

# **The Stochastic Intergalactic Attenuation and its Impact on High-Redshift Galaxies**

Dissertation  
zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultäten  
der Georg-August-Universität zu Göttingen

vorgelegt von  
**Thorsten Tepper-García**  
aus Mexiko-Stadt, Mexiko.

**Göttingen 2007**

D 7

Referentin: Prof. Dr. Uta Fritze

Korreferent: Prof. Dr. Wolfram Kollatschny

Tag der mündlichen Prüfung: 11. Juli 2007

---

# Zusammenfassung

---

In der vorliegenden Dissertation wird der Einfluss des intergalaktischen, neutralen Wasserstoffs auf die photometrischen Eigenschaften von hoch rotverschobenen Galaxien – auch *intergalaktische Attenuation* genannt – untersucht. Insbesondere wird auf die Unterschiede in den beobachteten Helligkeiten und Farben von Galaxien mit identischen spektralen Energieverteilungen aufgrund der unterschiedlichen physikalischen Eigenschaften der Wasserstoff – oder  $\text{Ly}\alpha$  – Wolken entlang verschiedener Sichtlinien eingegangen. Zuerst wird eine neue Methode entwickelt, mit der Profile verschiedener Absorptionslinien mit wenigem Rechenaufwand und hoher Präzision berechnet werden können. Zunächst wird, basierend auf dieser Methode, ein Modell für die Absorption im intergalaktischem Medium und deren Streuung aufgrund der unterschiedlichen Wolkenpopulationen entlang verschiedener Sichtlinien vorgestellt. Als Input für dieses Modell werden Verteilungsfunktionen für die physikalischen Eigenschaften – Dichte, Temperatur, und Rotverschiebung entlang der Sichtlinie – der  $\text{Ly}\alpha$  Wolken verwendet, wie sie in der Literatur vorliegen. Mit Hilfe der Monte Carlo Methode werden aus diesen Verteilungen zufällige Wolkenpopulationen und somit sogenannten Absorptionsmasken erzeugt, die die Absorption entlang einer einzigen Sichtlinie darstellen. Um die Güte verschiedener Sätze von Verteilungsfunktionen zu bewerten, werden zunächst Modellrechnungen für jeden Satz mit Beobachtungen verglichen. Zu diesem Zweck werden Messungen des kosmischen Flussdekrement herangezogen. Diese ist eine von der Bestimmung der Verteilungsfunktionen unabhängige Grösse, die die Menge an neutralem Wasserstoff entlang einer gegebenen Sichtlinie misst. Um den für diesen Vergleich benutzten Rotverschiebungsbereich zu erweitern, werden Messungen die in der Literatur gefunden werden mit unseren eigenen Messungen vervollständigt. Ein bestimmter Satz von Verteilungsfunktionen, für den die Modellrechnungen mit diesen Messungen hervorragend übereinstimmen, wird dann weiter für die Berechnung der Unterschiede in den Magnituden – auch intergalaktische  $k$ -Korrektur genannt – als Funktion der Rotverschiebung für verschiedene Galaxienspektren in verschiedenen Wellenlängenbereichen benutzt. Darüber hinaus wird der Einfluss der intergalaktischen Absorption auf die Bestimmung von photometrischen Rotverschiebungen durch Anwendung spektraler Energieverteilungen analysiert.



