increasing variance, dominating over the signal of the input model completely. We will therefore not consider results from $z = 2$ in this work.

The increase in variance is dominated by two factors. On the one hand the universe evolves and the growth of structure increases with decreasing redshift. This introduces stronger density contrasts between under- and overdense regions. On the other hand, the IGM becomes more transparent with decreasing redshift due to the increasing UV background and cosmic expansion reducing the mean density of the universe. Therefore at low redshifts, the Ly α forest traces denser structures than at high redshift.

From the simple test of this subsection we conclude that with randomly selected lines of sight of random origins, the input proximity effect model can be regained with the

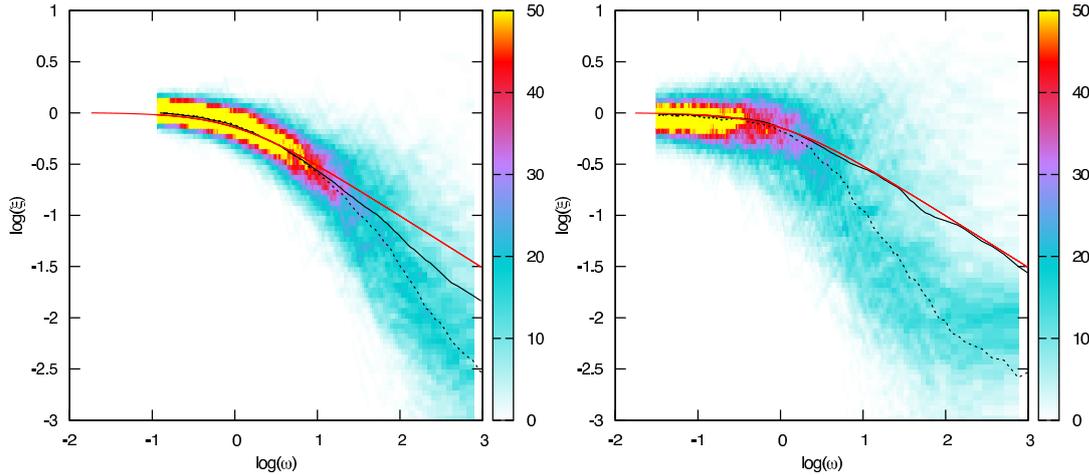


Figure 6.7.: Frequency of the normalised optical depth ξ using $\Delta \log \omega = 0.5$ calculated from 500 random lines of sight drawn from our simulation box, as a function of ω . The origins of these lines of sight were chosen not to be centred on any specific haloes, but on random points in the box. The left panel gives results for $z = 4.8$ and the right panel $z = 3$. The dashed line marks the input model used for generating the spectra including the proximity effect. The black solid line marks the mean ξ -profile, whereas the dotted black line marks the median profile.

mean ξ profile. In our sample this is valid for $z \geq 3$. The median profile however deviates more and more from the mean profile with decreasing redshift and cannot be used to measure the UV background from the proximity effect with the current analytical model, which is formulated for a mean IGM.

6.4.2. Dependence on the spectra signal to noise

We now want to discuss the effect of detector noise on the mean proximity effect profiles. Using the 500 lines of sight from the null hypothesis sample, noise is added to each spectrum using a signal to noise (S/N) of 10, 20, 50, 100, and 150. For each S/N sample we determine the mean ξ profile and normalise it to the S/N= 150 results. The influence of noise on the proximity effect signature is shown in Fig. 6.8, where the normalised ξ profiles are shown for $z = 4.8$ and $z = 3$ as a function of the signal to noise.

The largest influence noise has on the profile is at high ω values where only a small number of pixels contribute to ξ . However for ω values below $\log \omega \sim 1.5$, the strong influence of the noise on the profile disappears. At $z = 4.8$ an increase of noise effects can be noted at $\log \omega < 0$. There the results with a S/N of 10 produce deviations of up to 5% from high signal to noise spectra. However already with a signal to noise of 50, the ξ profile is regained with an accuracy of less than 1%. At $z = 3$ convergence with the low noise results for $\log \omega < 1.5$ is already achieved with a signal to noise of 20. In order to resolve the profile within 1% accuracy at higher ω values, a signal to noise of 50 is needed.

One has to keep in mind though that these results only apply to the combined sample and not to individual lines of sight. However this shows that with our approach of not considering noise in the spectra, high signal to noise results are very well approximated.

6. Large scale environmental bias on the LOS proximity effect.

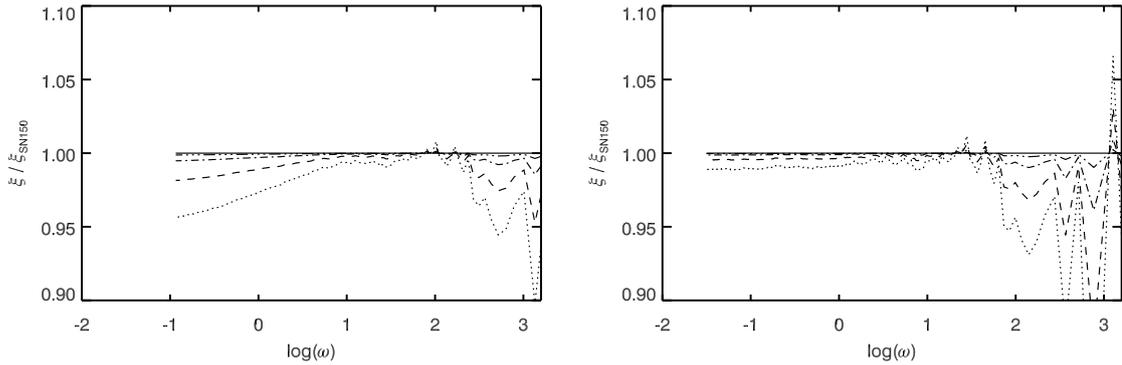


Figure 6.8.: The influence of the signal to noise ratio on the mean normalised optical depth ξ for the null hypothesis. Shown are results at redshifts $z = 4.8$ (left panel) and $z = 3$ (right panel). The ξ -profiles have been normalised to results using a signal to noise of 150. The dotted line denotes a S/N of 10, the dashed line a S/N of 20, the dash dotted line a S/N of 50, and the dash triple dotted line a S/N of 100.

6.4.3. Large scale environment

We now want to characterise the large scale density and velocity environment around our QSO host halos. It has been shown by Prada et al. (2006) and Faucher-Giguère et al. (2008b) that the mean density profile around massive halos does not reach the mean cosmic density before a halocentric distance of comoving $r_c \sim 10 h^{-1}$ Mpc. For smaller radii the mean DM density around halos can be up to 10 times above the mean density for halocentric distances of comoving $r_c \sim 1 h^{-1}$ Mpc. For even smaller radii the density profile is governed by the density profile of the host halo. A similar behaviour has been observed by P1. Faucher-Giguère et al. (2008b) have further determined the bias caused by Doppler shifting of absorption lines due to infalling gas in proximity effect measurements. The additional bias is found to be similarly important to the effect of overdensities. This effect is naturally included in our synthetic spectra taken from the simulations.

These large scale features are caused by large scale modes perturbing the density distribution at scales of several Mpc. These large modes still experience linear growth and increase with decreasing redshifts. Since halos with very large masses form preferably where these large modes show a positive amplitude, an increase in density above the mean cosmic density is expected around very massive halos. Since low mass halos also form in regions where there is less power on large scales, the density profiles are expected to be less influenced.

In Fig. 6.9 the smoothed and radially averaged overdensity profiles of our halo sample are shown as a function of distance and halo mass. A consistent picture to previous findings emerges. At $z = 4.8$ the profiles show a steep decline in density up to comoving distances of $r_c \sim 0.4 h^{-1}$ Mpc for the massive halos. This region directly shows the density profile of the host halo. For larger distances the slope with which the density decreases becomes shallower and a transition between the host halo profile and its surrounding density environment is seen. The profile then reaches the mean cosmic

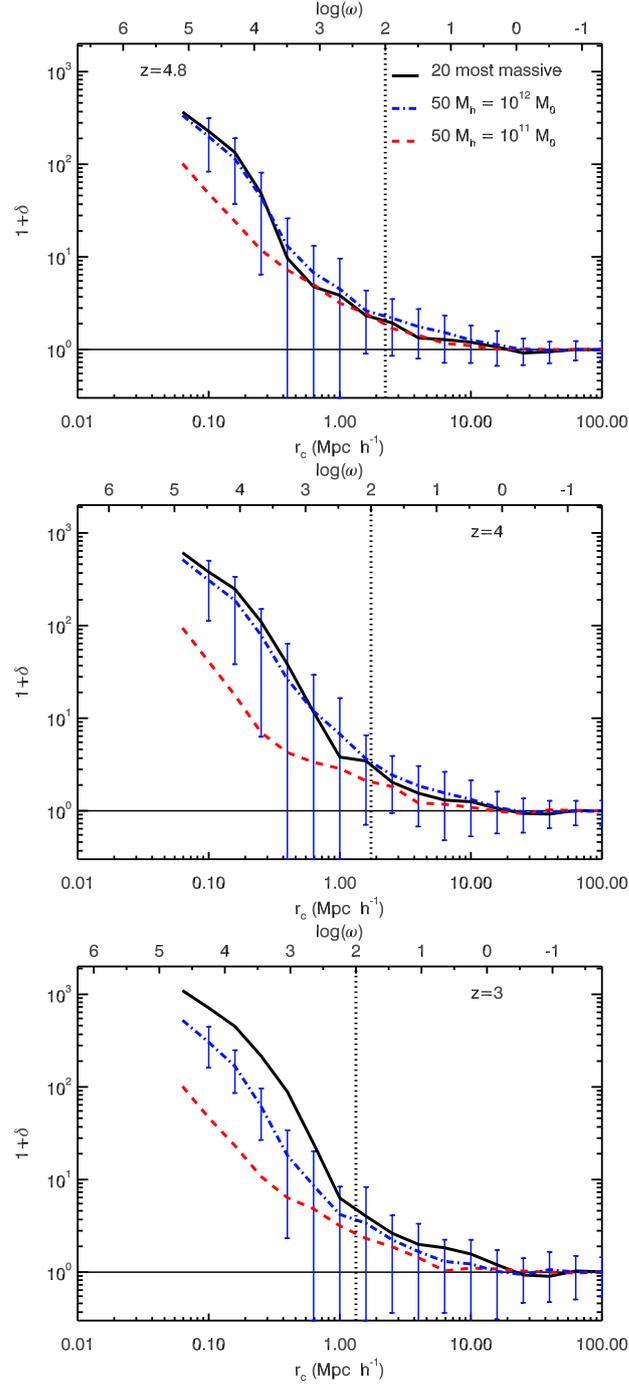


Figure 6.9.: Mean density environment around all the halos in each mass bin. The density profiles are smoothed for better visibility. Further the corresponding ω scale is given as reference. The black solid line denotes the mean profile around the 20 most massive halos, the blue dash-dotted line denotes the mean around the $M_{\text{halo}} = 10^{12} M_{\odot}$ halos, and the red dashed line marks the mean around the $M_{\text{halo}} = 10^{11} M_{\odot}$ halos. The standard deviation is given for the $M_{\text{halo}} = 10^{12} M_{\odot}$ halos exemplary for the other environments. The first panel shows $z = 4.8$, the second panel $z = 4$, and the last panel $z = 3$. The vertical dotted line marks the area of $\log \omega > 2$ which was excluded in the determination of the strength parameter.

6. Large scale environmental bias on the LOS proximity effect.

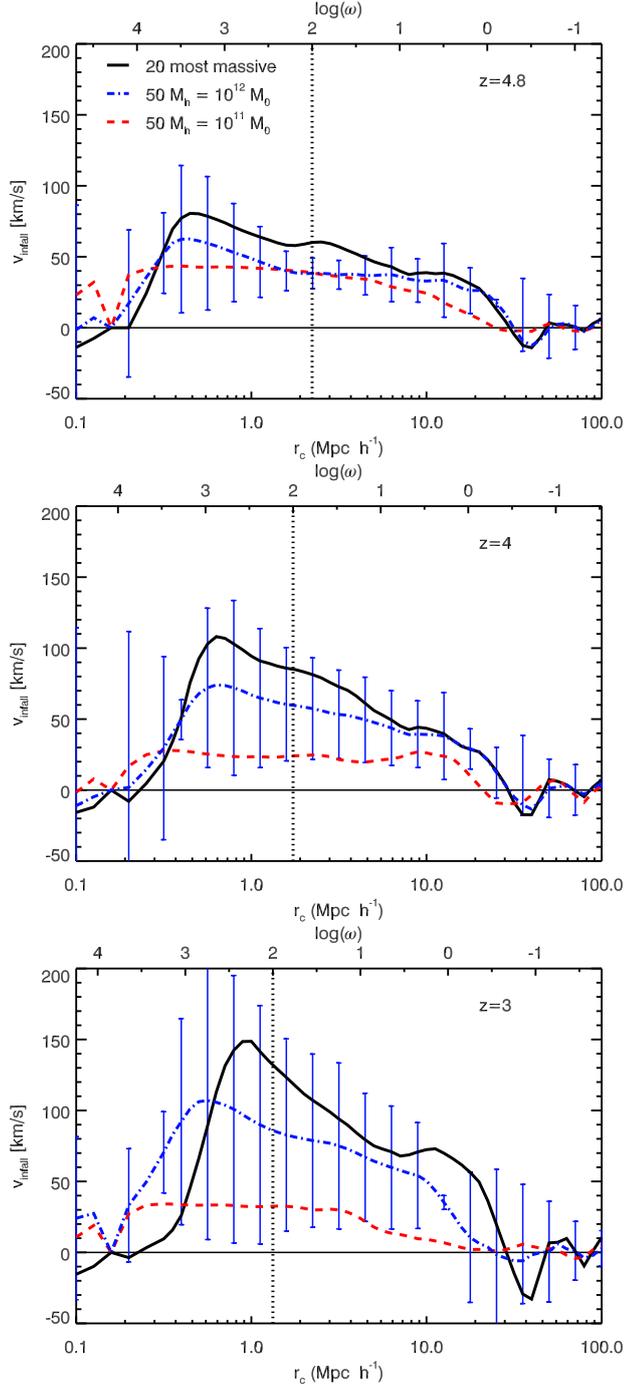


Figure 6.10.: Mean infall velocity around all the halos in each mass bin. The infall velocity profiles are smoothed for better visibility. Further the corresponding ω scale is given as reference. The black solid line denotes the mean infall velocity around the 20 most massive halos, the blue dash-dotted line denotes the mean around the $M_{\text{halo}} = 10^{12} M_{\odot}$ halos, and the red dashed line marks the mean around the $M_{\text{halo}} = 10^{11} M_{\odot}$ halos. The standard deviation is given for the $M_{\text{halo}} = 10^{12} M_{\odot}$ halos exemplary for the other environments. The first panel shows $z = 4.8$, the second panel $z = 4$, and the last panel $z = 3$. The vertical dotted line marks the area of $\log \omega > 2$ which was excluded in the determination of the strength parameter.

density at a distance above $r_c > 10 h^{-1}$ Mpc, independent of halo mass.

For $z = 4$ the picture is similar, however differences between the low mass halos and the more massive ones increase. The most massive halos themselves now leave their imprint up to a radius of $r_c \sim 1 h^{-1}$ Mpc, whereas the direct influence of the low mass host halos ends at $r_c \sim 0.3 h^{-1}$ Mpc. The low mass halos show a less pronounced large scale overdensity with a shallower decline in density than the more massive halos. The transition point where the density profile reaches the mean cosmic density is again found at a radius of around $r_c \sim 10 h^{-1}$ Mpc. Furthermore the large scale density around the halos shows slightly larger values than at $z = 4.8$, which is due to the ongoing structure formation and the linear growth of the corresponding large wavelength modes.

At $z = 3$ the differences between the various halo masses further increases. The transition point where the profile reaches the mean cosmic density starts to show a dependency on halo mass. The 20 most massive halos reach the mean density at a radius of $r_c > 20 h^{-1}$ Mpc, whereas the low mass halos already reach it at $r_c > 7 h^{-1}$ Mpc. Furthermore the large scale overdensity has grown in density. We therefore expect any influence of this large scale overdensity on the proximity effect measurement to increase with decreasing redshift. Additionally we do not expect a dependency of a possible bias with halo mass at high redshifts, since no large differences between the density profiles are seen. However at low redshifts, differences due to stronger large scale overdensities with increasing halo mass may leave an imprint in the proximity effect profile.

For completeness we show the mean infalling velocities as a function of distance to the host for each halo mass bin in Fig. 6.10. Analogously to the density environment, the infall velocities grow with decreasing redshift due to the onset of structure formation. The infall velocity profiles for the 20 most massive halos and the $M_{\text{halo}} = 10^{12} M_{\odot}$ sample behave similarly at all redshifts. Moving away from the halo, the infall velocity rises until it reaches its maximum at the position where the density profile shows a transition between the host halo and its surrounding density environment. At larger radii the infall velocity steadily decreases until no systematic infall towards the halo is seen anymore. The radii where the systematic infall vanishes is about the size of the large scale overdensity itself. In the case of the low mass halo though, the infall velocity remains constant at low velocities over the whole large scale overdensity.

6.4.4. Proximity effect strength as function of halo mass

We now want to test whether the proximity effect strength parameter a shows any dependency with the host halo mass. For each halo the mean ξ profile is calculated from 100 lines of sight centred on the halo position. Then Eq. 6.5 is fitted to the mean profile and a proximity effect strength parameter a is obtained for each halo in our sample. Any strength parameter with $\log a > 0$ results in an overestimation of the UV background photoionisation rate in real measurements, whereas a $\log a < 0$ resembles an underestimation.

In Fig. 6.11 each halo's mean strength parameter a is marked as a function of redshift and halo mass. The a parameter distribution is thus obtained for each halo mass bin, and its mean and standard deviation are marked by the vertical bold line and grey area in the plot. At each redshift, the mean values of the distribution show no clear dependency with the host halo mass within the 1σ fluctuations. For redshift $z = 4.8$, the different

6. Large scale environmental bias on the LOS proximity effect.

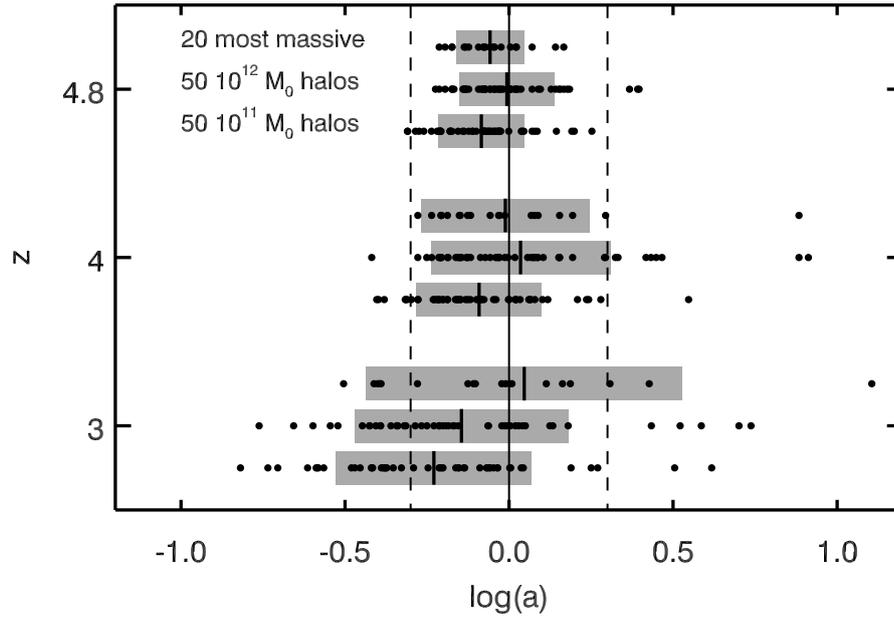


Figure 6.11.: The a parameter distribution as a function of halo mass and redshift. Each black dot represents the a parameter obtained using the mean halo profile of 100 lines of sight. For each redshift the first line shows results from the 20 most massive halos sample, the second line represents the $M_{\text{halo}} = 10^{12} M_{\odot}$ sample, and the third line the $M_{\text{halo}} = 10^{11} M_{\odot}$ sample. The bold black lines indicate the sample's mean a parameter, while the grey shaded areas indicate the corresponding 1σ standard deviation. The solid line marks the fiducial model, while the dashed lines indicate deviations of a factor of two in the UVB measurements.

halo bins yield similar mean values and standard deviations. Considering the variance in the distribution, the input model is regained for all the halo mass bins at $z = 4.8$.

This picture does not change with redshift. Against the expectations motivated by the overdensity profiles discussed in Sect. 6.4.3, the data shows no dependency on the halo mass. Even at $z = 3$, where the mean large scale overdensity was found to be lower around lower mass halos than around high mass ones, no significant dependency is seen. The mean values are thus consistent with the input model within 1σ . The halo to halo variance increases with decreasing redshift and the strength parameter distribution broadens. At $z = 4.8$ the standard deviation of the whole halo sample is $\sigma(\log a) = 0.08$, while at $z = 4$ we find $\sigma(\log a) = 0.20$ and $\sigma(\log a) = 0.36$ for $z = 3$. A similar increase in the variance has been previously observed by Dall’Aglio et al. (2008a) for individual lines of sight of a high-resolution QSO sample.

From these results we conclude that the proximity effect measurements do not show any host halo mass dependent bias. The variance in the signal between different halos increases with decreasing redshift and tighter constraints are obtained with higher redshift objects. However even at low redshifts, the mean strength parameter regained the input value within a factor of less than two.

6.4.5. Influence of the large scale overdensity

In the previous section we have shown that the proximity effect strength parameter is not affected by the host halo mass. However a large scatter around the mean strength parameter is seen in each mass bin, which increases with decreasing redshift. We now want to establish the physical cause of the scatter, and therefore check if this scatter is connected with the density enhancement seen on large scales of up to comoving radii of $r_c = 10 h^{-1}$ Mpc.

For each halo, the mean overdensity profile is calculated using the density distribution along each line of sight. Then the mean over-density $\langle \delta \rangle_R$ between the radii r_1 and r_2 is calculated from the mean density distribution around the halo using

$$\langle \delta \rangle_R = \frac{1}{r_2 - r_1} \int_{r_1}^{r_2} \delta(r) dr. \quad (6.6)$$

We choose $r_1 = 1 h^{-1}$ Mpc comoving in order to exclude the influence of the host halo on the density profile, since we are only interested in the large scale density distribution. The integral extends to comoving $r_2 = 10 h^{-1}$ Mpc which is the characteristic size of the large scale density enhancements (see Sect. 6.4.3).

In Fig. 6.12 the strength parameter a is correlated with the mean overdensity in the comoving radial interval $r_c = [1, 10] h^{-1}$ Mpc. The various symbols indicate the different halo mass samples. At all redshifts, a correlation between the mean density in the vicinity of the QSO and the strength parameter of the mean ξ -profile is seen. A linear regression $\log a = \log \langle 1 + \delta \rangle_0 + \alpha \log \langle 1 + \delta \rangle_{10\text{Mpc}}$ is derived from the data, where $\log \langle 1 + \delta \rangle_0$ is the normalisation point and α is the slope of the regression. The resulting fit parameters are listed in Table 6.3 together with the Pearson product-moment correlation coefficient r of the data set.

Looking at the $z = 4.8$ results, the data points align clearly on a power law relation. With a Pearson correlation coefficient of $r_{\text{Pearson}} = 0.74$ the data exhibits a tight correla-

6. Large scale environmental bias on the LOS proximity effect.

Table 6.3.: Table of linear regression results of the overdensity-strength parameter relation, together with the data’s Pearson product-moment correlation coefficients r . The uncertainties in the fitting parameters have been determined with the bootstrap technique.

z	$\log \langle 1 + \delta \rangle_0$	α	r_{Pearson}
4.8	-0.15 ± 0.07	0.76 ± 0.08	0.74
4.0	-0.19 ± 0.13	1.07 ± 0.16	0.68
3.0	-0.28 ± 0.19	1.11 ± 0.17	0.60

tion between the two quantities. The higher the mean density around a halo becomes, the stronger the proximity effect is biased towards larger strength parameters. At $z = 4.8$ a strength parameter of $a = 1$ (resembling an unbiased measurement of the photoionisation rate) is obtained for a mean overdensity of $\langle 1 + \delta \rangle_{10\text{Mpc}} \sim 1.6 \pm 0.35$. It is interesting to note that there is no strong segregation between the various halo mass bins along the density axis. However, the lower mass halos lie to slightly lower overdensities than the more massive ones. All the different mass bins cover a similar range of large scale densities. Again this indicates that there is no distinct connection between the halo mass and the mean overdensity on these large scales. Hence, only the amount of matter in the greater vicinity of a halo is responsible for the large scatter in the strength parameter a seen in the previous section. However, a small scatter in a of $\approx 20\%$ around the linear regression remains at $z = 4.8$.

Similar findings apply to the lower redshift results. The a parameters are still related through a power law with the large scale overdensity, however the strength of the correlation drops to $r_{\text{Pearson}} = 0.60$ at $z = 3$. Comparing the $z = 4.8$ results with the lower redshift ones shows a slight steepening of the relation from $\alpha = 0.76 \pm 0.08$ at $z = 4.8$ to $\alpha = 1.11 \pm 0.17$ at $z = 3$. The scatter around the power law increases to $\approx 35\%$ for $z = 4$ and 55% for $z = 3$. Further the scatter seems to spread with increasing density, however a larger halo sample from a larger simulation would be needed to conclusively determine this increase in the variance. Again no bias in the strength parameter is found for an overdensity of $\langle 1 + \delta \rangle_{10\text{Mpc}} \sim 1.5 \pm 0.52$ at $z = 4$ and $\langle 1 + \delta \rangle_{10\text{Mpc}} \sim 1.8 \pm 0.77$ at $z = 3$.

These results show that the large scale distribution of matter around a host halo affects the proximity effect strength parameter and biases the measurements of the UV background photoionisation rate. There is a power law correlation between the mean density surrounding a QSO host and the a parameter. The bias on the UV background photoionisation rate can be determined, if the mean overdensity around a QSO is inferred. However, density measurements from the Ly α forest are degenerated with the UV background photoionisation rate. The properties of the UV background have to be known in order to convert optical depths into densities.

Observationally, D’Odorico et al. (2008) have shown that QSOs at $z = 2.6$ are situated in regions showing a density excess on scales of 4 proper Mpc. Due to the dependency of the results on the UV background photoionisation rate, their results only constrain the mean density of the large scale overdensity to a factor of about 3. Nevertheless we estimate a mean density in a radius of 4 proper Mpc from their results. We obtain an upper value on the mean overdensity of $\langle 1 + \delta \rangle \sim 3.5$ and a lower value of $\langle 1 + \delta \rangle \sim 1.25$.

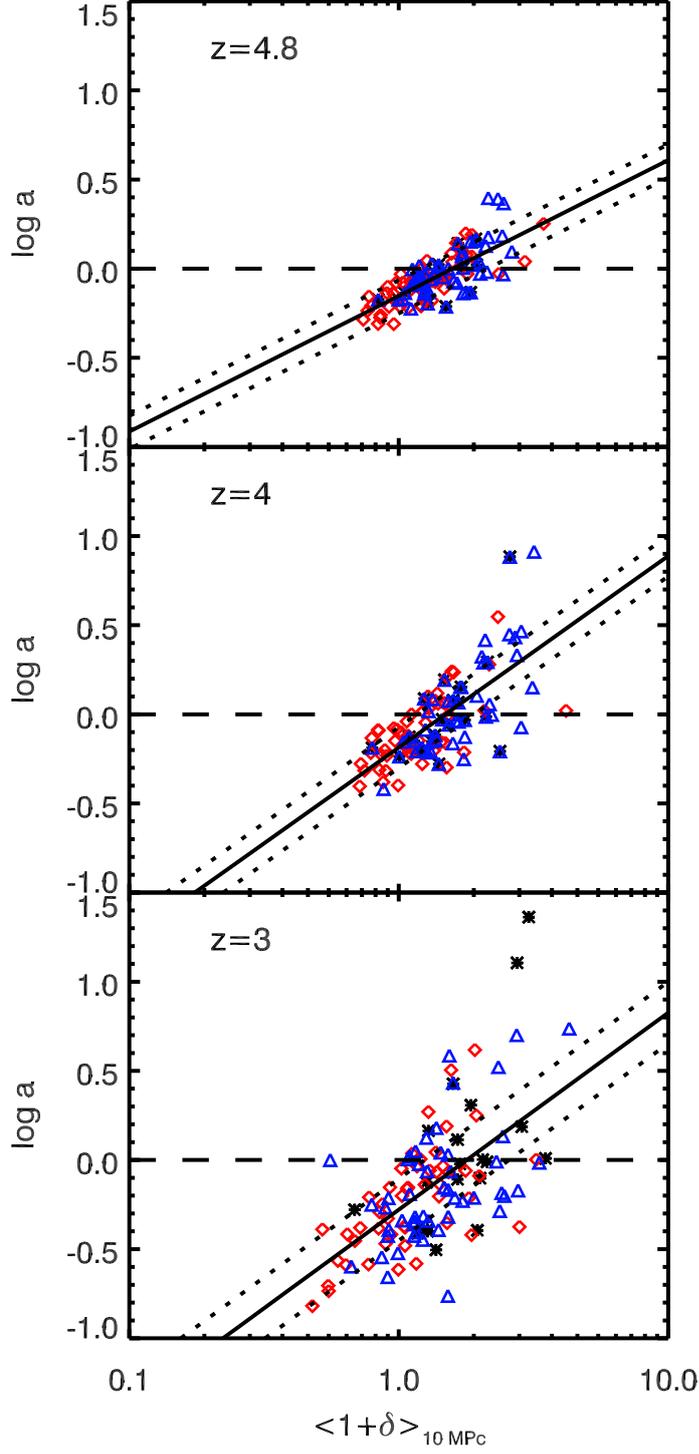


Figure 6.12.: Proximity effect strength parameter as a function of the mean overdensity in a shell of 1 to $10 h^{-1} \text{ Mpc}^{-1}$ comoving radius around the 20 most massive halos (black asterisks), $M_{\text{halo}} = 10^{12} M_{\odot}$ halos (blue triangles), and $M_{\text{halo}} = 10^{11} M_{\odot}$ halos (red diamonds). The black dashed line marks $a = 1$. The solid line represent linear regressions obtained from the dataset where the dotted lines mark the 1σ upper and lower boundaries.

6. Large scale environmental bias on the LOS proximity effect.

According to our results at $z = 3$, this would translate into a bias of the strength parameter of $a = 2.1$ for an overdensity of 3.5 and $a = 0.7$ for the lower value on the mean density around a QSO. Remember though that our results only apply to a QSO with a Lyman limit luminosity of $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$. According to P1 the bias becomes smaller for higher luminosity QSOs.

6.5. Conclusion

We have studied the effect of the host halo mass and the large scale density distribution on measurements of the proximity effect strength parameter. The strength parameter is observationally used to infer the UV background photoionisation rate. From a $64 h^{-1}$ Mpc sized dark-matter simulation, we picked the 20 most massive halos, 50 halos with a mass around $M_{\text{halo}} = 10^{12} M_{\odot}$, and another 50 halos with a mass around $M_{\text{halo}} = 10^{11} M_{\odot}$ at redshifts $z = 4.8, 4,$ and 3 . Each halo is individually assumed to host a QSO with a Lyman limit luminosity of $L_{\nu_{\text{LL}}} = 10^{31} \text{ erg Hz}^{-1} \text{ s}^{-1}$.

Around each halo, 100 random lines of sight of the density and velocity distribution are obtained. Using models of the IGM which we calibrated to observational constraints, the neutral hydrogen fractions and gas temperatures along each line of sight are inferred. To include the increase in ionisation of an additional QSO radiation field, the neutral hydrogen fractions are decreased proportional to geometric dilution of the QSO flux field (Partl et al. 2010). Then, for each line of sight, a Ly α forest spectrum is computed. From the spectrum we derive the mean proximity effect signature with observational methods and measure the proximity effect strength parameter a . Since the UV background photoionisation rate used in generating the Ly α spectra is known, the strength parameter quantifies the over- or underestimation of the UV background.

In order to assess whether the QSO host environment affects the proximity effect signal, a null hypothesis test was performed. From the simulation box, 500 lines of sight originating at random positions with random directions have been obtained. On each line of sight, the proximity effect was modelled. From these lines of sight, we obtained the distribution of normalised optical depths ξ as a function of normalised distance to the QSO ω . From the ξ -distribution, the mean and median proximity effect signals were determined. The mean profiles match the analytic model very well and the model UV photoionisation rates are regained. At large halocentric distances, the median and the mean profiles match. However the median increasingly deviates from the analytic model and the mean profile with decreasing distance to the QSO. This indicates that the ξ -distribution closely resembles a log normal distribution at large distances from the QSO, however it is increasingly skewed when nearing the source. Not only is the variance of the ξ -distribution found to increase with decreasing redshift, but its skewness increases as well.

We further determined the mean density distribution and infall velocity around all the halos in each mass bin as a function of redshift. For all redshifts the density profile does not immediately reach the mean cosmic density at the border of the dark matter halo, however it stays above the mean density up to comoving halocentric radii of $r_c \gtrsim 10 h^{-1}$ Mpc. For redshifts $z = 4.8$ and $z = 4$ the size of this large scale overdensity is independent of the halo mass. However higher mass halos show in the mean a larger

overdensity than lower mass halos. At redshift $z = 3$ the size of the large scale overdensity is smaller for the lower mass sample than for the more massive ones. The velocity field shows a mean infall up to distances of $30 h^{-1}$ Mpc around the halos, reaching for the most massive halos 80, 100, and 150 km/s for redshifts $z = 4.8$, 4, and 3, respectively. Overdensity and infall velocities act together in the derived bias of the proximity effect.

The mean strength parameter per halo mass bin does not show a dependency with the halo mass. For each halo mass bin, a mean strength parameter which is consistent with the input model is regained within 1σ standard deviation. However the halo to halo variance is found to increase with decreasing redshift from $\sigma(\log a) = 0.08$ at $z = 4.8$ to $\sigma(\log a) = 0.36$ at $z = 3$.

The strength parameter is found to correlate with the mean density measured in a shell of comoving $1 - 10 h^{-1}$ Mpc. We fit a power law to this correlation. Regions with a mean density below the cosmic mean show a stronger proximity effect than regions with densities above the cosmic mean. The correlation is tightest at $z = 4.8$ and the scatter around the power law increases with decreasing redshift. Furthermore, the power law is found to steepen slightly with decreasing redshift. From these results we find that an unbiased UV background photoionisation rate can only be obtained if the mean density around the QSO is between $\langle 1 + \delta \rangle_{10\text{Mpc}} \sim 1.5 \pm 0.52$ and 1.8 ± 0.77 .

If a possibility exists to determine the mean density around a QSO, the bias arising from the large scale overdensity can be corrected for. However density measurements from Ly α forest spectra are degenerate with the UV background photoionisation rate. Due to this degeneracy, previous measurements of the density distribution around QSOs, such as D'Odorico et al. (2008), have only been able to constrain the mean density up to a factor of 3. By inferring the bias of the proximity effect at $z \sim 3$ according to their density measurements, the strength parameter ranges from $a = 0.7$ to $a = 2.1$ for a QSO with a Lyman limit luminosity of $L_{\nu_{\text{LL}}} = 10^{31}$ erg Hz $^{-1}$ s $^{-1}$.

*Da steh ich nun, ich armer Tor, und bin so klug als wie
zuvor.*

"Faust" by Johann Wolfgang von Goethe

7

Summary & Conclusions

In this thesis I have focused on three characteristics of the intergalactic medium using simulations and observations. From observed high-resolution spectra, the properties of metal enriched H I absorbers have been determined and the density-metallicity relation was constrained. Additionally I have provided updated results on the distribution of H I absorbers in the Ly α forest. These findings will help to understand the evolution of the intergalactic ionising background field and provide tight constraints on the ionisation state of the intergalactic medium. Since the ionising background field is an important factor in galaxy formation, it is crucial to have good independent observational constraints of the UV background from different indicators. Therefore I contributed to improve measurements of the UV background photoionisation rate as given by the proximity effect, through modelling the effect using detailed radiative-transfer simulations. It turned out, that the impact of the large scale structure bias on proximity effect measurements is a dominant component and can bias the measurements of the UVB photoionisation rates. In order to obtain a more complete and detailed model of the UV background field and the ionisation state of the intergalactic medium in the future, the radiative transfer code CRASH2 was adapted to run on distributed memory clusters. This greatly extends the range of problems that can be studied with the parallel version of CRASH2 and can now be used to model for instance the ionising background radiation. The conclusions of this thesis can be summarised into the following points:

- **Updated absorber number density evolution and differential column density distribution for the Ly α forest at $1.9 < z < 3.2$:**

Using 18 high resolution spectra of the Ly α forest and results from an extensive Voigt profile analysis, updates on the absorber number density evolution and the differential column density distribution of H I absorbers have been obtained. Detailed knowledge about the absorber number density evolution measures the joint evolution of structure formation and the UV background field. Since the formation of structure is very well understood through numerical simulations, the number density evolution can be used to constrain the ionising background field. With the large redshift path coverage of the sample, it was possible to derive the mean absorber number density evolution dn/dz for dense systems ($10^{14} < N_{\text{HI}} < 10^{17} \text{ cm}^{-2}$), which are thought to trace the intergalactic medium near to galaxies, and low density systems ($10^{12.75} < N_{\text{HI}} < 10^{14} \text{ cm}^{-2}$), which trace a large fraction of the intergalactic gas. The absorber number density is found to gradually decrease with decreasing redshift and follows a power law. The gradual build up of clustered structure and the increasing ionisation state of the intergalactic gas results in the

7. Summary & Conclusions

declining absorber number density. The evolution of the dense systems show as well a gradual decline at $z > 2.3$, but at $2 < z < 2.3$ the evolution exhibits a dip indicating that the number of dense absorbers strongly decreases at $z \sim 2.1$. Since the star formation rate is highest at around this redshift, this feature indicates a change in the ionising background flux and with it, a change in the ionisation state of the intergalactic medium. This conclusion is strengthened by the column density distribution function, which describes the distribution of the ionised gas in the intergalactic medium. Generally, the distribution function follows a power law form over a large range of densities. However at intermediate densities of around $N_{\text{HI}} \sim 10^{14.5} \text{ cm}^{-2}$, the distribution function is known to deviate from the power law in the form of a depression. It was possible to show, that this depression evolves with redshift. At $2.7 < z < 3.2$ almost no deviation is seen from the power law, however at lower redshifts the dip becomes increasingly prominent. This feature directly links to the dip seen in the absorber number density evolution of dense systems and thus indicates an increase in the ionising background radiation at $z \sim 2.1$. Together with numerical simulations, these findings might lead to constraints on the ionising background radiation at $z < 2.5$, and might help to better understand the contribution of star forming galaxies to the UV background.