---

# Abstract

---

The effect of the absorption in the intergalactic medium due to neutral hydrogen randomly distributed along the line-of-sight – in the form of Ly $\alpha$  absorbers – on the photometric properties of high-redshift galaxies is analysed. In particular, the differences in the observed colors of galaxies with otherwise identical Spectral Energy Distributions, arising from the difference in the distribution and physical properties – density, temperature, redshift – of the neutral hydrogen along different lines-of-sight is investigated. As a very first step, a new method to synthesize absorption line-profiles – assuming that these are mathematically given by the Voigt-Hjerting function – is developed. This new method is compared to existing methods and within this context it is found to be computationally faster than more accurate methods, and more accurate in the parameter range of interest than faster methods. Using this method, a model that accounts for the intergalactic absorption and its scatter due to the difference in the populations of Ly $\alpha$  absorbers along different random lines-of-sight is presented. This model makes use of a set of input differential distribution functions for the physical properties of the Ly $\alpha$  absorbers, *i.e.* their number density, column density, and Doppler parameter distribution, as obtained from observations and reported in the literature. Following the Monte Carlo approach, absorption masks which account for the intergalactic absorption along single lines-of-sight are constructed. Within this approach, the number and properties of each absorber for each absorption mask are drawn from the corresponding input distributions, and their absorption line-profiles are computed using the method previously introduced. Different sets of input distribution functions are tested by comparing the model predictions against measurements of the cosmic flux decrement, an independent observable that directly relates to the amount of neutral hydrogen along a given line-of-sight. This comparison is done using measurements found in the literature together with our own measurements performed on a sample of quasars obtained from the SDSS data archive. A particular set of input distribution functions is found which satisfactorily reproduces the evolution of the cosmic flux decrement over a large redshift range. Using this set as input for our model, the magnitude changes – or intergalactic  $k$ -correction – as a function of redshift is computed for different input galaxy spectra and in different wavelength ranges. Furthermore, the effect of the stochastic intergalactic absorption on the estimates of photometric redshifts using Spectral Energy Distributions is explored.



---

# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Physics of Neutral Hydrogen</b>	<b>5</b>
2.1	The Lyman series . . . . .	5
2.2	Absorption due to H $\alpha$ Atoms . . . . .	6
2.2.1	The absorption coefficient of H $\alpha$ . . . . .	7
2.2.2	The Curve-of-Growth for H $\alpha$ . . . . .	9
<b>3</b>	<b>An Analytic Approximation to The Voigt-Hjerting Function</b>	<b>13</b>
3.1	Introduction . . . . .	13
3.2	The Voigt-Hjerting Function in the context of H $\alpha$ Absorption Lines . . . . .	15
3.2.1	The Absorption Coefficient . . . . .	15
3.2.2	The Absorption Coefficient of H $\alpha$ at Low Column Densities . . . . .	16
3.2.3	Higher Column Densities and First Order Term . . . . .	17
3.3	The Dawson Function Revisited . . . . .	18
3.3.1	Properties of the Dawson Function . . . . .	18
3.4	The analytic approximation $D_1$ . . . . .	19
3.4.1	The Voigt-Hjerting function to First Order . . . . .	20
3.5	Analysis . . . . .	20
3.5.1	Speed . . . . .	21
3.5.2	Precision . . . . .	21
3.5.3	Modeling of H $\alpha$ absorption profiles . . . . .	21
3.6	Application . . . . .	25
3.7	Summary . . . . .	27
3.8	Acknowledgments . . . . .	28
<b>4</b>	<b>The Intergalactic Medium</b>	<b>29</b>
4.1	Introduction . . . . .	29
4.2	The Background Cosmology . . . . .	30
4.3	The Large-Scale Structure of the Universe . . . . .	32
4.3.1	The $\Lambda$ -Cold Dark Matter ( $\Lambda$ CDM) Model . . . . .	32