- **Characteristics of C IV enriched H I absorbers in the Ly α forest at $2 < z < 3.5$:**

The Ly α forest is not only composed of H I absorbers, but also contains a small number of highly-ionised metal absorption lines. By associating C IV absorption features with H I absorption, a clear distinction between the metal enriched and unenriched intergalactic gas can be made. By characterising these two distinct populations, insights into the enrichment process of the intergalactic medium through galaxy feedback can be obtained. Additionally, characteristics of the spatial distribution of enriched gas can be inferred. Studying the absorber number density evolution of C IV enriched and unenriched H I Ly α absorbers, reveals that around 40 – 50% of dense systems show C IV enrichment, while at lower densities only 2 – 7% of the gas show signs of metal enrichment. If dense absorbers indeed trace the gas near galaxies and the low density systems trace gas that is far away from galaxies, these findings imply that only the surroundings of galaxies (i.e. the circum-galactic gas) is metal enriched. However, a small fraction of enriched low density systems exists, showing a slight increase in the enrichment fraction with decreasing redshift. Mostly, since these low density H I systems show large quantities of C IV, these systems are highly ionised with high metallicities and are thought to trace gas streams of enriched material originating in galaxies flowing into the intergalactic medium. By studying the column density distribution of the enriched and unenriched H I absorbers, it is clearly seen, that these two types of absorbers resemble distinct populations. The enriched systems show a completely different distribution than the unenriched ones. Further it is shown, that the enriched systems show almost no evolution with redshift, whereas the unenriched systems are seen to evolve. This again indicates, that metal enriched gas is located near galaxies, where they are dominated by the radiation stemming from the galaxy. However the unenriched gas which is thought to trace the intergalactic medium

is dominated by the ionising background radiation. Therefore, any evolution in the unenriched systems might indicate changes in the UV background. However to conclusively verify these findings, detailed modelling of the two populations is needed using numerical simulations which follow the metal enrichment of the intergalactic gas in combination with detailed radiative transfer calculation to derive the ionisation state of the gas and the enclosed metals.

- **Volume averaged density-metallicity relation in the Ly α forest at $2 < z < 3.5$:**

Numerical simulations of metal enrichment in the intergalactic medium predicted that the metallicity in the intergalactic gas remains almost constant for high densities down to a characteristic drop-off density. Below this characteristic drop-off density, the intergalactic low density regions are mostly unenriched. Using 17 high resolution spectra of the Ly α forest, the volume-averaged $N_{\text{HI}} - N_{\text{CIV}}$ relation has been measured. Again, C IV systems are used as tracers of metal enriched gas. However contrary to the study of C IV enriched H I, the focus is not laid on individual H I absorbers, but on volume averaged characteristics of enriched systems. The observed $N_{\text{HI}} - N_{\text{CIV}}$ relation shows a similar behaviour than the results from numerical simulations. At high densities the observed relation is shown to be constant, whereas it drops sharply at a column density of $N_{\text{HI}} \sim 10^{15.2} \text{ cm}^{-2}$. Only highly ionised systems do not follow this trend with H I. The $N_{\text{HI}} - N_{\text{CIV}}$ relation remains constant over the whole redshift range and only the highly ionised systems seem to become more numerous. These findings strengthen the conclusions from the study of individual enriched and unenriched H I absorbers and strongly indicate that only the circum-galactic gas is metal enriched, whereas the bulk of the Ly α forest remains without metals.

- **Analytic model by Bajtlik et al. (1988) captures the physical state of the IGM well:**

Since the ionising background radiation mainly determines the ionisation state in the intergalactic medium and is a crucial process in galaxy formation, it is important to understand the evolution of the background flux. One method of measuring the ionising background photoionisation rate is the QSO line of sight proximity effect. The additional ionising flux stemming from the QSO additionally increases the ionisation state of the surrounding intergalactic medium, reducing the number of H I absorbers in the Ly α forest when approaching the redshift of the QSO. If the ionising flux from the QSO is known, the proximity effect can be used to measure the UV background. Using detailed radiative transfer modelling of the QSO line of sight proximity effect at $3 < z < 5$, it was confirmed that the change in the ionisation fractions due to the additional radiation from the QSO is strictly proportional to geometric dilution. This confirms one of the assumptions needed in analytic formulations of the effect. Further, dense clumps which are able to self shield themselves from the QSO UV radiation are found to cast shadows behind them. In these shadows produced by such Lyman limit systems, the ionisation fraction of hydrogen cannot be increased by the QSO and the proximity effect cannot be detected.

- **Large scatter in the normalised optical depths of the proximity effect, which decreases with redshift and QSO luminosity:**

The normalised optical depth ξ , which relates the optical depth around a QSO as a function of distance to the QSO with the mean optical depth in the Ly α forest, shows strong fluctuations around the analytical Bajtlik et al. (1988) model. These fluctuations arise from the fluctuating matter density distribution and increase with decreasing redshift. However, the impact of the density fluctuations is reduced around QSOs with increased luminosity. The distribution function of the normalised optical depth as a function of distance to the QSO shows an increase in its skewness with decreasing distance. Far from the QSO, the ξ -distribution is found to resemble a log-normal distribution. However for distances shorter than the distance where the flux from the QSO equals the UV background flux, the ξ -distribution deviates from a log-normal function and becomes increasingly skewed towards smaller ξ values. The amount of how much the ξ -distribution is skewed increases with decreasing redshift.

- **The large scale mean density on scales of comoving radii of $10 h^{-1}$ Mpc strongly affects the proximity effect strength:**

The mean large scale density on distances of up to comoving radii of $10 h^{-1}$ Mpc from the QSO host halo is shown to deviate from the mean cosmic density. Assuming dark matter halos of different masses to host QSOs, the influence of the host halo mass and the mean large scale density around the host on the proximity effect strength was determined. No dependency of the proximity effect strength on the host halo mass is found. However the proximity effect strength is found to correlate with the mean large scale density around the QSO host on scales of comoving $10 h^{-1}$ Mpc. QSOs embedded in an environment with a mean density above the cosmic mean show a weakening of the proximity effect strength. If the UV background photoionisation rate is now determined from the proximity effect, this translates in an overestimation of the UV background. However if the QSO is embedded in an underdense environment, the proximity effect is stronger, resulting in an underestimation of the UV background. The mean density around the host relates with the proximity effect strength in the form of a power law and does not show any strong evolution with redshift. In order to obtain UV background measurements which are not biased by the large scale density distribution, the mean overdensity around a QSO needs to be determined. However, measurements of the density require a good knowledge of the UV background photoionisation rate, which poses a challenge to the correction of the large scale density bias in measurements of the UV background with the proximity effect.

- **Parallelisation of the cosmological radiative transfer code CRASH2:**

In order to obtain a detailed model of the ionising background radiation and constraints on the ionisation state of the metals in the intergalactic medium in the future, the cosmological Monte Carlo radiative transfer code CRASH2 has been adopted to run on distributed memory clusters using the MPI library. With a new code called pCRASH2, very good scaling properties can be achieved, as long as the number of sources exceed the number of cores used to run the simulations.

This was made possible by an optimal statical decomposition of the domain using a Peano-Hilbert curve. Along the Peano-Hilbert curve, the computational load estimator is integrated and the curve is cut, whenever the integral reaches a fraction of the total computational load, which is proportional to the number of used cores. As a work load estimator, the optically thin case is considered, which assumes the computational load to be proportional to the geometric dilution of the flux stemming from a point source. This assumption breaks down in the optically thick case, however it is shown, that good scaling results are still possible with this strategy in the optically thick regime. In order to reduce the amount of communication overhead, solving the radiative transfer with the static approximation has been segmented. Rays are only propagated to the borders of a subdomain in one time step and will only be continued by the adjacent subdomain in the next time step. This results in a highly efficient communication scheme. However it has to be assured, that the ionisation fronts never propagate faster then the smallest possible light crossing speed (i.e. the size of a cell divided by the length of a time step). With the parallelisation of CRASH2, a wealth of new problems that require large and high resolution radiative transfer simulations become accessible, such as for instance the evolution and nature of the UV background.

A

CRASH: Cosmological RAdiative transfer Scheme for Hydrodynamics

A.1. Overview of radiative transfer schemes

Radiative transfer (RT) has a longstanding history in astronomy. Especially in the field of stellar atmospheres a large amount of knowledge about the transport of photons has been accumulated over the last century (see e.g. Mihalas & Weibel Mihalas (1984)). Many of the radiation transfer codes used in stellar atmospheres obtain solutions in 1D by assuming radial symmetry.

Unfortunately, applying radial symmetry in a cosmological context is not feasible due to the filamentary nature of our universe on scales larger than galaxies. Studying radiative transfer in a cosmological context thus requires 3D methods that capture properly the density features arising during structure formation. This is only feasible in a simplified case by considering hydrogen (and in some codes helium) atom physics.

Over the last decade, a wealth of different numerical methods emerged solving radiative transfer in 3D coupled to hydrogen physics. Most effort has been focusing on 3D continuum radiative transfer that follows changes in the gas ionisation state. However, some hydrogen Lyman α line transfer codes for arbitrary geometries have been developed as well (Verhamme et al. 2006; Pierleoni et al. 2009). Table A.1 gives an overview of some of the codes available for solving the continuum radiative transfer problem.

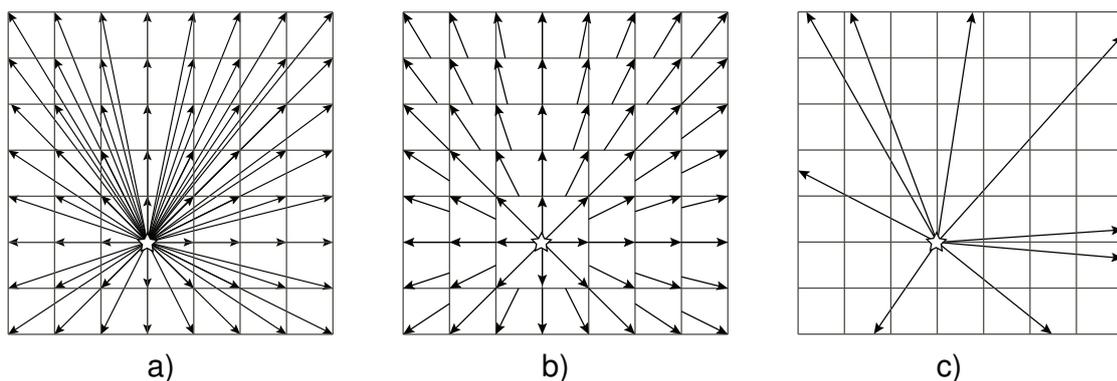


Figure A.1.: Comparison of the long (a) and short characteristics (b) method with the Monte Carlo ray tracing technique (c).

Most of these codes make use of one of the following methods (see Fig. A.1):

- **Ray tracing:**

Ray tracing methods reduce the RT problem to one dimension and solve the radiative transfer along a one dimensional ray. From each source rays need to be cast to each cell in the computational domain. This can be done in three different ways (see Fig. A.1):

- **Long characteristics method:**

In the long characteristics method, a ray is fully cast from each source to each cell in the domain. Each time a ray is cast, the optical depth in all the cells lying between the source and the destination cell needs to be evaluated. This causes a large computational overhead, since the same cells are evaluated multiple times.

- **Short characteristics method:**

The short characteristics method eases the main disadvantage of the long characteristics method, by determining the optical depth in each cell individually. Like in the long characteristics method, a ray is cast from the source to each cell. However the optical depth is not evaluated by calculating the full optical depth between the cell and the source, but only the contribution of the last cell to the ray's optical depth is calculated and stored. The total optical depth for any ray originating at the source to any cell is then determined by summing up and interpolating the contributions of the various individual cells to the ray's optical depth (see Figure A.1). In this scheme, each cell has to be evaluated only once per source, thus speeding up the computation substantially.

- **Monte Carlo ray tracing:**

Monte Carlo ray tracing is a simplification of the long characteristics method in such a way that the solution is approximated by stochastically emitting rays from each source in random directions. A computational speed up is achieved by constraining the solution's accuracy through a reduced sampling of the computational domain.

- **Variable Eddington tensor:**

In this approach the moments of the radiative transfer equations are used. This leads to a simplified radiative transfer equation that can be easily solved. The radiation field is then treated similarly to a fluid as an diffusion problem. This has the advantage that coupling to hydrodynamics is straight forward. One of the disadvantages though is the introduction of elongations when H II regions start to overlap (see Petkova & Springel (2009) for detailed discussion).

Besides these approaches, various other approaches exist. Some codes make use of a Delauney tessellation of the optical depth (Ritzerveld & Icke 2006) in such a way, that each point on the grid is separated by the same optical depth. This method has the disadvantage, that for each frequency bin a distinct tessellation needs to be found. A different approach is taken by Pawlik & Schaye (2008) for inclusion of radiative

transport in SPH codes, by transporting cones of light from particles to particles. Other strategies involve fast approximative formulations of the transport problem (Alvarez et al. 2006), directly solving the transfer equation in 3D (two spatial dimensions, one angular dimension) using a WENO (Weighted Essentially Non-Oscillatory) scheme for hyperbolic partial differential equations (Qiu et al. 2008), and using fast Fourier transforms to solve the RT equation (Cen 2002).

In order to fully understand how well all these different codes perform, a comparison project has been started with defined test problems that allow a simple and direct comparison of different codes (Iliev et al. 2006). Even though the number of different methods is large, all codes produce reasonably similar results. This picture is, however, altered when RT codes coupled with hydrodynamics are compared (Iliev et al. 2009). In this case, large differences between various codes arise. Special care needs to be taken when applying any radiative-hydrodynamics scheme to a specific problem. It needs to be checked whether a specific code is suited for a particular problem or not.

In the following section we summarise the **CRASH** scheme as it is currently implemented. It will become clear, why **CRASH** is a very powerful and flexible scheme for solving the radiative transfer equation. With **CRASH** it is especially easy to include new physical processes, such as X-ray radiation or H₂ physics. The latest version of **CRASH** supports full polychromatic radiation transport without resorting to monochromatic transport. Due to the way **CRASH** handles polychromatic transport, spectral filtering effects are included instantly. Another feature of **CRASH** is that the scheme does not rely on the on-the-spot approximation and that the diffuse recombination radiation is followed explicitly. The only disadvantage of **CRASH** however is the large number of photon samples that need to be followed to reduce the Monte Carlo noise when involving more elaborated physics. However, thanks to the parallel version of **CRASH**, a substantially larger sampling rate is achievable.

Before discussing **CRASH**, the foundations of radiation transport will be discussed. Further, we will establish the impact of the instant propagation approximation and elaborate, under which circumstances it can be safely applied and when caution is needed.

A.1.1. The radiative transfer equation

We will now derive the radiative transfer equation in an expanding universe. For this we follow Dautcourt (1976) and Gnedin & Ostriker (1997). First we start by defining the photon distribution function $F_\gamma(t, x^i, p^k)$ where x^i are comoving coordinates and p^k the comoving photon momenta

$$p^k = a \frac{h\nu}{c} n^k. \quad (\text{A.1})$$

Here $a = a(t)$ is the cosmological expansion factor, h is the Planck constant, ν is the photon frequency, and n^k denotes the unit vector pointing in the direction of the photon propagation. We thus can write the Boltzmann equation for the distribution function F_γ

$$\frac{\partial F_\gamma}{\partial t} + \frac{\partial}{\partial x^i} (\dot{x}^i F_\gamma) + \frac{\partial}{\partial p^k} (\dot{p}^k F_\gamma) = \frac{\partial F_\gamma}{\partial t} \Big|_{\text{sources}} - \frac{\partial F_\gamma}{\partial t} \Big|_{\text{sinks}} \quad (\text{A.2})$$

A. CRASH: Cosmological RADIative transfer Scheme for Hydrodynamics

Code (Author)	Type	H	He	rec. case	Grid type	speed of light	coupled to hydro	parallelisation
CRASH2 (Maselli et al. 2009)	MC ray tracing	x	x	A	fixed	instant	-	distributed
SPHRAY (Alhay et al. 2008)	MC ray tracing	x	x	A/B	particles	instant	-	-
RADAMESH (Cantalupo & Porciani 2010)	MC ray tracing	x	x	A	fixed/AMR	c	-	shared
LICORICE (Baek et al. 2009)	MC ray tracing	x	-	-	AMR	instant	x	shared
C^2 -Ray (Mellem et al. 2006)	short characteristics	x	in dev.	B	fixed/AMR	-	x	shared
RSPH (Susa & Umemura 2004)	long characteristics	x	-	B	particles	-	x	distributed
ART (Nakamoto et al. 2001) Zeus-MP	short characteristics	x	-	A	fixed	-	-	distributed
(Whalen & Norman 2006)	long characteristics	x	x	B	fixed	instant	x	distributed
FLASH-HC (Rijkhorst et al. 2006)	hybrid characteristics	x	-	B	fixed/AMR	-	x	distributed
START (Hasegawa & Umemura 2010)	long characteristics	x	-	B	particles	-	x	-
FTTE (Razoumov & Cardall 2005)	adaptive ray-splitting	x	x	A	fixed/AMR	-	x	-
Enzo-RT (Abel & Wandelt 2002)	adaptive ray-splitting	x	-	B	fixed/AMR	c	x	distributed
OTVET (Gnedin & Abel 2001)	variable Eddington tensor	x	x	A	fixed	reduced	x	shared
GADGET-3 (RT module) (Petkova & Springel 2009)	variable Eddington tensor	x	-	B	particles	-	-	distributed
ATON (Aubert & Teyssier 2008)	variable Eddington tensor	x	-	A	fixed/AMR	reduced or c	-	GPU
HART (Gnedin et al. 2009)	variable Eddington tensor	x	x	A	fixed/AMR	reduced	x	shared/distributed
TRAPHIC (Pawlik & Schaye 2010)	tracing photons from particle to particle	x	x	A	particles	c	x	distributed
SimpleX2 (Paardekooper et al. 2010)	Delaney tessellation of τ	x	-	A	unstructured	instant	-	distributed
IFT (Alvarez et al. 2006)	ionisation front tracking	x	-	B	fixed/AMR	-	-	-

Table A.1.: Overview of some RT codes available in the literature.

where the right hand side gives the sources due to emission processes and the sinks due to absorption of photons. We further relate the specific intensity $J_\nu(t, x^i, n^k)$ to the photon distribution function F_γ using the relations $a^3 d^3x = c dt dA$ and $d^3p = p^2 dp d\Omega$. We thus obtain

$$J_\nu = h\nu F_\gamma \frac{d^3x d^3p}{d\nu d\Omega dA dt} = \frac{h^4 \nu^3}{c^2} F_\gamma \quad (\text{A.3})$$

where dA denotes an infinitesimal area element and $d\Omega$ denotes the solid angle. By substituting F_γ in Eq. A.2 using Eq. A.3 we obtain the cosmological radiative transport equation

$$\frac{\partial J_\nu}{\partial t} + \frac{\partial}{\partial x^i} (\dot{x}^i J_\nu) - H \left(\nu \frac{\partial J_\nu}{\partial \nu} - 3J_\nu \right) = -k_\nu J_\nu + S_\nu \quad (\text{A.4})$$

where H is the Hubble constant $H = \dot{a}/a$, k_ν the absorption coefficient and S_ν the emission function describing the sources. By identifying \dot{x}^i with the velocity vector of photons $\dot{x} = n^k c$, we can rewrite Eq. A.4 into

$$\frac{1}{c} \frac{\partial J_\nu}{\partial t} + \frac{n^k}{a} \frac{\partial}{\partial x^i} (J_\nu) - \frac{H}{c} \left(\nu \frac{\partial J_\nu}{\partial \nu} - 3J_\nu \right) = -k_\nu J_\nu + S_\nu. \quad (\text{A.5})$$

If we assume that the mean free path of a photon to be much smaller than the horizon (which is assumed from now on), the radiative transfer equation turns into its classical form

$$\frac{\partial J_\nu}{\partial t} + n^k \frac{\partial}{\partial x^i} J_\nu = -k_\nu J_\nu + S_\nu. \quad (\text{A.6})$$

This is the equation that is solved in **CRASH**. However, care must be taken at all times however, that the mean free path much smaller than the horizon. In the preionisation universe when H II regions are still fairly small, this is always ensured. However, as soon as H II regions start to overlap and the mean free path increases substantially, one has to assess whether the inclusion of the cosmological expansion alters the result or not.