4.4	The Ly $\alpha$ Absorbers . . . . .	34
4.4.1	The Ly $\alpha$ Forest . . . . .	35
4.4.2	Lyman Limit Systems . . . . .	35
4.4.3	Damped Ly $\alpha$ Absorbers (DLA) . . . . .	36
4.5	Properties of the Ly $\alpha$ absorbers . . . . .	36
4.5.1	The proper path length . . . . .	36
4.5.2	The Number Density Evolution along the line-of-sight . . . . .	38
4.5.3	Line profile fitting . . . . .	41
4.5.4	The column density distribution function (CDDF) . . . . .	41
4.5.5	The Doppler parameter distribution . . . . .	43
4.5.6	The mean optical depth or effective optical depth . . . . .	44
4.5.7	Cosmic flux decrement . . . . .	45
4.5.8	Structure of the Ly $\alpha$ Forest . . . . .	45
<b>5</b>	<b>Modelling the absorption of light in the intergalactic medium: The Monte Carlo Approach</b>	<b>47</b>
5.1	The path of radiation through the IGM . . . . .	47
5.2	Gambling with the Ly $\alpha$ absorber parameters . . . . .	48
5.2.1	Some Notes on "arbitrarily"-distributed random numbers . . . . .	50
5.3	Numerical simulations of the intergalactic attenuation . . . . .	51
5.4	Cookbook for a Line-of-Sight with a random population of Ly $\alpha$ absorbers . . . . .	52
<b>6</b>	<b>The evolution of <math>D_A</math></b>	<b>55</b>
6.1	Introduction . . . . .	55
6.1.1	Methods and Input Distributions: A Brief Review . . . . .	56
6.2	The Cosmic Flux Decrement Revisited . . . . .	58
6.3	Modelling The Intergalactic Attenuation . . . . .	59
6.3.1	The transmission factor $\Phi$ . . . . .	61
6.4	Our $D_A$ Measurements . . . . .	62
6.4.1	Continuum fit . . . . .	62
6.4.2	Measurement of $D_A$ . . . . .	64
6.5	Results & Discussion . . . . .	64
6.5.1	Observations vs. Models . . . . .	64
6.5.2	Distribution of $D_A$ . . . . .	67
6.5.3	Scatter in $D_A$ . . . . .	73
6.6	Summary & Conclusions . . . . .	74
6.7	Acknowledgments . . . . .	75
<b>7</b>	<b>Photometric Redshifts</b>	<b>77</b>
7.1	Introduction . . . . .	77
7.2	Model Description & Input Physics . . . . .	78
7.2.1	Absorption masks . . . . .	78
7.2.2	The Statistics . . . . .	81
7.2.3	The Model Galaxy Spectra . . . . .	82
7.3	Variations in magnitudes . . . . .	82
7.3.1	The intergalactic $k$ -correction . . . . .	83
7.4	Implications for the estimates of photometric redshifts . . . . .	88
7.4.1	Effects on the photometric redshifts . . . . .	89
7.5	Summary & Conclusions . . . . .	92
<b>8</b>	<b>Summary &amp; Outlook</b>	<b>93</b>

---

<b>A</b>	<b>The Lyman series</b>	<b>104</b>
<b>B</b>	<b>The Gauß approximation to the Voigt-Hjerting function</b>	<b>106</b>
<b>C</b>	<b>Dependence of the density on the scale factor</b>	<b>109</b>
<b>D</b>	<b>The proper path length</b>	<b>111</b>
<b>E</b>	<b>Evolutionary parameters of the <math>Ly\alpha</math> absorbers</b>	<b>113</b>

---

## List of Figures

---

2.1	Transitions from the ground to various excited states of the hydrogen atom . . . . .	6
2.2	Curve-of-growth for the first six Lyman transitions of H $\alpha$ . . . . .	10
2.3	Curve-of-growth for the Ly $\alpha$ for different Doppler parameters . . . . .	10
2.4	Ly $\alpha$ Voigt profiles for different column densities and a Doppler parameter $b = 36 \text{ km s}^{-1}$ . . . . .	11
2.5	Effect of the Doppler parameter on the Voigt profile of the Ly $\alpha$ line at a fixed column density . . . . .	12
3.1	Value of the damping parameter $a$ for different elements in different ionisation stages typically found in QSO spectra . . . . .	16
3.2	Departure of the Voigt-Hjerting function from a pure Gaussian function . . . . .	17
3.3	Contribution from the zeroth and first order terms of the series (3.3) to the absorption coefficient of H $\alpha$ . . . . .	18
3.4	Precision of different methods to compute the Voigt-Hjerting function, relative to Monaghan's differential method . . . . .	22
3.5	Lower limit to the precision of our method to synthesise Voigt profiles . . . . .	23
3.6	Ly $\alpha$ absorption line profiles for three different column densities . . . . .	24
3.7	Distribution in column density and Doppler parameter of the absorbing systems along the line-of-sight towards the source QSO J2233-606 . . . . .	25
3.8	Observed and synthetic spectrum of the quasar HDF-S QSO J2233-606 . . . . .	26
4.1	Galaxy redshift distribution from the 2dFGRS . . . . .	33
4.2	Spectrum of the quasar PKS0454+039 at $z_{em} = 1.34$ (Charlton et al. 2000) . . . . .	37
4.3	Comparison of the Ly $\alpha$ absorption on QSO spectra at two different redshifts (Charlton et al. 2000) . . . . .	38
4.4	Number density evolution of the Ly $\alpha$ forest (Kim et al. 2002a) . . . . .	39
4.5	Number density evolution of the Lyman limit systems (Sargent et al. 1989) . . . . .	40
4.6	Number density evolution of the damped Ly $\alpha$ absorbers (Wolfe et al. 2005) . . . . .	40
4.7	Column density distribution function of the Ly $\alpha$ absorbers (Kim et al. 2002a) . . . . .	42
5.1	Absorption of a single Ly $\alpha$ absorber on a constant input spectrum . . . . .	49
6.1	Spectrum of the QSO SDSS J112107.99+513005.4 at $z_{em} = 2.843$ . . . . .	64

6.2	Evolution of $D_A$ computed using the MMC, the BMC, and MTC models, compared to observations performed over the past twenty years by different groups and different methods . . . . .	66
6.3	Evolution of the mean $D_A$ according to the MMC model in the range $0.2z_{em} < 6$ . . . . .	66
6.4	Quantile-quantile plot for the MMC model . . . . .	69
6.5	Distribution of the $D_A$ values at a given redshift computed from simulations based on the MMC model . . . . .	70
6.6	Values of $\delta$ for the MMC model assuming a lognormal and Gaussian parent distribution . . . . .	72
6.7	Curve of growth of the Ly $\alpha$ absorption line, for three typical values of the Doppler parameter . . .	73
6.8	Evolution of the intrinsic scatter of $D_A$ due to the stochastic nature of the absorption in the intergalactic medium for the two competing models MMC and BMC . . . . .	74
7.2	Example of an absorption mask for a single line of sight out to a redshift $z_{em} = 3.5$ . . . . .	80
7.1	Absorption line strength at line centre, relative to Ly $\alpha$ . . . . .	81
7.3	Comparison of the mean absorption computed according to the MMC model and the Madau model . . . . .	81
7.4	Intergalactic $k$ -correction in Johnson's $U$ , $B$ , $V$ , and $R$ bands for an Sd-type galaxy . . . . .	84
7.5	Evolution of $U - B$ an Sd-type galaxy including different models for the intergalactic attenuation . . . . .	85
7.6	Evolution of $B - V$ an Sd-type galaxy including different models for the intergalactic attenuation . . . . .	85
7.7	Evolution of $V - R$ an Sd-type galaxy including different models for the intergalactic attenuation . . . . .	85
7.8	Comparison of the bandwidth forof two different sets of filters: Standard Johnson's $U$ , $B$ , $V$ , and $R$ and Strömgren $u$ , $v$ , $b$ , and $y$ . . . . .	86
7.9	Intergalactic $k$ -correction in Strömgren's $u$ , $v$ , $b$ , and $y$ bands for an Sd-type galaxy . . . . .	87
7.10	Photometric redshifts for a sample of mock SEDs computed in Johnson's $U$ , $B$ , $V$ , and $R$ bands for an Sd-type galaxy . . . . .	91
7.11	Photometric redshifts for a sample of mock SEDs computed in Johnson's $U$ , $B$ , $V$ , and $R$ for an Sd-type galaxy, compared to a grid of model SEDs computed in the same filter system for an Sa-type galaxy . . . . .	92

---

## List of Tables

---

4.1	State-of-the-art measurements of some cosmological parameters, obtained from the three-year WMAP Data Release ( <a href="#">Spergel et al. 2006</a> ). . . . .	32
4.2	Geometry of Ly $\alpha$ absorbers, inferred from observations along the lines-of-sight to QSO pairs and/or to multiple images of one and the lensed QSO ( <a href="#">Rauch 1998</a> ) . . . . .	45
6.1	Types of absorbers and their corresponding parameters adopted from <a href="#">Madau (1995)</a> . . . . .	60
6.2	Types of absorbers and their corresponding parameters, adopted from <a href="#">Bershady et al. (1999)</a> . . . . .	60
6.3	QSO sample selected from the SDSS DR5 . . . . .	63
7.1	Characteristic redshifts for two different sets of filters: Standard Johnson's $U$ , $B$ , $V$ , and $R$ and Strömgren $u$ , $v$ , $b$ , and $y$ . . . . .	88
A.1	Some characteristic quantities for the first 24 Lyman transitions . . . . .	105
E.1	Compilation of the evolutionary parameters of the Ly $\alpha$ absorbers found in the literature . . . . .	114