One additional simplification can be applied to the radiative transfer equation regarding the propagation speed of photons. If changes in k_ν and S_ν are on timescales larger than the light crossing time of the simulation box (or the timescales of interest), photons can be assumed to instantly propagate through the simulation volume. This is the static approximation to the transfer equation, which is currently used by many RT schemes (compare Table A.1). The RT equation A.6 thus simplifies to

$$n^k \frac{\partial}{\partial x^i} J_\nu = -k_\nu J_\nu + S_\nu. \quad (\text{A.7})$$

Since radiation is assumed to propagate instantaneously through space, unphysical results can arise in certain situations. It was pointed out by Abel et al. (1999) that very close to a source, ionisation fronts might travel artificially faster than the speed of light. In this early formation phase of an H II regions, the results cannot be relied on, if the instant propagation approximation is used. The present version of **CRASH** uses this approximation. However, with the parallelised version in place, changes allowing the photons to propagate at speed of light are feasible.

A.2. How it all works - The CRASH RT scheme

The following sections describe the latest implementation of CRASH which was published in the following papers: Ciardi et al. (2001), Maselli et al. (2003), Maselli & Ferrara (2005), Maselli et al. (2009), Chapter 4, and Schleicher et al. in preparation. Currently three versions of CRASH exist, which are planned to merge into one code base in the future. These are the classical CRASH code, the parallel pCRASH2 code and a version implementing X-Rays called xCRASH. We will give a combined description of these versions below, outlining differences when necessary. A detailed description of pCRASH2 is found in Section 4.

The basis of the CRASH RT scheme is simple and straight forward. Given a distribution of sources together with their characteristics and an initial hydrogen and/or helium density field combined with a temperature field, photon packages are followed through the computational domain, altering the ionisation state and temperature of the gas on their way through the grid. Each source emits photon packets stochastically. These photon packets represent a group of photons, since it is impossible to track individual photons.

At each time step, photon packets are sent out by each source in random directions. The angular distribution with which a source produces photon packets can be easily controlled and modified. With this modification, anisotropically emitting sources can be implemented naturally. The total number of packets produced by each source determines the angular sampling resolution. The more packets each source emits, the better the solution is constrained and the Monte Carlo simulation noise is reduced.

Each photon packet represents a polychromatic beam of light, which we call "coloured packet". Each packet has a multiple of wavelength bins sampling the spectral energy distribution of the source. These coloured packets are then propagated through the box until they either exit the box or get completely absorbed. The number of photons absorbed per wavelength bin is calculated each time a package crosses a cell and is subtracted from the propagating package. With this procedure, wavelength dependent filtering effects are naturally followed by CRASH.

It is also straight forward to solve for the diffuse radiation field produced by recombining electrons. In each cell the number of electrons recombining with hydrogen and helium are counted. If a certain threshold of recombinations has been reached, a diffuse monochromatic photon packet according to the corresponding Milne spectrum is released from the cell and the cell's recombination counter is set to zero.

A.2.1. Coloured photon packets

CRASH discretises the photon field into coloured photon packets. Each packet resembles a discrete photon energy content, which is distributed according to a given spectral energy distribution S_ν into discrete frequency bins $\Delta\nu$. Quantities continuous in frequency we denote with subscript ν and frequency discretised quantities with subscript $\Delta\nu$. CRASH does not assume a fixed frequency discretisation; it reads arbitrary frequency binnings from input files.

For point sources, the emitted energy per time step j and source s reads

$$\Delta E_{j,\Delta\nu,s} = \int_{t_{j-1}}^{t_j} \int_{\Delta\nu} L_{0,s}(t) S_{\nu,s} d\nu dt, \quad (\text{A.8})$$

where $t_j = j \Delta t$ ($j = 1, \dots, N_t$), N_t the number of time steps, and $L_{0,s}(t)$ denotes the time dependent normalisation point for a source's spectrum $S_{\nu,s}$ with bolometric luminosity $L_s(t)$. The time step is given by $\Delta t = t_s/N_t$, where t_s is the total simulation time. In CRASH and xCRASH, each source only emits one packet per time step, resulting in $N_t = N_p$, where N_p is the number of packets produced per source. In pCRASH2 a source can emit multiple photons per time step and the emitted energy per time step $\Delta E_{j,\Delta\nu,s}$ is equally divided among the number of packets per time step $(N_p/N_t) \in \mathbb{N}$.

Internally CRASH does not use radiation energy to follow the radiative transport equation. It is more convenient to use the number of photons in each $\Delta\nu$ bin

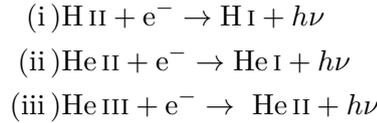
$$N_{\gamma,j,\Delta\nu,s} = \frac{\Delta E_{j,\Delta\nu,s}}{h\nu} \quad (\text{A.9})$$

where h is the Plank constant. For point sources, the packets originate at the source's position \mathbf{P} , and a random propagation direction (θ, φ) is chosen according to the source's angular emissivity function.

Hence, a coloured photon packet is characterised by the following parameters: the origin \mathbf{P} of the packet, the propagation direction (θ, φ) , the frequency bin array $\Delta\nu_i$, and the number of photons per frequency bin array $N_{\gamma,j,\Delta\nu}$.

Diffuse recombination radiation

CRASH accounts for a diffuse radiation field produced by recombining electrons with hydrogen or helium. The following recombination processes of H and He are treated:



As with point sources, the diffuse recombination radiation is discretised into coloured packets. Since the emission of recombining electrons is isotropic, a random propagation direction is assigned to the packet accounting for an isotropic distribution on a sphere. Then the spectral energy distribution is calculated. By assuming local thermal equilibrium, the Milne spectrum

$$S_{d,\nu} = \frac{2h\nu^3}{c^2} \frac{g_k}{g_{k+1}} \left(\frac{h^2}{2\pi m_p k T_e} \right)^{3/2} \sigma_{\Xi,\nu} e^{-h(\nu-\nu_{\text{th}})/(k_B T_e)} n_{\Xi} n_e \quad (\text{A.10})$$

is evaluated, where g_k and g_{k+1} are the species ground-level statistical weights, T_e and n_e are the electron temperature and number density, n_{Ξ} gives the number density of species $\Xi \in \{\text{H II}, \text{He II}, \text{He III}\}$, $\sigma_{\Xi,\nu}$ is the species cross section, and all the other symbols

have their usual meaning. Eq. A.10 only holds for $\nu \geq \nu_{\text{th}}$, where ν_{th} is the ionisation frequency of a given species. In CRASH and xCRASH Eq. A.10 is sampled in a Monte Carlo fashion to assign a frequency and intensity to a monochromatic photon packet. In pCRASH2 we use equally spaced frequency bins to evaluate Eq. A.10 and produce a coloured packet.

The intensity of the recombination radiation is determined as follows. In each cell of the domain, the number of recombination events for each species $N_{\Xi, \text{rec}}$ is counted. Whenever a packet crosses a cell at time t_c , the number of recombined electrons after the last cell crossing by a packet t_{c-1} is added to $N_{\Xi, \text{rec}}$ using

$$\Delta N_{\Xi, \text{rec}} \approx \alpha_{\Xi}(T_{c-1}) n_{e, c-1} n_{\Xi, c-1} (t_c - t_{c-1}) (\Delta x)^3 \quad (\text{A.11})$$

where T_{c-1} , $n_{e, c-1}$, and $n_{\Xi, c-1}$ are the cell temperature, electron number density and number density of species Ξ at the last cell crossing time t_{c-1} . $\alpha_{\Xi}(T_{c-1})$ is the species total recombination coefficient and $(\Delta x)^3$ the volume of one cell. Whenever $N_{\Xi, \text{rec}}$ reaches a fraction of the total numbers of atoms of species Ξ in the cell $N_{\Xi, a}$

$$N_{\Xi, \text{rec}, c} > f_r N_{\Xi, a}, \quad (\text{A.12})$$

with $f_r \in [0, 1]$. This gives the intensity of the diffuse photon packet.

Since $\alpha_{\Xi}(T_{c-1})$ captures all recombining electrons, including the ones recombining not only to the ground level but also to different energy levels, we need to determine the fraction of photons capable of ionising further atoms. Each time Eq. A.12 is satisfied, the probability of recombining to the ground level α_1/α is Monte Carlo sampled, where α_1 is the recombination coefficient to the ground level only. Only recombination photons capable of ionising another atom are considered. If recombination occurs to the ground level for any of the species, a diffuse packet according to the Milne spectrum is emitted. For helium, a slightly more complicated approach has to be taken, since de-excitation following recombination to higher levels of He II or He III is capable of ionising H I or He I. De-excitation can proceed through various channels. However, CRASH approximates the relevant transitions by considering mean values between the $2s \rightarrow 1s$ ($h\nu = 19.8$ eV) and the $2p \rightarrow 1s$ ($h\nu = 21.2$ eV) channels for He I, assuming equipartition between the two processes. For He II, the only relevant process is $2p \rightarrow 1s$ ($h\nu = 40.7$ eV). A more detailed treatment could be implemented into pCRASH2, since it allows for more elaborated physics to be followed. After emission of the diffuse packet, N_{rec} is set to zero.

A.2.2. Package propagation

Whenever a packet was produced, it is propagated through the box. In CRASH and xCRASH, this happens immediately after a packet was produced. Due to the parallelisation strategy, pCRASH2 first produces all the packets and adds them onto a list, which will then be processed after all the packets have been produced and/or have been communicated with neighbouring domains.

In the process of creating a packet, its initial position \mathbf{P}_0 and propagation direction (θ, φ) are assigned. The packet is then propagated along a straight line to the l -th cell at

distance r according to

$$\mathbf{P}_l = \mathbf{P}_0 + r \cdot \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}. \quad (\text{A.13})$$

In order to properly progress from cell to cell and to obtain the intersection length Δx_l of the ray with the l -th cell, the fast voxel traversal algorithm by Amanatides & Woo (1987) is used.

At each cell crossing, a fraction of the coloured packet is absorbed as a function of frequency. The absorption probability P_{abs} through a cell l with an optical depth τ_l is given by

$$P_{\text{abs}}(\tau_{l,\Delta\nu}) = 1 - e^{-\tau_{l,\Delta\nu}}. \quad (\text{A.14})$$

The optical depth at each frequency bin in the coloured packet is thus given by

$$\tau_{l,\Delta\nu} = \sum_{\Xi} \tau_{l,\Xi,\Delta\nu} = \sum_{\Xi} \sigma_{\Xi,\Delta\nu} n_{l,\Xi} \Delta x_l, \quad (\text{A.15})$$

where $n_{l,\Xi}$ is the species number density in cell l . Using Eq. A.14 the number of photons absorbed in cell l becomes

$$N_{l,\gamma,\Delta\nu}^{\text{abs}} = N_{l-1,\gamma,\Delta\nu}^{\text{trans}} (1 - e^{-\tau_{l,\Delta\nu}}), \quad (\text{A.16})$$

where $N_{l-1,\gamma,\Delta\nu}^{\text{trans}}$ denotes the number of photons transmitted up to the preceding cell $l-1$. Due to numerical noise, care has to be taken, that the number of absorbed photons does not exceed the total number of photons in the frequency bin of the packet. Since the absorption probability is a function of frequency, each cell crossing will modify the spectral shape of a packet. Typically, photons close to the ionisation edge of a species Ξ will be depleted first and more strongly. This scheme captures spectral filtering naturally.

Photons are propagated until they either exit the box (when no periodic boundary conditions are imposed) or until they are fully absorbed. This is the case, when the number of photons in all the frequency bins fall below a threshold $N_{l,\gamma,\Delta\nu} \leq 10^{-f_{\text{abs}}} N_{\gamma,\Delta\nu}$. This guarantees energy conservation up to an accuracy of $10^{-f_{\text{abs}}}$. Typically a $f_{\text{abs}} \in [4, 9]$ is adopted, depending on the required accuracy.

A.2.3. Solving for the chemistry

After the total number of absorbed photons in the cell is known, the ionisation fractions x_{Ξ} of the various species and the gas temperature T need to be determined. The system

of coupled ordinary differential equations used by CRASH are

$$n_{\text{H}} \frac{dx_{\text{HII}}}{dt} = \gamma_{\text{HI}}(T)n_{\text{HI}}n_e - \alpha_{\text{HII}}(T)n_{\text{HII}}n_e + \Gamma_{\text{HI}}n_{\text{HI}} = I_{\text{HII}} \quad (\text{A.17})$$

$$n_{\text{He}} \frac{dx_{\text{HeII}}}{dt} = \gamma_{\text{HeI}}(T)n_{\text{HeI}}n_e - \gamma_{\text{HeII}}(T)n_{\text{HeII}}n_e - \alpha_{\text{HeII}}(T)n_{\text{HeII}}n_e + \alpha_{\text{HeIII}}(T)n_{\text{HeIII}}n_e + \Gamma_{\text{HeI}}n_{\text{HeI}} = I_{\text{HeII}} \quad (\text{A.18})$$

$$n_{\text{He}} \frac{dx_{\text{HeIII}}}{dt} = \gamma_{\text{HeII}}(T)n_{\text{HeII}}n_e - \alpha_{\text{HeIII}}(T)n_{\text{HeIII}}n_e + \Gamma_{\text{HeII}}n_{\text{HeII}} = I_{\text{HeIII}} \quad (\text{A.19})$$

$$\frac{dT}{dt} = \frac{2}{3k_{\text{B}}n_{l,\text{tot}}} \left[-k_{\text{B}}T \frac{dn_{l,\text{tot}}}{dt} + \sum_{\Xi} \mathcal{H}(T, x_{\Xi}) - \sum_{\Xi} \Lambda(T, x_{\Xi}) \right]. \quad (\text{A.20})$$

Here γ_{Ξ} denote collisional ionisation rates, α_{Ξ} are the recombination rates and Γ_{Ξ} stands for the photoionisation rates. In xCRASH secondary ionisation through energetic electrons Ψ_{Ξ} is also included in Eqs. A.17, A.18, and A.19. Further, $\mathcal{H}(T, x_{\Xi})$ and $\Lambda(T, x_{\Xi})$ are the heating and cooling functions which capture collisional ionisation cooling, recombination cooling, collisional excitation cooling, Bremsstrahlung cooling, Compton cooling and heating, and photoionisation heating. Furthermore, Coulomb heating of thermal electrons by fast electrons is considered in xCRASH. All the corresponding rates and cross sections used in CRASH are given in Section A.2.4.

The set of equations A.17 - A.20 needs to be discretised in time, and we obtain

$$\begin{aligned} x_{\text{HII}}(t + \Delta t) &= x_{\text{HII}}(t) + I_{\text{HII}}(t)\Delta t/n_{\text{H}} \\ x_{\text{HeII}}(t + \Delta t) &= x_{\text{HeII}}(t) + I_{\text{HeII}}(t)\Delta t/n_{\text{He}} \\ x_{\text{HeIII}}(t + \Delta t) &= x_{\text{HeIII}}(t) + I_{\text{HeIII}}(t)\Delta t/n_{\text{He}} \\ T(t + \Delta t) &= T(t) + \\ &\quad \frac{2}{3k_{\text{B}}n_{l,\text{tot}}} \left[-k_{\text{B}}T\Delta n_e + \Delta t \left\{ \sum_{\Xi} \mathcal{H}(T, x_{\Xi}) - \sum_{\Xi} \Lambda(T, x_{\Xi}) \right\} \right] \end{aligned} \quad (\text{A.21})$$

where $\Delta n_e = n_e(t + \Delta t) - n_e(t)$. This system is solved every time a photon packet crosses a cell. The numerical integration time-step Δt , being the time interval between two subsequent crossings of a cell by a photon packet, is not constant. Therefore, for each cell the last crossing time t_{c-1} of a photon packet needs to be saved. Thus $\Delta t = t_c - t_{c-1}$.

Most of the processes in Eq. A.21 are treated continuously, except for the processes connected with photoionisation. Since the photon field is discretised into photon packets, photoionisation is a discrete process in CRASH. In order to determine the photoionisation rate Γ_{Ξ} , the total number of absorbed photons $N_{l,\gamma,\Delta\nu}^{\text{abs}}$ (see Eq. A.16) is distributed among the species proportional to the species' optical depth (Mellema, Ringberg conference on

reionisation 2009). We thus obtain

$$\begin{aligned}
\Delta x_{l,\text{H I}} &= \frac{n_{\text{H I}}}{n_{\text{H}}} \Gamma_{\text{H I}} \Delta t = \frac{1}{n_{l,\text{H}} (\Delta x)^3} \sum_{\Delta\nu} N_{l,\gamma,\Delta\nu}^{\text{abs}} \frac{\tau_{l,\Delta\nu,\text{H I}}}{\tau_{l,\Delta\nu}} \\
\Delta x_{l,\text{He I}} &= \frac{n_{\text{He I}}}{n_{\text{He}}} \Gamma_{\text{He I}} \Delta t = \frac{1}{n_{l,\text{He}} (\Delta x)^3} \sum_{\Delta\nu} N_{l,\gamma,\Delta\nu}^{\text{abs}} \frac{\tau_{l,\Delta\nu,\text{He I}}}{\tau_{l,\Delta\nu}} \\
\Delta x_{l,\text{He II}} &= \frac{n_{\text{He II}}}{n_{\text{He}}} \Gamma_{\text{He II}} \Delta t = \frac{1}{n_{l,\text{He}} (\Delta x)^3} \sum_{\Delta\nu} N_{l,\gamma,\Delta\nu}^{\text{abs}} \frac{\tau_{l,\Delta\nu,\text{He II}}}{\tau_{l,\Delta\nu}} \\
\Delta T_l &= \frac{2}{3k_{\text{B}} n_{l,\text{tot}}} \left[-k_{\text{B}} T_l \Delta n_e + (\Delta x)^{-3} \sum_{\Xi,\Delta\nu} N_{l,\gamma,\Delta\nu}^{\text{abs}} \frac{\tau_{l,\Delta\nu,\Xi}}{\tau_{l,\Delta\nu}} (h(\Delta\nu - \nu_{\text{th},\Xi})) \right]
\end{aligned} \tag{A.22}$$

All other processes such as recombination, collisional ionisation, and the cooling function are not discretised quantities and are thus treated as continuous processes. In order to correctly capture these with a simple Eulerian integration scheme, Δt needs to be much smaller than the characteristic time-scale of these processes for all species Ξ . It has to be ensured at all times, that $\Delta t \ll t_{\text{min}} = \min [t_{\text{rec},\Xi}, t_{\text{coll},\Xi}, t_{\text{cool}}]$ is fulfilled. If this is not the case, the integration is subsampled using $n_s = \text{int} [\Delta t / (f_s t_{\text{min}})]$ steps, with f_s being a fudge factor. It is generally chosen to be $f_s \in [50, 100]$.

A.2.4. Cross-sections and rates

Photoionisation cross-sections

CRASH, pCRASH2 (**Osterbrock & Ferland 2006**) Cross-sections are given in [cm^2].

$$\begin{aligned}
\text{H I} : \sigma_{\text{H I}}(\nu) &= 6.3 \times 10^{-18} (\nu/\nu_{\text{th},\text{H I}})^{-3} \\
\text{He I} : \sigma_{\text{He I}}(\nu) &= 7.2 \times 10^{-18} \left[1.66 (\nu/\nu_{\text{th},\text{He I}})^{-2.05} + 0.66 (\nu/\nu_{\text{th},\text{He I}})^{-3.05} \right] \\
\text{He II} : \sigma_{\text{He II}}(\nu) &= 1.58 \times 10^{-18} (\nu/\nu_{\text{th},\text{He II}})^{-3}
\end{aligned}$$

xCRASH (**Verner & Yakovlev 1995**) Cross-sections are given in [Mb] (megabarn).

$$\begin{aligned}
y &= E/E_0 \\
Q &= 5.5 + l - 0.5P \\
F(y) &= \left[(y-1)^2 + y_{\text{w}}^2 \right] y^{-Q} \left(1 + \sqrt{(y/y_{\text{a}})} \right)^{-P} \\
\sigma_{nl}(E) &= \sigma_0 F(E/E_0)
\end{aligned}$$

with

	E_0	P	σ_0	y_w	y_a	l
H I	4.298×10^{-1}	2.963	5.475×10^4	0	32.88	0
He I	5.996	6.098	4.47×10^3	0	2.199	0
He II	1.720	2.963	1.369×10^4	0	32.88	0

Recombination rates (Cen 1992)

Rates are given in $[\text{cm}^3\text{s}^{-1}]$, T in $[K]$.

$$\text{H II} : \alpha_{\text{H II}}(T) = 8.40 \times 10^{-11} T^{-1/2} \left(\frac{T}{10^3} \right)^{-0.2} \left[1 + \left(\frac{T}{10^6} \right)^{0.7} \right]^{-1}$$

$$\text{He II} : \alpha_{\text{He II}}(T) = 1.50 \times 10^{-10} T^{-0.6353}$$

$$\text{He III} : \alpha_{\text{He III}}(T) = 3.36 \times 10^{-10} T^{-1/2} \left(\frac{T}{10^3} \right)^{-0.2} \left[1 + \left(\frac{T}{10^6} \right)^{0.7} \right]^{-1}$$

Collisional ionisation rates (Cen 1992)

Rates are given in $[\text{cm}^3\text{s}^{-1}]$.

$$\text{H I} : \gamma_{\text{H I}}(T) = 5.85 \times 10^{-11} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-157809.1/T}$$

$$\text{He I} : \gamma_{\text{He I}}(T) = 2.38 \times 10^{-11} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-285335.4/T}$$

$$\text{He II} : \gamma_{\text{He II}}(T) = 5.68 \times 10^{-12} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-631515.0/T}$$

Secondary ionisation rates (Dalgarno et al. 1999)

CRASH, pCRASH2 Not included.

xCRASH The number of ionisation events stemming from fast electrons is given by

$$N_{\text{sec.ion}} = \frac{E_e}{W(E)},$$

where E_e is the electrons' energy and

$$W(E) = W_0(E) \left(1 + C(E)x_e^{\alpha(E)} \right)$$

is the fit to the mean energy per ion pair and x_e is the electron fraction. The fitting parameters for a H-He gas are for H II:

E (eV)	W_0 (eV)	α	C
30	60.0	1.06	221
50	48.9	1.01	87.9
100	42.6	0.964	41.0
200	40.3	0.928	24.5
500	39.8	0.890	15.5
1000	39.8	0.866	12.2

for He I:

E (eV)	W_0 (eV)	α	C
30	6740	1.08	522
50	2150	1.07	98.4
100	1030	1.03	40.9
200	700	1.02	24.1
500	540	1.00	15.6
1000	487	0.994	12.5

and for He II:

E (eV)	W_0 (eV)	α	C
30			
50			
100	688000	1.09	59.6
200	78200	1.06	23.6
500	25400	1.05	14.7
1000	16400	1.05	11.7

The coefficients are linearly interpolated to obtain values for intermediate energies. According to Shull & van Steenberg (1985) the secondary ionisation fractions stay constant above energies much larger than 100 eV. We thus assume them to be constant above 1000 eV. The fits by (Dalgarno et al. 1999) reproduce detailed calculations within 4% accuracy for $x \leq 0.1$.

These contributions are added to Eqs. A.17, A.18, and A.19.

Collisional ionisation cooling rate (Cen 1992)

Rates are given in $[\text{erg cm}^3\text{s}^{-1}]$.

$$\begin{aligned} \text{H I} : \zeta_{\text{H I}}(T) &= 1.27 \times 10^{-21} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-157809.1/T} \\ \text{He I} : \zeta_{\text{He I}}(T) &= 9.38 \times 10^{-22} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-285335.4/T} \\ \text{He II} : \zeta_{\text{He II}}(T) &= 4.95 \times 10^{-22} T^{-1/2} \left[1 + \left(\frac{T}{10^5} \right)^{1/2} \right]^{-1} e^{-631515.0/T} \end{aligned}$$

Recombination cooling rate (Cen 1992)

Rates are given in $[\text{erg cm}^3\text{s}^{-1}]$.

$$\begin{aligned} \text{H II} : \eta_{\text{H II}}(T) &= 8.70 \times 10^{-27} T^{1/2} \left(\frac{T}{10^3}\right)^{-0.2} \left[1 + \left(\frac{T}{10^6}\right)^{0.7}\right]^{-1} \\ \text{He II} : \eta_{\text{He II}}(T) &= 1.55 \times 10^{-26} T^{0.3647} \\ \text{He III} : \eta_{\text{He III}}(T) &= 3.48 \times 10^{-26} T^{1/2} \left(\frac{T}{10^3}\right)^{-0.2} \left[1 + \left(\frac{T}{10^6}\right)^{0.7}\right]^{-1} \end{aligned}$$

Collisional excitation cooling rate (Cen 1992)

Rates are given in $[\text{erg cm}^3\text{s}^{-1}]$, except $\psi_{\text{He I}}$ is $[\text{erg cm}^6\text{s}^{-1}]$ (due to the two electrons in neutral helium). The expressions are for excitations to all levels for H I, to the $n = 2$ level for He II, and to the $n = 2, 3, 4$ triplets (of the $\text{He}^0(2^3\text{S})$ state, which is assumed to be populated by He II recombinations) for He I.

$$\begin{aligned} \text{H I} : \psi_{\text{H I}}(T) &= 7.5 \times 10^{-19} \left[1 + \left(\frac{T}{10^5}\right)^{1/2}\right]^{-1} e^{-118348.0/T} \\ \text{He I} : \psi_{\text{He I}}(T) &= 9.10 \times 10^{-27} T^{-0.1687} \left[1 + \left(\frac{T}{10^5}\right)^{1/2}\right]^{-1} e^{-13179.0/T} \\ \text{He II} : \psi_{\text{He II}}(T) &= 5.54 \times 10^{-17} T^{-0.397} \left[1 + \left(\frac{T}{10^5}\right)^{1/2}\right]^{-1} e^{-473638.0/T} \end{aligned}$$

Bremsstrahlung cooling (Black 1981)

Rate is given in $[\text{erg cm}^{-3}\text{s}^{-1}]$.

$$\beta(T) = 1.42 \times 10^{-27} T^{1/2} [n_{\text{H II}} + n_{\text{He II}} + 4n_{\text{He III}}] n_e$$

Compton cooling/heating through CMB photons (Haiman et al. 1996)

Rate is given in $[\text{erg cm}^{-3}\text{s}^{-1}]$.

$$\varpi(T) = 1.017 \times 10^{-37} T_{\gamma, \text{CMB}}^4 [T - T_{\gamma, \text{CMB}}^4] n_e$$

Coulomb heating of thermal electrons (Dalgarno et al. 1999)

CRASH, pCRASH2 Not included.

xCRASH The photo heating term in Eq. A.21 is modified to account for the energy deposited as heat due to secondary ionisation. This is done by multiplying the photo heating terms in Eq. A.21 with a heating efficiency factor

$$A(x_e, E) = 1 + \frac{\eta_0(E) - 1}{1 + C(E)x_e^{\alpha(E)}}$$

where for a H-He mixed gas the coefficients are

E (eV)	η_0	α	C
30	0.259	0.984	181
50	0.186	0.834	53.4
100	0.148	0.781	25.8
200	0.132	0.743	15.8
500	0.122	0.703	10.1
1000	0.117	0.678	7.95

The coefficients are linearly interpolated to obtain values for intermediate energies. For energies above 1000 eV the heating efficiency is assumed to lose its dependence on the energy, as was discussed above in the secondary ionisation rates section. The fits by (Dalgarno et al. 1999) reproduce detailed calculations of the heating efficiency within 15% accuracy for $x \leq 0.1$.

Bibliography

Additionally to the bibliographic references the page numbers where the respective works are cited in this thesis are also given at the end of each entry. Where applicable, uniform resource locators are also provided for online material.

- Abel, T., Bryan, G. L., & Norman, M. L., *The Formation of the First Star in the Universe*. 2002, *Science*, 295, 93–4
- Abel, T., Norman, M. L., & Madau, P., *Photon-conserving Radiative Transfer around Point Sources in Multidimensional Numerical Cosmology*. 1999, *ApJ*, 523, 66–76, 77, 171
- Abel, T., & Wandelt, B. D., *Adaptive ray tracing for radiative transfer around point sources*. 2002, *MNRAS*, 330, L53–69, 170
- Adelberger, K. L., Shapley, A. E., Steidel, C. C., Pettini, M., Erb, D. K., & Reddy, N. A., *The Connection between Galaxies and Intergalactic Absorption Lines at Redshift $2 < z < 3$* . 2005, *ApJ*, 629, 636–14, 32, 33, 57
- Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M., *Galaxies and Intergalactic Matter at Redshift $z \sim 3$: Overview*. 2003, *ApJ*, 584, 45–9
- Agafonova, I. I., Levshakov, S. A., Reimers, D., Fechner, C., Tytler, D., Simcoe, R. A., & Songaila, A., *Spectral shape of the UV ionizing background and He II absorption at redshifts $1.8 < z < 2.9$* . 2007, *A&A*, 461, 893–25
- Aguirre, A., Hernquist, L., Katz, N., Gardner, J., & Weinberg, D., *Enrichment of the Intergalactic Medium by Radiation Pressure-driven Dust Efflux*. 2001a, *ApJ*, 556, L11–14
- Aguirre, A., Hernquist, L., Schaye, J., Weinberg, D. H., Katz, N., & Gardner, J., *Metal Enrichment of the Intergalactic Medium at $z=3$ by Galactic Winds*. 2001b, *ApJ*, 560, 599–9, 10, 14, 25, 26, 27, 32, 38, 56
- Ahn, K., & Shapiro, P. R., *Does radiative feedback by the first stars promote or prevent second generation star formation?* 2007, *MNRAS*, 375, 881–80
- Alme, H. J., Rodrigue, G. H., & Zimmerman, G. B., *Domain decomposition models for parallel monte carlo transport*. 2001, *The Journal of Supercomputing*, 18, 5–72

Bibliography

- Altay, G., Croft, R. A. C., & Pelupessy, I., *SPHRAY: a smoothed particle hydrodynamics ray tracer for radiative transfer*. 2008, MNRAS, 386, 1931–1970
- Alvarez, M. A., Bromm, V., & Shapiro, P. R., *The H II Region of the First Star*. 2006, ApJ, 639, 621–669, 170
- Amanatides, J., & Woo, A. 1987, in *Eurographics '87* (Elsevier Science Publishers), 3–10–109, 175
- Amanullah, R., Lidman, C., Rubin, D., Aldering, G., Astier, P., Barbary, K., Burns, M. S., Conley, A., Dawson, K. S., Deustua, S. E., Doi, M., Fabbro, S., Faccioli, L., Fakhouri, H. K., Folatelli, G., Fruchter, A. S., Furusawa, H., Garavini, G., Goldhaber, G., Goobar, A., Groom, D. E., Hook, I., Howell, D. A., Kashikawa, N., Kim, A. G., Knop, R. A., Kowalski, M., Linder, E., Meyers, J., Morokuma, T., Nobili, S., Nordin, J., Nugent, P. E., Östman, L., Pain, R., Panagia, N., Perlmutter, S., Raux, J., Ruiz-Lapuente, P., Spadafora, A. L., Strovink, M., Suzuki, N., Wang, L., Wood-Vasey, W. M., Yasuda, N., & Supernova Cosmology Project, T., *Spectra and Hubble Space Telescope Light Curves of Six Type Ia Supernovae at $0.511 < z < 1.12$ and the Union2 Compilation*. 2010, ApJ, 716, 712–720
- Amdahl, G. M. 1967, in *AFIPS '67 (Spring): Proceedings of the April 18-20, 1967, spring joint computer conference* (New York, NY, USA: ACM), 483–485–89
- Aracil, B., Tripp, T. M., Bowen, D. V., Prochaska, J. X., Chen, H., & Frye, B. L., *High-metallicity, photoionized gas in intergalactic large-scale filaments*. 2006, MNRAS, 367, 139–145
- Aubert, D., & Teyssier, R., *A radiative transfer scheme for cosmological reionization based on a local Eddington tensor*. 2008, MNRAS, 387, 295–315, 170
- Baek, S., Di Matteo, P., Semelin, B., Combes, F., & Revaz, Y., *The simulated 21 cm signal during the epoch of reionization: full modeling of the Ly- α pumping*. 2009, A&A, 495, 389–405, 70, 170
- Bahcall, J. N., & Peebles, P. J. E., *Statistical Tests for the Origin of Absorption Lines Observed in Quasi-Stellar Sources*. 1969, ApJ, 156, L7+–L11, 43, 44
- Bajtlik, S., Duncan, R. C., & Ostriker, J. P., *Quasar ionization of Lyman-alpha clouds - The proximity effect, a probe of the ultraviolet background at high redshift*. 1988, ApJ, 327, 570–585, 102, 103, 112, 136, 142, 163, 164
- Barkana, R., & Loeb, A., *In the beginning: the first sources of light and the reionization of the universe*. 2001, Phys. Rep., 349, 125–185
- Becker, G. D., Bolton, J. S., Haehnelt, M. G., & Sargent, W. L. W., *Detection of extended He II reionization in the temperature evolution of the intergalactic medium*. 2011, MNRAS, 410, 1096–1108
- Becker, G. D., Rauch, M., & Sargent, W. L. W., *The Evolution of Optical Depth in the Ly α Forest: Evidence Against Reionization at $z \sim 6$* . 2007, ApJ, 662, 72–85, 140, 141, 142, 143, 144

- Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G., *The clustering evolution of the galaxy distribution*. 2001, MNRAS, 327, 1041–3
- Bi, H., & Davidsen, A. F., *Evolution of Structure in the Intergalactic Medium and the Nature of the Ly α Forest*. 1997, ApJ, 479, 523–8, 13
- Bianchi, S., Cristiani, S., & Kim, T.-S., *The contribution of galaxies to the UV ionising background and the evolution of the Lyman forest*. 2001, A&A, 376, 1–7, 39, 64, 106, 138
- Bianchi, S., Ferrara, A., & Giovanardi, C., *Monte Carlo Simulations of Dusty Spiral Galaxies: Extinction and Polarization Properties*. 1996, ApJ, 465, 127–72
- Black, J. H., *The physical state of primordial intergalactic clouds*. 1981, MNRAS, 197, 553–180
- Bolton, J. S., & Becker, G. D., *Resolving the high redshift Ly α forest in smoothed particle hydrodynamics simulations*. 2009, MNRAS, 398, L26–104
- Bolton, J. S., & Haehnelt, M. G., *The observed ionization rate of the intergalactic medium and the ionizing emissivity at $z \geq 5$: evidence for a photon-starved and extended epoch of reionization*. 2007, MNRAS, 382, 325–136
- Bolton, J. S., Haehnelt, M. G., Viel, M., & Springel, V., *The Lyman α forest opacity and the metagalactic hydrogen ionization rate at $z \sim 2-4$* . 2005, MNRAS, 357, 1178–8, 11, 106, 107, 136, 138, 140, 141
- Bolton, J. S., Oh, S. P., & Furlanetto, S. R., *Photoheating and the fate of hard photons during the reionization of HeII by quasars*. 2009, MNRAS, 395, 736–8
- Bolton, J. S., Viel, M., Kim, T.-S., Haehnelt, M. G., & Carswell, R. F., *Possible evidence for an inverted temperature-density relation in the intergalactic medium from the flux distribution of the Ly α forest*. 2008, MNRAS, 386, 1131–26, 48
- Bromm, V., Coppi, P. S., & Larson, R. B., *The Formation of the First Stars. I. The Primordial Star-forming Cloud*. 2002, ApJ, 564, 23–4
- Bromm, V., Ferrara, A., Coppi, P. S., & Larson, R. B., *The fragmentation of pre-enriched primordial objects*. 2001, MNRAS, 328, 969–27
- Bromm, V., & Larson, R. B., *The First Stars*. 2004, ARA&A, 42, 79–4, 5
- Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F., *The formation of the first stars and galaxies*. 2009, Nature, 459, 49–5
- Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S., *Measurements of the UV background at $4.6 < z < 6.4$ using the quasar proximity effect*. 2010, ArXiv e-prints, arxiv:1011.5850–8, 136, 140, 141
- Cantalupo, S., & Porciani, C., *RADAMESH: Cosmological Radiative Transfer for Adaptive Mesh Refinement Simulations*. 2010, ArXiv e-prints, arxiv:1009.1625–170

Bibliography

- Carlberg, R. G., Yee, H. K. C., Ellingson, E., Morris, S. L., Abraham, R., Gravel, P., Pritchet, C. J., Smecker-Hane, T., Hartwick, F. D. A., Hesser, J. E., Hutchings, J. B., & Oke, J. B., *The Average Mass Profile of Galaxy Clusters*. 1997, ApJ, 485, L13+ 1
- Carroll, S. M., *The Cosmological Constant*. 2001, Living Reviews in Relativity, 4, 1 2
- Carswell, B., Schaye, J., & Kim, T., *The Enrichment History of the Intergalactic Medium: O VI in Ly α Forest Systems at Redshift $z \sim 2$* . 2002, ApJ, 578, 43 17, 25, 60, 66
- Carswell, R. F., Webb, J. K., Baldwin, J. A., & Atwood, B., *High-redshift QSO absorbing clouds and the background ionizing source*. 1987, ApJ, 319, 709 102
- Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A., & Tytler, D., *Observations of the spectra of Q0122-380 and Q1101-264*. 1982, MNRAS, 198, 91 136
- Cen, R., *A hydrodynamic approach to cosmology - Methodology*. 1992, ApJS, 78, 341 178, 179, 180
- , *A Fast, Accurate, and Robust Algorithm for Transferring Radiation in Three-dimensional Space*. 2002, ApJS, 141, 211 169
- Cen, R., & McDonald, P., *Evolution of the Ionizing Radiation Background and Star Formation in the Aftermath of Cosmological Reionization*. 2002, ApJ, 570, 457 7
- Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M., *Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest*. 1994, ApJ, 437, L9 13, 38
- Chen, H., Lanzetta, K. M., & Webb, J. K., *The Origin of C IV Absorption Systems at Redshifts $z < 1$: Discovery of Extended C IV Envelopes around Galaxies*. 2001, ApJ, 556, 158 32, 33, 34
- Chenyang, L., Hong, Z., & Nengchao, W., *Fast n-dimensional hilbert mapping algorithm*. 2008, Computational Science and its Applications, International Conference, 0, 507 76, 98
- Churchill, C. W., Kacprzak, G. G., & Steidel, C. C. 2005, in *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, ed. P. Williams, C.-G. Shu, & B. Menard, 24–41 33
- Ciardi, B., & Ferrara, A., *The First Cosmic Structures and Their Effects*. 2005, Space Sci. Rev., 116, 625 4
- Ciardi, B., Ferrara, A., Governato, F., & Jenkins, A., *Inhomogeneous reionization of the intergalactic medium regulated by radiative and stellar feedbacks*. 2000, MNRAS, 314, 611 5
- Ciardi, B., Ferrara, A., Marri, S., & Raimondo, G., *Cosmological reionization around the first stars: Monte Carlo radiative transfer*. 2001, MNRAS, 324, 381 70, 77, 108, 172
- Ciardi, B., Ferrara, A., & White, S. D. M., *Early reionization by the first galaxies*. 2003, MNRAS, 344, L7 5