---

## Introduction

---

*The upshot of all this is that we live in a universe whose age we can't quite compute, surrounded by stars whose distances from us and each other we don't altogether know, filled with matter we can't identify, operating in conformance with physical laws whose properties we don't truly understand.*

[Bill Bryson, *A Short History of Nearly Everything*]

From the moment science was born, mankind has continuously pursued a quest for a so-called Theory-of-Everything – Heisenberg's celebrated *Weltformel* –, *i.e.* for a system of physical laws which consistently accounts for all the phenomena observed in the World. In the last century in particular, this quest has led science into two, apparently mutually exclusive and opposite directions: the study of the Universe on smallest scales, the so-called micro-cosmos, and the study of the macro-cosmos, *i.e.* of the Universe at the largest imaginable scales.

The first approach has led to the discovery and development of theories such as Quantum Mechanics, which is as fascinating as it is puzzling. Leaving the practical difficulties inherent to its mathematical complexity aside, it appears as if Quantum Mechanics could account for every single phenomenon observed at the atomic level. The second approach, *i.e.* the study of the (macro-)Cosmos – referred to as *Cosmology* –, has been developed on the basis of the Theory of General Relativity, one of greatest intellectual achievements in the history of mankind. General Relativity is the theory

which most accurately describes the gravitational force to date, which in turn is the dominant interaction at the largest scales in the Universe. *Ergo*, one might think that the description of the Universe is also complete at this scales. However, the situation is a little more complex than this.

In its modern conception, Cosmology deals more with the evolution of the Universe as a whole, with its structure and its contents, and rather less with the question/problem about its creation, leaving it aside mainly because of the failure of physical laws to deal with the description of extreme conditions in Nature, and the creation of the Universe is certainly one of this kind! Nonetheless, Cosmology remains an exciting subject, powered in part by the many puzzles raised in the last decade through the improvement of the observational techniques. Among these puzzles, the putative existence of the so-called Dark Matter and of a mysterious Dark Energy are the most striking.

In the last decade or so, a consistent description about the phenomena observed at cosmological scales has been developed which is known as Concordance Cosmology. This model accounts for the three major pieces of evidence for the Big Bang theory: the existence of a cosmic microwave background radiation, the expansion of the Universe, and its observed large-scale structure. However, despite its success, and in addition to the puzzles mentioned above, an important open question still remains: the problem of the formation of galaxies and their evolution.

This subject is of particular interest for the develop-

ment of Cosmology, mainly because of the complex feed-back mechanisms at galactic scales which determine both the fate of galaxies and their surroundings. It is now widely accepted that galaxy formation is a direct consequence of the structure formation in the Universe. Nevertheless, it is not truly understood how galaxies *actually* form, neither is it known how they evolve in time or what exactly is the origin of the variety of types observed in deep images – or Deep Field Observations – of the Universe.

Present-day observational facilities like the Hubble Space Telescope (HST) and ground-based 8-10 m class telescopes provide an unprecedented wealth of data on galaxies at significant redshifts, *i.e.* at very large distances. During the last decade, in particular, with multi-wavelength Deep Field Observations (such as the Hubble-Deep Fields North & South, Hubble-Ultra Deep Field, etc) in combination with spectral and photometric redshift determinations, it has become possible to study the evolution of representative samples of galaxies - not just the brightest few percent at any redshift/distance/epoch - over a wide range of redshifts, and hence, look-back times.

Thus, for the first time the interpretation of a statistically relevant number of high redshift galaxies (*i.e.* several thousand galaxies at redshifts between 2.5 and 3.5 and several hundreds at  $z > 3.5$ ) can be performed, with which a deeper insight into the processes involved in the formation and evolution of galaxies can be gained. It should be noted that redshifts between 2.5 and 3.5 correspond to look-back times of 80% to 90% of the age of the Universe in Concordance Cosmology. Hence, the wealth of data at hand opens the possibility of studying the properties of galaxies at their youngest evolutionary stages.