- Clowe, D., Gonzalez, A., & Markevitch, M., *Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter*. 2004, ApJ, 604, 596 1
- Cole, S., Hatton, S., Weinberg, D. H., & Frenk, C. S., *Mock 2dF and SDSS galaxy redshift surveys*. 1998, MNRAS, 300, 945 3
- Colgate, S. A., *Supernovae as a standard candle for cosmology*. 1979, ApJ, 232, 404 2
- Cooke, A. J., Espey, B., & Carswell, R. F., *Evolution of the ionizing background at high redshifts*. 1997, MNRAS, 284, 552 136
- Couchman, H. M. P., & Rees, M. J., *Pregalactic evolution in cosmologies with cold dark matter*. 1986, MNRAS, 221, 53 4
- Cowie, L. L., Songaila, A., Kim, T., & Hu, E. M., *The metallicity and internal structure of the Lyman-alpha forest clouds*. 1995, AJ, 109, 1522 9, 14, 25, 38, 56
- Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., & Savaglio, S., *The space distribution of the Lyman alpha clouds in the line of sight to the $z=3.66$ QSO 0055-269*. 1995, MNRAS, 273, 1016 136
- Croom, S. M., Richards, G. T., Shanks, T., Boyle, B. J., Strauss, M. A., Myers, A. D., Nichol, R. C., Pimbblet, K. A., Ross, N. P., Schneider, D. P., Sharp, R. G., & Wake, D. A., *The 2dF-SDSS LRG and QSO survey: the QSO luminosity function at $0.4 < z < 2.6$* . 2009, MNRAS, 399, 1755 8
- da Ângela, J., Shanks, T., Croom, S. M., Weilbacher, P., Brunner, R. J., Couch, W. J., Miller, L., Myers, A. D., Nichol, R. C., Pimbblet, K. A., de Propris, R., Richards, G. T., Ross, N. P., Schneider, D. P., & Wake, D., *The 2dF-SDSS LRG and QSO survey: QSO clustering and the L-z degeneracy*. 2008, MNRAS, 383, 565 136, 145
- Dalgarno, A., Yan, M., & Liu, W., *Electron Energy Deposition in a Gas Mixture of Atomic and Molecular Hydrogen and Helium*. 1999, ApJS, 125, 237 178, 179, 180, 181
- Dall'Aglio, A., & Gnedin, N. Y., *Analysis of methods for detecting the proximity effect in quasar spectra*. 2010, ApJ, 722, 699 123
- Dall'Aglio, A., Wisotzki, L., & Worseck, G., *An unbiased measurement of the UV background and its evolution via the proximity effect in quasar spectra*. 2008a, A&A, 491, 465 8, 11, 102, 106, 110, 118, 119, 127, 129, 131, 136, 140, 146, 155
- , *The line-of-sight proximity effect in individual quasar spectra*. 2008b, A&A, 480, 359 102, 103, 118, 119, 136, 138, 141, 145, 146
- , *The UV background photoionization rate at $2.3 < z < 4.6$ as measured from the Sloan Digital Sky Survey*. 2009, ArXiv e-prints, arxiv:0906.1484 113, 136, 138, 140
- Dautcourt, G. 1976, *Relativistische Astrophysik. Kosmische Hintergrundstrahlung.*, 2nd edn., ed. Dautcourt, G. (Berlin (GDR): Akademie-Verlag) 169

Bibliography

- Davé, R., Hellsten, U., Hernquist, L., Katz, N., & Weinberg, D. H., *Constraining the Metallicity of the Low-Density Ly α Forest Using O VI Absorption*. 1998, ApJ, 509, 661–674, 38
- Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H., *The Low-Redshift Ly α Forest in Cold Dark Matter Cosmologies*. 1999, ApJ, 511, 521–538, 14, 38, 39, 47, 64, 102
- de Bernardis, P., Ade, P. A. R., Bock, J. J., Bond, J. R., Borrill, J., Boscaleri, A., Coble, K., Crill, B. P., De Gasperis, G., Farese, P. C., Ferreira, P. G., Ganga, K., Giacometti, M., Hivon, E., Hristov, V. V., Iacoangeli, A., Jaffe, A. H., Lange, A. E., Martinis, L., Masi, S., Mason, P. V., Mauskopf, P. D., Melchiorri, A., Miglio, L., Montroy, T., Netterfield, C. B., Pascale, E., Piacentini, F., Pogosyan, D., Prunet, S., Rao, S., Romeo, G., Ruhl, J. E., Scaramuzzi, F., Sforna, D., & Vittorio, N., *A flat Universe from high-resolution maps of the cosmic microwave background radiation*. 2000, Nature, 404, 955–962, 2
- de Blok, W. J. G., & Bosma, A., *High-resolution rotation curves of low surface brightness galaxies*. 2002, A&A, 385, 816–821, 1
- Dessauges-Zavadsky, M., Péroux, C., Kim, T., D’Odorico, S., & McMahon, R. G., *A homogeneous sample of sub-damped Lyman α systems - I. Construction of the sample and chemical abundance measurements*. 2003, MNRAS, 345, 447–458, 25
- D’Odorico, V., Bruscoli, M., Saitta, F., Fontanot, F., Viel, M., Cristiani, S., & Monaco, P., *The quasar proximity effect at redshift $\langle z \rangle \simeq 2.6$ with the From Lines to Overdensities approach*. 2008, MNRAS, 389, 1727–1742, 136, 156, 159
- D’Odorico, V., Calura, F., Cristiani, S., & Viel, M., *The rise of the CIV mass density at $z < 2.5$* . 2010, MNRAS, 401, 2715–2724, 10
- D’Odorico, V., Viel, M., Saitta, F., Cristiani, S., Bianchi, S., Boyle, B., Lopez, S., Maza, J., & Outram, P., *Tomography of the intergalactic medium with Ly α forests in close QSO pairs*. 2006, MNRAS, 372, 1333–1343, 20
- Ellison, S. L., Songaila, A., Schaye, J., & Pettini, M., *The Enrichment History of the Intergalactic Medium-Measuring the C IV/H I Ratio in the Ly α Forest*. 2000, AJ, 120, 1175–1184, 9, 14, 15, 27, 29, 64
- Evrard, A. E., Metzler, C. A., & Navarro, J. F., *Mass Estimates of X-Ray Clusters*. 1996, ApJ, 469, 494–501, 1
- Fan, X., Carilli, C. L., & Keating, B., *Observational Constraints on Cosmic Reionization*. 2006a, ARA&A, 44, 415–444, 7
- Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., & Rix, H., *Evolution of the Ionizing Background and the Epoch of Reionization from the Spectra of $z \sim 6$ Quasars*. 2002, AJ, 123, 1247–1254, 7
- Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., & Fukugita, M., *Constraining the Evolution of*

- the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars.* 2006b, AJ, 132, 117–7
- Fardal, M. A., Giroux, M. L., & Shull, J. M., *The High-Redshift He II Gunn-Peterson Effect: Implications and Future Prospects.* 1998, AJ, 115, 2206–102
- Faucher-Giguère, C., Lidz, A., Hernquist, L., & Zaldarriaga, M., *Evolution of the Intergalactic Opacity: Implications for the Ionizing Background, Cosmic Star Formation, and Quasar Activity.* 2008a, ApJ, 688, 85–7, 8, 11, 136
- Faucher-Giguère, C., Lidz, A., Zaldarriaga, M., & Hernquist, L., *A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization.* 2009, ApJ, 703, 1416–7
- Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L., *The Line-of-Sight Proximity Effect and the Mass of Quasar Host Halos.* 2008b, ApJ, 673, 39–11, 102, 121, 130, 136, 137, 150
- Fechner, C., Baade, R., & Reimers, D., *The associated system of HE 2347-4342.* 2004, A&A, 418, 857–41, 119
- Fechner, C., Reimers, D., Kriss, G. A., Baade, R., Blair, W. P., Giroux, M. L., Green, R. F., Moos, H. W., Morton, D. C., Scott, J. E., Shull, J. M., Simcoe, R., Songaila, A., & Zheng, W., *The UV spectrum of HS 1700+6416. II. FUSE observations of the HeII Lyman alpha forest.* 2006, A&A, 455, 91–109
- Forero-Romero, J. E., Yepes, G., Gottlöber, S., Knollmann, S. R., Khalatyan, A., Cuesta, A. J., & Prada, F., *Simulated versus observed UV emission at high redshift: a hint for a clumpy interstellar medium?* 2010, MNRAS, 403, L31–94
- Furlanetto, S. R., *Fluctuations in the Ionizing Background During and After Helium Reionization.* 2009, ApJ, 703, 702–7
- Furlanetto, S. R., & Dixon, K. L., *Large-Scale Fluctuations in the He II Ly α Forest and He II Reionization.* 2010, ApJ, 714, 355–8
- Giallongo, E., Cristiani, S., D’Odorico, S., Fontana, A., & Savaglio, S., *The Proximity Effect, the UV Background, and the Statistics of the LY alpha Lines at High Resolution.* 1996, ApJ, 466, 46–47, 136
- Glover, S. C. O., & Brand, P. W. J. L., *On the photodissociation of H₂ by the first stars.* 2001, MNRAS, 321, 385–4
- Gnedin, N. Y., *Metal enrichment of the intergalactic medium.* 1998a, MNRAS, 294, 407–9, 14, 25, 26, 32
- , *Probing the Universe with the Ly α forest - II. The column density distribution.* 1998b, MNRAS, 299, 392–8
- , *Cosmological Reionization by Stellar Sources.* 2000, ApJ, 535, 530–5, 7

Bibliography

- Gnedin, N. Y., & Abel, T., *Multi-dimensional cosmological radiative transfer with a Variable Eddington Tensor formalism*. 2001, *New Astronomy*, 6, 437–69, 77, 102, 170
- Gnedin, N. Y., & Hamilton, A. J. S., *Matter power spectrum from the Lyman-alpha forest: myth or reality?* 2002, *MNRAS*, 334, 107–9
- Gnedin, N. Y., Kravtsov, A. V., & Chen, H., *Escape of Ionizing Radiation from High-Redshift Galaxies*. 2008, *ApJ*, 672, 765–6
- Gnedin, N. Y., & Ostriker, J. P., *Reionization of the Universe and the Early Production of Metals*. 1997, *ApJ*, 486, 581–14, 32, 169
- Gnedin, N. Y., Tassis, K., & Kravtsov, A. V., *Modeling Molecular Hydrogen and Star Formation in Cosmological Simulations*. 2009, *ApJ*, 697, 55–170
- Gottlöber, S., Hoffman, Y., & Yepes, G. 2010, in *Proceedings of "High Performance Computing in Science and Engineering"*, ed. S. Wagner, M. Steinmetz, A. Bode, & M. Müller, Vol. arxiv:1005.2687 (Springer-Verlag), 309–3, 137
- Guimarães, R., Petitjean, P., Rollinde, E., de Carvalho, R. R., Djorgovski, S. G., Srianand, R., Aghaee, A., & Castro, S., *Evidence for overdensity around $z \leq 4$ quasars from the proximity effect*. 2007, *MNRAS*, 377, 657–102, 110, 136
- Gunn, J. E., & Peterson, B. A., *On the Density of Neutral Hydrogen in Intergalactic Space*. 1965, *ApJ*, 142, 1633–6, 103
- Haardt, F., & Madau, P., *Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background*. 1996, *ApJ*, 461, 20–7, 102, 137, 145
- Haardt, F., & Madau, P. 2001, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, ed. D. M. Neumann & J. T. V. Tran 102, 106, 107, 138, 140, 141
- , *New Population Synthesis Models of the Cosmic UV Background*. 2011, ArXiv e-prints, arxiv:1103.5226–136
- Haehnelt, M. G., Rauch, M., & Steinmetz, M., *Non-equilibrium effects on line-of-sight size estimates of QSO absorption systems*. 1996a, *MNRAS*, 283, 1055–15
- Haehnelt, M. G., Steinmetz, M., & Rauch, M., *C IV Absorption from Galaxies in the Process of Formation*. 1996b, *ApJ*, 465, L95+–14
- , *Damped Ly α Absorber at High Redshift: Large Disks or Galactic Building Blocks?* 1998, *ApJ*, 495, 647–9
- Haiman, Z., & Loeb, A., *Signatures of Stellar Reionization of the Universe*. 1997, *ApJ*, 483, 21–4, 14
- Haiman, Z., Rees, M. J., & Loeb, A., *Destruction of Molecular Hydrogen during Cosmological Reionization*. 1997, *ApJ*, 476, 458–4
- Haiman, Z., Thoul, A. A., & Loeb, A., *Cosmological Formation of Low-Mass Objects*. 1996, *ApJ*, 464, 523–180

- Hanany, S., Ade, P., Balbi, A., Bock, J., Borrill, J., Boscaleri, A., de Bernardis, P., Ferreira, P. G., Hristov, V. V., Jaffe, A. H., Lange, A. E., Lee, A. T., Mauskopf, P. D., Netterfield, C. B., Oh, S., Pascale, E., Rabii, B., Richards, P. L., Smoot, G. F., Stompor, R., Winant, C. D., & Wu, J. H. P., *MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10^{-5} °*. 2000, ApJ, 545, L5 2
- Hasegawa, K., & Umemura, M., *START: smoothed particle hydrodynamics with tree-based accelerated radiative transfer*. 2010, MNRAS, 407, 2632 69, 170
- Hawkins, E., Maddox, S., Cole, S., Lahav, O., Madgwick, D. S., Norberg, P., Peacock, J. A., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Jones, B., Lewis, I., Lumsden, S., Percival, W., Peterson, B. A., Sutherland, W., & Taylor, K., *The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe*. 2003, MNRAS, 346, 78 3
- Heckman, T. M., Lehnert, M. D., Strickland, D. K., & Armus, L., *Absorption-Line Probes of Gas and Dust in Galactic Superwinds*. 2000, ApJS, 129, 493 14
- Heger, A., & Woosley, S. E., *The Nucleosynthetic Signature of Population III*. 2002, ApJ, 567, 532 5
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J., *The Lyman-Alpha Forest in the Cold Dark Matter Model*. 1996, ApJ, 457, L51+ 8
- Hinshaw, G., Weiland, J. L., Hill, R. S., Odegard, N., Larson, D., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R., Jarosik, N., Komatsu, E., Nolte, M. R., Page, L., Spergel, D. N., Wollack, E., Halpern, M., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., & Wright, E. L., *Five-Year Wilkinson Microwave Anisotropy Probe Observations: Data Processing, Sky Maps, and Basic Results*. 2009, ApJS, 180, 225 15, 137
- Hoefl, M., Yepes, G., Gottlöber, S., & Springel, V., *Dwarf galaxies in voids: suppressing star formation with photoheating*. 2006, MNRAS, 371, 401 102
- Hu, E. M., Kim, T., Cowie, L. L., Songaila, A., & Rauch, M., *The Distribution of Column Densities and B Values in the Lyman-Alpha Forest*. 1995, AJ, 110, 1526 45, 49, 50
- Hui, L., & Gnedin, N. Y., *Equation of state of the photoionized intergalactic medium*. 1997, MNRAS, 292, 27 14, 26, 48, 105, 135, 138
- Hui, L., Gnedin, N. Y., & Zhang, Y., *The Statistics of Density Peaks and the Column Density Distribution of the Ly alpha Forest*. 1997, ApJ, 486, 599 8, 9, 103, 104, 105, 109, 135, 138
- Iliev, I., Ciardi, B., Alvarez, M., Maselli, A., Ferrara, A., Gnedin, N., Mellema, G., Nakamoto, T., Norman, M., Razoumov, A., Rijkhorst, E.-J., Ritzerveld, J., Shapiro, P., Susa, H., Umemura, M., & Whalen, D., *Cosmological radiative transfer codes comparison project - i. the static density field tests*. 2006, Mon. Not. R. Astron. Soc., 371, 1057 70, 78, 82, 86, 97, 169

Bibliography

- Iliev, I. T., Mellema, G., Pen, U., Merz, H., Shapiro, P. R., & Alvarez, M. A., *Simulating cosmic reionization at large scales - I. The geometry of reionization*. 2006, MNRAS, 369, 1625–5, 70
- Iliev, I. T., Mellema, G., Shapiro, P. R., & Pen, U.-L., *Self-regulated reionization*. 2007, MNRAS, 376, 534–103
- Iliev, I. T., Whalen, D., Mellema, G., Ahn, K., Baek, S., Gnedin, N. Y., Kravtsov, A. V., Norman, M., Raicevic, M., Reynolds, D. R., Sato, D., Shapiro, P. R., Semelin, B., Smidt, J., Susa, H., Theuns, T., & Umemura, M., *Cosmological radiative transfer comparison project - II. The radiation-hydrodynamic tests*. 2009, MNRAS, 400, 1283–70, 169
- Iwata, I., Inoue, A. K., Matsuda, Y., Furusawa, H., Hayashino, T., Kousai, K., Akiyama, M., Yamada, T., Burgarella, D., & Deharveng, J., *Detections of Lyman Continuum from Star-Forming Galaxies at $z \sim 3$ through Subaru/Suprime-Cam Narrow-Band Imaging*. 2009, ApJ, 692, 1287–38
- Janknecht, E., Reimers, D., Lopez, S., & Tytler, D., *The evolution of Lyman α absorbers in the redshift range $0.5 < z < 1.9$* . 2006, A&A, 458, 427–45, 47, 48
- Jarosik, N., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R., Halpern, M., Hill, R. S., Hinshaw, G., Kogut, A., Komatsu, E., Larson, D., Limon, M., Meyer, S. S., Nolte, M. R., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results*. 2011, ApJS, 192, 14–2, 3, 6, 38
- Jeans, J. H. 1928, *Astronomy and cosmogony*, ed. Jeans, J. H. 3
- Jena, T., Norman, M. L., Tytler, D., Kirkman, D., Suzuki, N., Chapman, A., Melis, C., Paschos, P., O’Shea, B., So, G., Lubin, D., Lin, W., Reimers, D., Janknecht, E., & Fechner, C., *A concordance model of the Lyman α forest at $z = 1.95$* . 2005, MNRAS, 361, 70–136
- Jonsson, P., *SUNRISE: polychromatic dust radiative transfer in arbitrary geometries*. 2006, MNRAS, 372, 2–72
- Juvela, M., *Efficient Monte Carlo methods for continuum radiative transfer*. 2005, A&A, 440, 531–72
- Kim, T., Carswell, R. F., Cristiani, S., D’Odorico, S., & Giallongo, E., *The physical properties of the Ly α forest at $z > 1.5$* . 2002, MNRAS, 335, 555–38, 39, 40, 42, 45, 46, 47
- Kim, T., Hu, E. M., Cowie, L. L., & Songaila, A., *The Redshift Evolution of the Ly alpha Forest*. 1997, AJ, 114, 1–39, 42, 45, 49
- Kim, T., Viel, M., Haehnelt, M. G., Carswell, R. F., & Cristiani, S., *The power spectrum of the flux distribution in the Lyman α forest of a large sample of UVES QSO absorption spectra (LUQAS)*. 2004, MNRAS, 347, 355–15, 40, 41

- Kim, T.-S., Bolton, J. S., Viel, M., Haehnelt, M. G., & Carswell, R. F., *An improved measurement of the flux distribution of the Ly α forest in QSO absorption spectra: the effect of continuum fitting, metal contamination and noise properties*. 2007, MNRAS, 382, 1657–17, 40, 106, 107, 119, 138, 140, 141, 142
- Kim, T.-S., Cristiani, S., & D’Odorico, S., *The Ly α forest at $1.5 < z < 4$* . 2001, A&A, 373, 757–9, 39, 119, 146
- Kim, Y., & Croft, R. A. C., *Constraining quasar host halo masses with the strength of nearby Ly α forest absorption*. 2008, MNRAS, 387, 377–102
- Kirkman, D., & Tytler, D., *Intrinsic Properties of the $\langle z \rangle = 2.7$ Ly α Forest from Keck Spectra of Quasar HS 1946+7658*. 1997, ApJ, 484, 672–45
- , *The transverse proximity effect in the $z \sim 2$ Lyman α forest suggests quasi-stellar object episodic lifetimes of ~ 1 Myr*. 2008, MNRAS, 391, 1457–102
- Kirkman, D., Tytler, D., Suzuki, N., Melis, C., Hollywood, S., James, K., So, G., Lubin, D., Jena, T., Norman, M. L., & Paschos, P., *The HI opacity of the intergalactic medium at redshifts $1.6 < z < 3.2$* . 2005, MNRAS, 360, 1373–136
- Klypin, A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M., *Galaxies in N-Body Simulations: Overcoming the Overmerging Problem*. 1999, ApJ, 516, 530–137
- Knollmann, S. R., & Knebe, A., *AHF: Amiga’s Halo Finder*. 2009, ApJS, 182, 608–76
- Kohler, K., Gnedin, N. Y., & Hamilton, A. J. S., *Large-Scale Simulations of Reionization*. 2007, ApJ, 657, 15–103
- Kolb, E. W., & Turner, M. S. 1990, *The early universe*, ed. Kolb, E. W. & Turner, M. S. 2
- Kujat, J., Linn, A. M., Scherrer, R. J., & Weinberg, D. H., *Prospects for Determining the Equation of State of the Dark Energy: What Can Be Learned from Multiple Observables?* 2002, ApJ, 572, 1–8
- Larson, R. B. 2003, in *Astronomical Society of the Pacific Conference Series*, Vol. 287, *Galactic Star Formation Across the Stellar Mass Spectrum*, ed. J. M. De Buizer & N. S. van der Bliek, 65–80–5
- Lee, K., Cen, R., Gott, III, J. R., & Trac, H., *The Topology of Cosmological Reionization*. 2008, ApJ, 675, 8–5
- Lehner, N., Savage, B. D., Richter, P., Sembach, K. R., Tripp, T. M., & Wakker, B. P., *Physical Properties, Baryon Content, and Evolution of the Ly α Forest: New Insights from High-Resolution Observations at $z < \sim 0.4$* . 2007, ApJ, 658, 680–45
- Levshakov, S. A., Agafonova, I. I., Centurión, M., & Molaro, P., *Extremely metal-poor Lyman limit system at $z_{abs} = 2.917$ toward the quasar HE 0940-1050*. 2003a, A&A, 397, 851–25

Bibliography

- Levshakov, S. A., Agafonova, I. I., D'Odorico, S., Wolfe, A. M., & Dessauges-Zavadsky, M., *Metal Abundances and Kinematics of Quasar Absorbers. II. Absorption Systems toward Q0347-3819 and APM BR J0307-4945*. 2003b, ApJ, 582, 596–25
- Lidz, A., Faucher-Giguère, C., Dall'Aglio, A., McQuinn, M., Fechner, C., Zaldarriaga, M., Hernquist, L., & Dutta, S., *A Measurement of Small-scale Structure in the $2.2 \leq z \leq 4.2$ Ly α Forest*. 2010, ApJ, 718, 199–104, 138, 140, 141, 142
- Lidz, A., Hui, L., Zaldarriaga, M., & Scoccimarro, R., *How Neutral Is the Intergalactic Medium at $z \sim 6$?* 2002, ApJ, 579, 491–7
- Liske, J., & Williger, G. M., *The proximity effect in a close group of QSOs*. 2001, MNRAS, 328, 653–11, 102, 112, 136, 145, 146
- Loeb, A., *First Light*. 2006, ArXiv Astrophysics e-prints, arxiv:0603.360–3
- Loeb, A., & Barkana, R., *The Reionization of the Universe by the First Stars and Quasars*. 2001, ARA&A, 39, 19–2, 4
- Loeb, A., & Eisenstein, D. J., *Probing Early Clustering with LY alpha Absorption Lines beyond the Quasar Redshift*. 1995, ApJ, 448, 17–102, 136
- Lu, L., *The carbon abundance of the Lyman-alpha clouds*. 1991, ApJ, 379, 99–9
- Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M., *The Lyman-Alpha Forest at Z approximately 4: Keck HIRES Observations of Q0000-26*. 1996, ApJ, 472, 509–45, 102
- Lu, L., Wolfe, A. M., & Turnshek, D. A., *The redshift distribution of Ly-alpha clouds and the proximity effect*. 1991, ApJ, 367, 19–102, 136
- Lynds, R., *The Absorption-Line Spectrum of 4c 05.34*. 1971, ApJ, 164, L73+–8
- Madau, P., *Radiative transfer in a clumpy universe: The colors of high-redshift galaxies*. 1995, ApJ, 441, 18–7
- Madau, P., Ferrara, A., & Rees, M. J., *Early Metal Enrichment of the Intergalactic Medium by Pregalactic Outflows*. 2001, ApJ, 555, 92–9, 14
- Madau, P., Haardt, F., & Rees, M. J., *Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources*. 1999, ApJ, 514, 648–7
- Madau, P., Pozzetti, L., & Dickinson, M., *The Star Formation History of Field Galaxies*. 1998, ApJ, 498, 106–53
- Marakis, J., Chamico, J., Brenner, G., & Durst, F., *Parallel ray tracing for radiative heat transfer: Application in a distributed computing environment*. 2001, International Journal of Numerical Methods for Heat & Fluid Flow, 11, 663–72
- Martin, C. L., *Mapping Large-Scale Gaseous Outflows in Ultraluminous Galaxies with Keck II ESI Spectra: Variations in Outflow Velocity with Galactic Mass*. 2005, ApJ, 621, 227–14

- , *Mapping Large-Scale Gaseous Outflows in Ultraluminous Infrared Galaxies with Keck II ESI Spectra: Spatial Extent of the Outflow*. 2006, ApJ, 647, 222–14
- Mascagni, M., & Srinivasan, A., *Algorithm 806: Sprng: a scalable library for pseudorandom number generation*. 2000, ACM Trans. Math. Softw., 26, 436–78
- Maselli, A., Ciardi, B., & Kanekar, A., *CRASH2: coloured packets and other updates*. 2009, MNRAS, 393, 171–11, 70, 71, 86, 97, 108, 170, 172
- Maselli, A., & Ferrara, A., *Radiative transfer effects on the Ly α forest*. 2005, MNRAS, 364, 1429–7, 70, 71, 103, 109, 110, 172
- Maselli, A., Ferrara, A., & Ciardi, B., *CRASH: a radiative transfer scheme*. 2003, MNRAS, 345, 379–11, 70, 71, 78, 97, 102, 108, 109, 111, 172
- Maselli, A., Gallerani, S., Ferrara, A., & Choudhury, T. R., *On the size of HII regions around high-redshift quasars*. 2007, MNRAS, 376, L34–103
- Mather, J. C., Cheng, E. S., Cottingham, D. A., Eplee, Jr., R. E., Fixsen, D. J., Hewagama, T., Isaacman, R. B., Jensen, K. A., Meyer, S. S., Noerdlinger, P. D., Read, S. M., Rosen, L. P., Shafer, R. A., Wright, E. L., Bennett, C. L., Boggess, N. W., Hauser, M. G., Kelsall, T., Moseley, Jr., S. H., Silverberg, R. F., Smoot, G. F., Weiss, R., & Wilkinson, D. T., *Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument*. 1994, ApJ, 420, 439–2
- McDonald, P., & Miralda-Escudé, J., *The Ly α Forest Flux Distribution at $Z \sim 5.2$ and the Evolution of the Ionizing Background*. 2001, ApJ, 549, L11–136
- McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., & Cen, R., *A Measurement of the Temperature-Density Relation in the Intergalactic Medium Using a New Ly α Absorption-Line Fitting Method*. 2001, ApJ, 562, 52–104
- McQuinn, M., Lidz, A., Zahn, O., Dutta, S., Hernquist, L., & Zaldarriaga, M., *The morphology of HII regions during reionization*. 2007, MNRAS, 377, 1043–5, 70
- McQuinn, M., Lidz, A., Zaldarriaga, M., Hernquist, L., Hopkins, P. F., Dutta, S., & Faucher-Giguère, C., *He II Reionization and its Effect on the Intergalactic Medium*. 2009, ApJ, 694, 842–8, 112
- Meiksin, A., Tittley, E. R., & Brown, C. K., *Helium reionization and the thermal proximity effect*. 2010, MNRAS, 401, 77–112
- Meiksin, A., & White, M., *Particle-mesh simulations of the Ly α forest*. 2001, MNRAS, 324, 141–137
- , *Constraints on the ultraviolet metagalactic emissivity using the Ly α forest*. 2003, MNRAS, 342, 1205–136
- , *The effects of ultraviolet background correlations on Ly α forest flux statistics*. 2004, MNRAS, 350, 1107–7

Bibliography

- Meiksin, A. A., *The physics of the intergalactic medium*. 2009, *Reviews of Modern Physics*, 81, 1405–135
- Mellema, G., Iliev, I. T., Alvarez, M. A., & Shapiro, P. R., *C²-ray: A new method for photon-conserving transport of ionizing radiation*. 2006, *New Astronomy*, 11, 374–69, 70, 102, 170
- Merritt, D., *The distribution of dark matter in the coma cluster*. 1987, *ApJ*, 313, 121–1
- Mesinger, A., & Furlanetto, S., *The inhomogeneous ionizing background following reionization*. 2009, *MNRAS*, 400, 1461–7
- Mihalas, D., & Weibel Mihalas, B. 1984, *Foundations of radiation hydrodynamics*, ed. Mihalas, D. & Weibel Mihalas, B. 72, 167
- Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M., *The Ly alpha Forest from Gravitational Collapse in the Cold Dark Matter + Λ Model*. 1996, *ApJ*, 471, 582–8, 53, 135
- Miralda-Escudé, J., Haehnelt, M., & Rees, M. J., *Reionization of the Inhomogeneous Universe*. 2000, *ApJ*, 530, 1–5, 7, 8
- Mo, H., van den Bosch, F. C., & White, S. 2010, *Galaxy Formation and Evolution*, ed. Mo, H., van den Bosch, F. C., & White, S. 4
- Murdoch, H. S., Hunstead, R. W., Pettini, M., & Blades, J. C., *Absorption spectrum of the $Z = 3.78$ QSO 2000-330. II - The redshift and equivalent width distributions of primordial hydrogen clouds*. 1986, *ApJ*, 309, 19–136
- Nakamoto, T., Umemura, M., & Susa, H., *The effects of radiative transfer on the reionization of an inhomogeneous universe*. 2001, *MNRAS*, 321, 593–103, 170
- Norris, J., Peterson, B. A., & Hartwick, F. D. A., *The oxygen abundance in the clouds producing the Ly-alpha-Ly-beta absorption systems in the quasars 4C 5.34 (0805+046) and OQ 172 (1442+101)*. 1983, *ApJ*, 273, 450–9
- O'Meara, J. M., Prochaska, J. X., Burles, S., Prochter, G., Bernstein, R. A., & Burgess, K. M., *The Keck+Magellan Survey for Lyman Limit Absorption. I. The Frequency Distribution of Super Lyman Limit Systems*. 2007, *ApJ*, 656, 666–48, 50, 53, 54
- Oppenheimer, B. D., & Davé, R., *Cosmological simulations of intergalactic medium enrichment from galactic outflows*. 2006, *MNRAS*, 373, 1265–9, 14, 25, 38, 56
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of gaseous nebulae and active galactic nuclei*, ed. Osterbrock, D. E. & Ferland, G. J. 8, 177
- Ostriker, J. P., & Gnedin, N. Y., *Reheating of the Universe and Population III*. 1996, *ApJ*, 472, L63+–4, 14
- Paardekooper, J., Kruip, C. J. H., & Icke, V., *SimpleX2: radiative transfer on an unstructured, dynamic grid*. 2010, *A&A*, 515, A79+–69, 170

- Page, L., Hinshaw, G., Komatsu, E., Nolta, M. R., Spergel, D. N., Bennett, C. L., Barnes, C., Bean, R., Doré, O., Dunkley, J., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Peiris, H. V., Tucker, G. S., Verde, L., Weiland, J. L., Wollack, E., & Wright, E. L., *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis*. 2007, ApJS, 170, 335 4, 6
- Partl, A. M., *Radiative transfer effects on the Ly α forest*. 2007, University of Vienna, Master thesis: <http://textfeld.ac.at/pdf/991.pdf> 109
- Partl, A. M., Dall’Aglio, A., Müller, V., & Hensler, G., *Cosmological radiative transfer for the line-of-sight proximity effect*. 2010, A&A, 524, A85+ 73, 77, 136, 137, 158
- Pawlik, A. H., & Schaye, J., *TRAPHIC - radiative transfer for smoothed particle hydrodynamics simulations*. 2008, MNRAS, 389, 651 69, 102, 168
- , *Multi-frequency, thermally coupled radiative transfer with TRAPHIC: Method and tests*. 2010, ArXiv e-prints, arxiv:1008.1071 170
- Peacock, J. A. 1999, *Cosmological Physics*, ed. Peacock, J. A. 3, 4
- Peebles, P. J. E. 1993, *Principles of physical cosmology*, ed. Peebles, P. J. E. 2, 4
- Penton, S. V., Shull, J. M., & Stocke, J. T., *The Local Ly α Forest. II. Distribution of H I Absorbers, Doppler Widths, and Baryon Content*. 2000, ApJ, 544, 150 45
- Penzias, A. A., & Wilson, R. W., *A Measurement of Excess Antenna Temperature at 4080 Mc/s*. 1965, ApJ, 142, 419 2
- Péroux, C., Dessauges-Zavadsky, M., D’Odorico, S., Kim, T., & McMahon, R. G., *A homogeneous sample of sub-damped Lyman systems - IV. Global metallicity evolution*. 2007, MNRAS, 382, 177 25
- Petitjean, P., Mückel, J. P., & Kates, R. E., *The LY alpha forest at low redshift: Tracing the dark matter filaments*. 1995, A&A, 295, L9 8, 13, 104, 135, 137
- Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F., & Lanzetta, K., *Evidence for structure in the H I column density distribution of QSO absorbers*. 1993, MNRAS, 262, 499 10, 48, 49, 50, 66
- Petkova, M., & Springel, V., *An implementation of radiative transfer in the cosmological simulation code GADGET*. 2009, MNRAS, 396, 1383 69, 168, 170
- Pettini, M., Kellogg, M., Steidel, C. C., Dickinson, M., Adelberger, K. L., & Giavalisco, M., *Infrared Observations of Nebular Emission Lines from Galaxies at $Z \sim 3$* . 1998, ApJ, 508, 539 14
- Pieri, M. M., & Haehnelt, M. G., *Pixel correlation searches for OVI in the Lyman α forest and the volume filling factor of metals in the intergalactic medium at $z \sim 2-3.5$* . 2004, MNRAS, 347, 985 30
- Pierleoni, M., Maselli, A., & Ciardi, B., *CRASH α : coupling continuum and line radiative transfer*. 2009, MNRAS, 393, 872 167

Bibliography

- Prada, F., Klypin, A. A., Simonneau, E., Betancort-Rijo, J., Patiri, S., Gottlöber, S., & Sanchez-Conde, M. A., *How Far Do They Go? The Outer Structure of Galactic Dark Matter Halos*. 2006, ApJ, 645, 1001–1050
- Qiu, J., Shu, C., Liu, J., & Fang, L., *A WENO algorithm for the growth of ionized regions at the reionization epoch*. 2008, New Astronomy, 13, 1–169
- Qiu, J.-M., Feng, L.-L., Shu, C.-W., & Fang, L.-Z., *A WENO algorithm of the temperature and ionization profiles around a point source*. 2007, New Astronomy, 12, 398–403
- Rauch, M., *The Lyman Alpha Forest in the Spectra of QSOs*. 1998, ARA&A, 36, 267–308, 9, 10, 101
- Rauch, M., Becker, G. D., Viel, M., Sargent, W. L. W., Smette, A., Simcoe, R. A., Barlow, T. A., & Haehnelt, M. G., *Expansion and Collapse in the Cosmic Web*. 2005, ApJ, 632, 58–20
- Rauch, M., Haehnelt, M. G., & Steinmetz, M., *QSO Metal Absorption Systems at High Redshift and the Signature of Hierarchical Galaxy Formation*. 1997a, ApJ, 481, 601–9, 14, 20, 23
- Rauch, M., Miralda-Escude, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., & Ostriker, J. P., *The Opacity of the Ly α Forest and Implications for Ω_b and the Ionizing Background*. 1997b, ApJ, 489, 7–8, 38, 136
- Razoumov, A. O., & Cardall, C. Y., *Fully threaded transport engine: new method for multi-scale radiative transfer*. 2005, MNRAS, 362, 1413–102, 170
- Razoumov, A. O., Norman, M. L., Abel, T., & Scott, D., *Cosmological Hydrogen Reionization with Three-dimensional Radiative Transfer*. 2002, ApJ, 572, 695–69
- Razoumov, A. O., & Sommer-Larsen, J., *Ionizing Radiation from $z = 4$ -10 Galaxies*. 2010, ApJ, 710, 1239–6
- Rees, M. J., & Ostriker, J. P., *Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters*. 1977, MNRAS, 179, 541–4
- Reimers, D., Kohler, S., Wisotzki, L., Groote, D., Rodriguez-Pascual, P., & Wamsteker, W., *Patchy intergalactic He II absorption in HE 2347-4342. II. The possible discovery of the epoch of He-reionization*. 1997, A&A, 327, 890–8
- Richter, P., Savage, B. D., Tripp, T. M., & Sembach, K. R., *FUSE and STIS Observations of the Warm-hot Intergalactic Medium toward PG 1259+593*. 2004, ApJS, 153, 165–45
- Ricotti, M., Gnedin, N. Y., & Shull, J. M., *The Evolution of the Effective Equation of State of the Intergalactic Medium*. 2000, ApJ, 534, 41–106, 138
- , *The Fate of the First Galaxies. II. Effects of Radiative Feedback*. 2002, ApJ, 575, 49–80