However, the analysis and interpretation of data, especially at the highest redshifts, is of course not straightforward, since the light emitted by galaxies – and hence their spectro-photometric properties – is affected by different factors along its way through the Universe. First, due to the finite velocity of light, distant galaxies are seen in younger evolutionary stages having formed only a fraction of the stars that similar galaxies in the Local Universe would have, an effect that is known as *evolutionary effect*. Correspondingly, their spectra and their fluxes measured in a given wavelength range are different from those of local galaxies, making a direct comparison of their properties virtually impossible.

Moreover, according to Big-Bang Nucleosynthesis,

only hydrogen (H), helium (He) and some light elements up to lithium (Li) were produced shortly after the Big Bang, while all other heavier elements present in the Universe were synthesized within earlier generations of stars and set free to the so-called interstellar medium (ISM) at the end of the stars' life by stellar winds and/or planetary nebulae (PNe) or Supernovae (SNe) explosions. As a consequence, successive stellar generations are formed from increasingly enriched material (with respect to the content of elements heavier than He). Accordingly, galaxies observed at high redshifts and hence in younger evolutionary stages do not have reached the enrichment levels observed in the Local Universe. Since the chemical abundances affect the lifetimes, gas- and heavy element-output rates, evolutionary tracks, colors, luminosities and spectral properties of the stellar population, the lower enrichment of galaxies at higher redshifts correspondingly modifies their observed photometric properties with respect to local, more enriched galaxies. Both the evolutionary effect and the effect of lower enrichment are intrinsic to the evolution of the galaxies.

Due to the expansion of the Universe, the wavelength of the light emitted by galaxies at very large distances is redshifted and their flux dimmed, their colors being consequently reddened and their luminosities weakened. In addition to this so-called *cosmological effect*, the spectra of distant sources are also affected by the fact that, being extreme ultra-high-vacuum by terrestrial standards, the space between galaxies – properly called *intergalactic medium* or IGM – is by no means empty, but is filled with a very low density gas. Due to the strong radiation background from quasars and star forming galaxies, this intergalactic gas – mainly of primordial abundance – is highly ionised. It contains, however, a small amount of neutral hydrogen (around one part in ten thousand) distributed along the line-of-sight, which significantly absorbs the light from distant galaxies and quasars at rest-frame wavelengths  $\lambda \leq 121.567$  nm which correspond to the Lyman resonant transitions of H $\alpha$ . Hence, the intergalactic attenuation adds to the cosmological effect, contributing to the dimming and reddening of the colors of high-redshift galaxies, an effect that has become to be known as *intergalactic attenuation*.

In order to account for the effect of the intergalactic attenuation on the spectro-photometric properties of galaxies, a detailed knowledge of the physical state (density, temperature) and of the distribution of intergalactic H $\alpha$  embedded in the IGM is required. Ob-

servations in different wavelength ranges of large volumes of space, so-called redshift surveys (*e.g.* 2dFGRS, SDSS), in combination with numerical simulations of structure formation based on the concordance (or  $\Lambda$ CDM) paradigm, reveal that the distribution of galaxies in the Universe trace at all scales the underlying distribution of Dark Matter (DM). In this scenario, the potential wells of DM act as a sink for the baryonic matter in the IGM—in particular for the diffuse, highly ionised intergalactic gas—, leading to the formation of structures by means of gravitational collapse. As a consequence, the matter distribution in the Universe, and in particular the distribution of gas in the IGM, becomes clumpy, with large voids (*i.e.* empty regions in space) alternating with filamentary structures and denser knots at which galaxy clusters form.

When the line-of-sight to a background source happens to intersect these structures, the distribution of H I gives rise to a series of discrete absorption lines on its spectrum. The properties (column densities, Doppler parameters) of these absorption features, commonly known as Ly $\alpha$  forest, encode the physical conditions of the medium where the absorption takes place. Thus, the properties of the intergalactic H I can be inferred from the analysis of the absorption lines observed in the spectra of bright background sources (*e.g.* a quasar). However, in order to extract the information from these absorption lines, a detailed analysis of their profiles is required. In this regard, this work makes a contribution to improve the analysis of this type of line profiles by providing a mathematical tool with which this profiles can be accurately described for a wide range of values for the column densities and Doppler parameters characteristic to intergalactic H I.

Once the properties characterising the intergalactic H I and their evolution are known, a model that account for its effect on the spectra of galaxies can be constructed. The first step toward such a model was developed by Madau (1995). In this model, Madau developed an analytic approach with which the spectrum of a galaxy can be corrected for the *mean* absorption by intergalactic H I, and its mean colors and luminosities can directly be computed. A more recent, and fundamentally different model of the intergalactic attenuation was given by Bershadsky et al. (1999). In contrast to Madau's analytical approach, this model accounts for the neutral hydrogen absorption on the basis of numerical Monte-Carlo simulations. Hereby, an ensemble of lines-of-sight with randomly generated populations of absorbers is created and colors and luminosities are cal-

culated for each line-of-sight independently. Mean luminosities and colors are then calculated as a arithmetical sample average of the former.

It turns out that a consistent description of the redshift evolution of the galaxies including evolutionary and cosmological effects, as well as the intergalactic attenuation, can only be achieved by using Evolutionary Synthesis (ES) models that describe the evolution of a galaxy on the basis of its Star Formation History (SFH) in terms of the time evolution of its stellar population, its integrated colors and luminosities in various wavelength regions (or filter bands), and in terms of the time evolution of the integrated galaxy spectra (like the GALEV code by Fritze-v. Alvensleben 1999) – provided a cosmological model is appropriately included. ES models are based on the fact that - when averaged over long enough periods of time and integrated over the entire galaxy - SFHs are fairly well constrained by a multitude of observations for respective galaxy types or templates (E, S0, Sa-Sd) and different enough from one type to the other to discriminate between types, provided the observational wavelength baseline is long enough (see *e.g.* Arimoto & Yoshii 1986, Arimoto & Jablonka 1991, Bruzual A. & Charlot 1993, Fritze-v. Alvensleben 1999, Vázquez et al. 2003, Le Borgne et al. 2004, Maraston 2005, Delgado et al. 2005, Mollá 2007).

In its present form, the Göttingen GALEV Code includes Madau's approach to account for the mean attenuation of distant galaxy spectra. However, due to the random distribution of H I along the line-of-sight and the corresponding stochastic absorption along different lines-of-sight, this model does not allows a consistent comparison with observations, in which the stochastic nature of the absorption is unavoidably present.

Hence, the specific goal of this work is to develop a model that accounts for the stochastic attenuation due to the random distribution of intergalactic H I along the line-of-sight. This model is necessary to address the impact of the stochastic intergalactic attenuation on the photometric properties of galaxies, and to compute the corresponding dispersion in their luminosity and color evolution. Before this can be carried out, this model is tested against observations, in particular, to the observed evolution of the cosmic flux decrement, using different input distribution functions for the absorber properties. In a second step, these predictions shall be compared to Deep Field observations, in order to determine which fraction of the galaxy population in certain redshift intervals can be accommodated within the range of normal galaxy models and what are the char-

acteristics of the others.

In what follows I give a brief summary of the contents of this work: Some basics on the physics of the hydrogen atom and the interaction between matter and radiation are given in the first chapter. The second chapter deals with an approximation to the Voigt-Hjerting function which is of great importance in the modelling and synthesis of absorption line profiles due to intergalactic neutral hydrogen. The third chapter gives a brief summary of the currently accepted picture of the Universe, its evolution, and introduces and explains the properties of the intergalactic medium and the Ly $\alpha$  absorbers. In the fourth chapter, our model of the stochastic absorption of the light of background sources due to intergalactic neutral hydrogen based on the Monte Carlo approach to model large numbers of individual lines-of-sight is introduced. The basic input of this model is a set of distribution functions of the physical parameters for the Ly $\alpha$  absorbers. Since the observational determination of this type of distribution functions are not ultimately settled nowadays, we use our model with two different sets of distributions constrained from observations to compute the redshift evolution of the cosmic flux decrement. Contrasting our results to observations allows us to identify one set of distributions that best matches the data. The computation of the evolution of the cosmic flux decrement and its scatter due to the stochastic nature of the absorption using our model, and a comparison between our predictions and observations is discussed in the fifth chapter. Since the observations used to constrain the distribution of the parameters for the Ly $\alpha$  absorbers and those used to measure the redshift evolution of the cosmic flux decrement are orthogonal, we consider that the method in addition to this set of input distribution very accurately accounts for the intergalactic attenuation. On this basis, we use our model to accurately assess the stochastic effect of intergalactic absorption on the spectra of high-redshift galaxies, and its impact on the estimates of photometric redshifts. A summary and an outlook are given in the last chapter.