- Rijkhorst, E.-J., Plewa, T., Dubey, A., & Mellema, G., *Hybrid characteristics: 3D radiative transfer for parallel adaptive mesh refinement hydrodynamics*. 2006, A&A, 452, 907–70, 102, 170
- Rijkhorst, E.-J., Plewa, T., Dubey, A., & Mellema, G., *Hybrid characteristics: 3d radiative transfer for parallel adaptive mesh refinement hydrodynamics*. 2006, Astronomy & Astrophysics, 452, 907–76
- Ripley, B. D., & Thompson, M., *Regression techniques for the detection of analytical bias*. 1987, Analyst, 112, 377–45
- Ritzerveld, J., *The diffuse nature of Strömgren spheres*. 2005, A&A, 439, L23–84, 115
- Ritzerveld, J., & Icke, V., *Transport on adaptive random lattices*. 2006, Phys. Rev. E, 74, 026704–168
- Ritzerveld, J., Icke, V., & Rijkhorst, E., *Triangulating Radiation: Radiative Transfer on Unstructured Grids*. 2003, ArXiv Astrophysics e-prints, arxiv:0312.301–69
- Rollinde, E., Srikanand, R., Theuns, T., Petitjean, P., & Chand, H., *The density structure around quasars from optical depth statistics*. 2005, MNRAS, 361, 1015–102, 110, 136
- Rubin, V. C., & Ford, Jr., W. K., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*. 1970, ApJ, 159, 379–1
- Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D., *The distribution of Lyman-alpha absorption lines in the spectra of six QSOs - Evidence for an intergalactic origin*. 1980, ApJS, 42, 41–8, 101
- Savaglio, S., Ferguson, H. C., Brown, T. M., Espey, B. R., Sahu, K. C., Baum, S. A., Carollo, C. M., Kaiser, M. E., Stiavelli, M., Williams, R. E., & Wilson, J., *The LYalpha Forest of the Quasar in the Hubble Deep Field South*. 1999, ApJ, 515, L5–45
- Schaye, J., *Model-independent Insights into the Nature of the Ly α Forest and the Distribution of Matter in the Universe*. 2001, ApJ, 559, 507–14, 26, 38, 48, 50
- Schaye, J., Aguirre, A., Kim, T., Theuns, T., Rauch, M., & Sargent, W. L. W., *Metallicity of the Intergalactic Medium Using Pixel Statistics. II. The Distribution of Metals as Traced by C IV*. 2003, ApJ, 596, 768–9, 10, 14, 15, 23, 27, 28, 30, 38, 56, 64, 106, 107, 138, 140, 142
- Schaye, J., Carswell, R. F., & Kim, T., *A large population of metal-rich, compact, intergalactic CIV absorbers - evidence for poor small-scale metal mixing*. 2007, MNRAS, 379, 1169–14, 25, 60, 66
- Schaye, J., Rauch, M., Sargent, W. L. W., & Kim, T., *The Detection of Oxygen in the Low-Density Intergalactic Medium*. 2000a, ApJ, 541, L1–14, 56
- Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W., *The thermal history of the intergalactic medium*. 2000b, MNRAS, 318, 817–38, 106, 107, 138, 140, 141, 142

Bibliography

- Schneider, R., Ferrara, A., Natarajan, P., & Omukai, K., *First Stars, Very Massive Black Holes, and Metals*. 2002, ApJ, 571, 30–4
- Schneider, R., Salvaterra, R., Ferrara, A., & Ciardi, B., *Constraints on the initial mass function of the first stars*. 2006, MNRAS, 369, 825–5
- Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P., *A Uniform Analysis of the Ly α Forest at $z=0-5$. II. Measuring the Mean Intensity of the Extragalactic Ionizing Background Using the Proximity Effect*. 2000, ApJS, 130, 67–102, 110, 112, 136
- Scott, J., Bechtold, J., Morita, M., Dobrzycki, A., & Kulkarni, V. P., *A Uniform Analysis of the Ly α Forest at $z=0-5$. V. The Extragalactic Ionizing Background at Low Redshift*. 2002, ApJ, 571, 665–102
- Scott, J. E., Kriss, G. A., Brotherton, M., Green, R. F., Hutchings, J., Shull, J. M., & Zheng, W., *A Composite Extreme-Ultraviolet QSO Spectrum from FUSE*. 2004, ApJ, 615, 135–110
- Sembach, K. R., Tripp, T. M., Savage, B. D., & Richter, P., *Physical Properties and Baryonic Content of Low-Redshift Intergalactic Ly α and O VI Absorption Line Systems: The PG 1116+215 Sight Line*. 2004, ApJS, 155, 351–45
- Shandarin, S. F., & Zel'dovich, Y. B., *The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium*. 1989, Reviews of Modern Physics, 61, 185–3
- Shapiro, P. R., Iliev, I. T., & Raga, A. C., *Photoevaporation of cosmological minihaloes during reionization*. 2004, MNRAS, 348, 753–103
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L., *Rest-Frame Ultraviolet Spectra of $z \sim 3$ Lyman Break Galaxies*. 2003, ApJ, 588, 65–14
- Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K., *The Direct Detection of Lyman Continuum Emission from Star-forming Galaxies at $z \sim 3$* . 2006, ApJ, 651, 688–38
- Shin, M., Trac, H., & Cen, R., *Cosmological H II Bubble Growth during Reionization*. 2008, ApJ, 681, 756–5
- Shull, J. M., France, K., Danforth, C. W., Smith, B., & Tumlinson, J., *HST/COS Observations of the Quasar HE 2347-4342: Probing the Epoch of He II Patchy Reionization at Redshifts $z = 2.4-2.9$* . 2010, ApJ, 722, 1312–6, 8
- Shull, J. M., Tumlinson, J., Giroux, M. L., Kriss, G. A., & Reimers, D., *The Fluctuating Intergalactic Radiation Field at Redshifts $z = 2.3-2.9$ from He II and H I Absorption toward HE 2347-4342*. 2004, ApJ, 600, 570–8
- Shull, J. M., & van Steenberg, M. E., *X-ray secondary heating and ionization in quasar emission-line clouds*. 1985, ApJ, 298, 268–179

- Siana, B., Teplitz, H. I., Ferguson, H. C., Brown, T. M., Giavalisco, M., Dickinson, M., Chary, R., de Mello, D. F., Conselice, C. J., Bridge, C. R., Gardner, J. P., Colbert, J. W., & Scarlata, C., *A Deep Hubble Space Telescope Search for Escaping Lyman Continuum Flux at $z \sim 1.3$: Evidence for an Evolving Ionizing Emissivity*. 2010, ApJ, 723, 241–38
- Silk, J., *On the fragmentation of cosmic gas clouds. I - The formation of galaxies and the first generation of stars*. 1977, ApJ, 211, 638–4
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M., *The Distribution of Metallicity in the Intergalactic Medium at $z \sim 2.5$: O VI and C IV Absorption in the Spectra of Seven QSOs*. 2004, ApJ, 606, 92–15, 27
- Songaila, A., *The Redshift Evolution of the Metagalactic Ionizing Flux Inferred from Metal Line Ratios in the Lyman Forest*. 1998, AJ, 115, 2184–14, 56
- , *The Minimum Universal Metal Density between Redshifts of 1.5 and 5.5*. 2001, ApJ, 561, L153–9
- , *The Evolution of the Intergalactic Medium Transmission to Redshift 6*. 2004, AJ, 127, 2598–6
- Songaila, A., & Cowie, L. L., *Metal enrichment and Ionization Balance in the Lyman Alpha Forest at $Z = 3$* . 1996, AJ, 112, 335–9, 14, 25
- Songaila, A., Hu, E. M., Cowie, L. L., & McMahon, R. G., *Limits on the Gunn-Peterson Effect at $Z = 5$* . 1999, ApJ, 525, L5–136
- Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L., *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology*. 2007, ApJS, 170, 377–104
- Springel, V., *The cosmological simulation code GADGET-2*. 2005, MNRAS, 364, 1105–76, 104, 137
- Springel, V., & Hernquist, L., *Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation*. 2003, MNRAS, 339, 289–10, 14, 25, 137
- Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F., *Simulations of the formation, evolution and clustering of galaxies and quasars*. 2005, Nature, 435, 629–3
- Srianand, R., & Khare, P., *Analysis of Ly α absorption lines in the vicinity of QSOs*. 1996, MNRAS, 280, 767–136
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M., *Lyman-Break Galaxies at $z > \sim 4$ and the Evolution of the Ultraviolet Luminosity Density at High Redshift*. 1999, ApJ, 519, 1–32

Bibliography

- Steidel, C. C., Erb, D. K., Shapley, A. E., Pettini, M., Reddy, N., Bogosavljević, M., Rudie, G. C., & Rakic, O., *The Structure and Kinematics of the Circumgalactic Medium from Far-ultraviolet Spectra of $z \sim 2-3$ Galaxies*. 2010, ApJ, 717, 289–39, 57, 64
- Storrie-Lombardi, L. J., & Wolfe, A. M., *Surveys for $z > 3$ Damped Ly α Absorption Systems: The Evolution of Neutral Gas*. 2000, ApJ, 543, 552–48, 50, 53, 54
- Strickland, D. K., Heckman, T. M., Colbert, E. J. M., Hoopes, C. G., & Weaver, K. A., *A High Spatial Resolution X-Ray and H α Study of Hot Gas in the Halos of Star-forming Disk Galaxies. II. Quantifying Supernova Feedback*. 2004, ApJ, 606, 829–14
- Strömberg, B., *The Physical State of Interstellar Hydrogen*. 1939, ApJ, 89, 526–5
- Susa, H., & Umemura, M., *Formation of Dwarf Galaxies during the Cosmic Reionization*. 2004, ApJ, 600, 1–170
- , *Secondary Star Formation in a Population III Object*. 2006, ApJ, 645, L93–80
- Syphers, D., Anderson, S. F., Zheng, W., Meiksin, A., Haggard, D., Schneider, D. P., & York, D. G., *Hubble Space Telescope Spectral Observations Near the He II Ly α Break: Implications for He II Reionization*. 2011, ApJ, 726, 111–8
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F., *The Rest-Frame Extreme-Ultraviolet Spectral Properties of Quasi-stellar Objects*. 2002, ApJ, 565, 773–110, 145
- Tepper-García, T., *Voigt profile fitting to quasar absorption lines: an analytic approximation to the Voigt-Hjerting function*. 2006, MNRAS, 369, 2025–8
- Tescari, E., Viel, M., D’Odorico, V., Cristiani, S., Calura, F., Borgani, S., & Tornatore, L., *Cosmic evolution of the CIV in high-resolution hydrodynamic simulations*. 2010, ArXiv e-prints, arxiv:1007.1628–10
- Teyssier, R., *Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES*. 2002, A&A, 385, 337–76
- Theuns, T., Leonard, A., & Efstathiou, G., *The low-redshift evolution of the Lyman-alpha forest*. 1998a, MNRAS, 297, L49–38, 39, 47, 53, 64
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A., *P^3M -SPH simulations of the Ly α forest*. 1998b, MNRAS, 301, 478–38, 104, 136
- Trac, H., & Cen, R., *Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and Initial Results*. 2007, ApJ, 671, 1–5, 70, 103
- Trammell, G. B., Vanden Berk, D. E., Schneider, D. P., Richards, G. T., Hall, P. B., Anderson, S. F., & Brinkmann, J., *The UV Properties of SDSS-Selected Quasars*. 2007, AJ, 133, 1780–110, 121
- Tytler, D., *The redshift distribution of QSO Lyman-alpha absorption systems*. 1987, ApJ, 321, 69–136

- Tytler, D., & Fan, X., *Systematic QSO emission-line velocity shifts and new unbiased redshifts*. 1992, ApJS, 79, 1–41
- Tytler, D., Fan, X., Burles, S., Cottrell, L., Davis, C., Kirkman, D., & Zuo, L. 1995, in *QSO Absorption Lines*, ed. G. Meylan, 289–+ 9, 14
- Tytler, D., Kirkman, D., O’Meara, J. M., Suzuki, N., Orin, A., Lubin, D., Paschos, P., Jena, T., Lin, W., Norman, M. L., & Meiksin, A., *Cosmological Parameters σ_8 , the Baryon Density Ω_b , the Vacuum Energy Density Ω_Λ , the Hubble Constant and the UV Background Intensity from a Calibrated Measurement of H I Ly α Absorption at $z=1.9$* . 2004, ApJ, 617, 1–136
- Tytler, D., O’Meara, J. M., Suzuki, N., & Lubin, D., *Review of Big Bang Nucleosynthesis and Primordial Abundances*. 2000, Physica Scripta Volume T, 85, 12–2
- Tytler, D., Paschos, P., Kirkman, D., Norman, M. L., & Jena, T., *The effect of large-scale power on simulated spectra of the Ly α forest*. 2009, MNRAS, 393, 723–104
- Unsöld, A. 1955, *Physik der Sternatmosphären, MIT besonderer Berücksichtigung der Sonne.*, ed. Unsöld, A. 8
- van Albada, T. S., Bahcall, J. N., Begeman, K., & Sancisi, R., *Distribution of dark matter in the spiral galaxy NGC 3198*. 1985, ApJ, 295, 305–1
- Vanden Berk, D. E., Richards, G. T., Bauer, A., Strauss, M. A., Schneider, D. P., Heckman, T. M., York, D. G., Hall, P. B., Fan, X., Knapp, G. R., Anderson, S. F., & Annis, J., *Composite Quasar Spectra from the Sloan Digital Sky Survey*. 2001, AJ, 122, 549–41
- Venkatesan, A., Tumlinson, J., & Shull, J. M., *Evolving Spectra of Population III Stars: Consequences for Cosmological Reionization*. 2003, ApJ, 584, 621–7
- Verhamme, A., Schaerer, D., & Maselli, A., *3D Ly α radiation transfer. I. Understanding Ly α line profile morphologies*. 2006, A&A, 460, 397–167
- Verner, D. A., & Yakovlev, D. G., *Analytic FITS for partial photoionization cross sections*. 1995, A&AS, 109, 125–177
- Viel, M., Matarrese, S., Mo, H. J., Theuns, T., & Haehnelt, M. G., *Modelling the IGM and the Ly α forest at high redshift from the dark matter distribution*. 2002, MNRAS, 336, 685–104
- Viel, M., Matarrese, S., Theuns, T., Munshi, D., & Wang, Y., *Dark energy effects on the Lyman α forest*. 2003, MNRAS, 340, L47–8
- Volders, L. M. J. S., *Neutral hydrogen in M 33 and M 101*. 1959, Bull. Astron. Inst. Netherlands, 14, 323–1
- von Benda-Beckmann, A. M., D’Onghia, E., Gottlöber, S., Hoefft, M., Khalatyan, A., Klypin, A., & Müller, V., *The fossil phase in the life of a galaxy group*. 2008, MNRAS, 386, 2345–104

Bibliography

- Warren, S. J., Hewett, P. C., & Osmer, P. S., *A wide-field multicolor survey for high-redshift quasars, Z greater than or equal to 2.2. 3: The luminosity function.* 1994, ApJ, 421, 412–8
- Weymann, R. J., Jannuzi, B. T., Lu, L., Bahcall, J. N., Bergeron, J., Boksenberg, A., Hartig, G. F., Kirhakos, S., Sargent, W. L. W., Savage, B. D., Schneider, D. P., Turnshek, D. A., & Wolfe, A. M., *The Hubble Space Telescope Quasar Absorption Line Key Project. XIV. The Evolution of Ly α Absorption Lines in the Redshift Interval $Z = 0-1.5$.* 1998, ApJ, 506, 1–39, 45
- Whalen, D., & Norman, M. L., *A Multistep Algorithm for the Radiation Hydrodynamical Transport of Cosmological Ionization Fronts and Ionized Flows.* 2006, ApJS, 162, 281–70, 170
- White, S. D. M., Navarro, J. F., Evrard, A. E., & Frenk, C. S., *The baryon content of galaxy clusters: a challenge to cosmological orthodoxy.* 1993, Nature, 366, 429–1
- Williger, G. M., Baldwin, J. A., Carswell, R. F., Cooke, A. J., Hazard, C., Irwin, M. J., McMahon, R. G., & Storrie-Lombardi, L. J., *Lyman-alpha absorption in the spectrum of the $Z = 4.5$ QSO BR 1033-0327.* 1994, ApJ, 428, 574–136
- Williger, G. M., Carswell, R. F., Weymann, R. J., Jenkins, E. B., Sembach, K. R., Tripp, T. M., Davé, R., Haberzettl, L., & Heap, S. R., *The low-redshift Ly α forest towards 3C 273.* 2010, MNRAS, 405, 1736–45
- Williger, G. M., Heap, S. R., Weymann, R. J., Davé, R., Ellingson, E., Carswell, R. F., Tripp, T. M., & Jenkins, E. B., *The Low-Redshift Ly α Forest toward PKS 0405-123.* 2006, ApJ, 636, 631–45
- Wise, J. H., & Abel, T., *How Very Massive Metal-Free Stars Start Cosmological Reionization.* 2008, ApJ, 684, 1–4
- Wise, J. H., & Cen, R., *Ionizing Photon Escape Fractions From High-Redshift Dwarf Galaxies.* 2009, ApJ, 693, 984–6
- Wolf, C., Wisotzki, L., Borch, A., Dye, S., Kleinheinrich, M., & Meisenheimer, K., *The evolution of faint AGN between $z \sim 1$ and $z \sim 5$ from the COMBO-17 survey.* 2003, A&A, 408, 499–8
- Wolfe, A. M., Gawiser, E., & Prochaska, J. X., *Damped Ly α Systems.* 2005, ARA&A, 43, 861–9
- Worseck, G., Fechner, C., Wisotzki, L., & Dall’Aglia, A., *The transverse proximity effect in spectral hardness on the line of sight towards HE 2347-4342.* 2007, A&A, 473, 805–119
- Wyithe, J. S. B., & Cen, R., *The Extended Star Formation History of the First Generation of Stars and the Reionization of Cosmic Hydrogen.* 2007, ApJ, 659, 890–5

- Yajima, H., Choi, J., & Nagamine, K., *Escape fraction of ionizing photons from high-redshift galaxies in cosmological SPH simulations*. 2010, ArXiv e-prints, arxiv:1002.3346 6
- Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N., *Simulations of Early Structure Formation: Primordial Gas Clouds*. 2003, ApJ, 592, 645 4
- Zahn, O., Lidz, A., McQuinn, M., Dutta, S., Hernquist, L., Zaldarriaga, M., & Furlanetto, S. R., *Simulations and Analytic Calculations of Bubble Growth during Hydrogen Reionization*. 2007, ApJ, 654, 12 5, 103
- Zel'dovich, Y. B., *Gravitational instability: An approximate theory for large density perturbations*. 1970, A&A, 5, 84 3
- Zhang, Y., Meiksin, A., Anninos, P., & Norman, M. L., *Physical Properties of the Ly alpha Forest in a Cold Dark Matter Cosmology*. 1998, ApJ, 495, 63 104
- Zwicky, F., *Die Rotverschiebung von extragalaktischen Nebeln*. 1933, Helvetica Physica Acta, 6, 110 1
- , *On the Masses of Nebulae and of Clusters of Nebulae*. 1937, ApJ, 86, 217 1

Acknowledgments

This thesis would not have been possible without the support and encouragement of many people. First of all I would like to thank Volker Müller for his patient and motivating supervision, without which this work would have never been possible. I also thank the AIP for giving me the opportunity to work at the institute. I especially enjoyed the stimulating environment and the many helpful discussions with my colleagues Aldo, Anne, Claudio, Jaime, Jochen, Juan-Carlos, Manee, Noam, Pratika, and all the others I missed in this list. Am am especially indebted to Tae-Sun Kim for introducing me to the world of observations. Additionally, I am thankful for discussions and useful advice by Lutz Wisotzki.

Further I thank for the hospitality I was able to enjoy at MPA Garching. Thank you Antonella Maselli, Benedetta Ciardi, and Dominik Schleicher for the wonderful collaboration! I am also indebted to Luca for showing me how to code IBM style. I also thank Andrea Ferrara for being a great host at SISSA Trieste and I am as well grateful for the hospitality received during the stay at CINECA, especially by R. Brunino and C. Gheller. Parts of this work was carried out under the HPC-EUROPA++ project (project number: 211437), with the support of the European Community - Research Infrastructure Action of the FP7 "Coordination and support action" Programme (application 1218).

I am also thankful for the time I was able to enjoy at UAM Madrid. Thank you Steffen and Gustavo for providing all the stimulating discussions and possibilities in Madrid. Alexander, I remember all the good hours we spent in that legendary Sherry bar. I also acknowledge the support by ASTROSIM.

Hanns, I will never forget the encouragement you gave me whenever I was in doubt. The stays with you in Vienna were soothing to the soul. I am as well grateful to Chatchai (aka Pom) for inventing the Cookie paradise, which reminded me of the relaxing times in Thailand. Last but not least I am forever grateful to the support I received from Christina. You were always there for me, especially in times of trouble and doubt. Without you, live would be like a world without a sun that lights up all these wonders around us.

Some simulations have been carried out at the Konrad-Zuse-Zentrum für Informationstechnik in Berlin, Germany using the AstroGrid-D. A.P. acknowledges support in parts by the German Ministry for Education and Research (BMBF) under grant FKZ 05 AC7BAA.