It should be noted that some results and equations may be repeated through the various chapters, some of which are already published or on in the process of acceptance/publication as independent papers. I decided to reproduce this papers in completeness, and ask the reader to kindly apologise for this redundancy.

---

## The Physics of Neutral Hydrogen

---

In this chapter, some theoretical results on the interaction between neutral hydrogen and radiation relevant for the next chapters are presented.

It is well known that neutral hydrogen interacts with radiation by absorbing or emitting photons. In this work, only the absorption processes are of importance and hence, emission processes will not be discussed. On the other hand, absorption processes relevant for this work will be treated with some detail, in particular the origin of the Lyman-Series.

### 2.1 THE LYMAN SERIES

Neutral hydrogen is known to be excited from low to higher energy states by the absorption (or scattering) of photons with certain wavelengths, by means of which the electron is pushed from a lower to a higher energy level. The set of transitions from ground level to the first, second, etc. excited state, or equivalently from these excited states to the ground state, specifically the transitions  $np \rightarrow 1s$ , are known as the Lyman series<sup>1</sup> (see Fig. 2.1). These transitions are quantised, *i.e.* they take place in principle at a well defined set of discrete energies. The energy differences between the ground level and each of the excited states correspond to a given resonant wavelength of the incident photon, as can be seen from the relation  $\Delta E = h\nu = hc/\lambda$ . The energy of

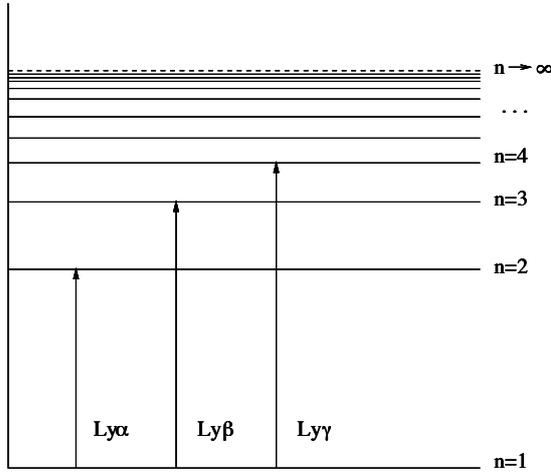
the  $n$ th excited state of the hydrogen atom is given by

$$E_n = -\frac{m_e e^2}{2(4\pi\epsilon_0 \hbar)^2} \frac{1}{n^2} = -\frac{13.6}{n^2} [\text{eV}], \quad (2.1)$$

where  $m_e$  is the mass of the electron,  $e$  is the fundamental charge unit,  $\epsilon_0$  is the vacuum permittivity (also called electric constant), and  $\hbar$  is Planck's constant  $h$  divided by  $2\pi$ .<sup>2</sup> Recall that  $1 \text{ eV} = 1.602 \cdot 10^{-19} \text{ J}$ . The transition between the ground level (which corresponds to an energy of  $E_0 \approx -13.6 \text{ eV}$ ) and the first excited state (with an energy of  $E_1 \approx -3.4 \text{ eV}$ ) is known as the Ly $\alpha$  transition. This has a corresponding resonant wavelength of approximately  $121.567 \text{ nm}$  (or, equivalently, an energy of  $E_1 \approx -10.2 \text{ eV}$ ). The transition from the ground level to the second, third, etc. excited states are correspondingly known as Ly $\beta$ -, Ly $\gamma$ -, etc. transitions, and they have associated energies which correspond to the resonant wavelengths  $\lambda_\beta = 102.572 \text{ nm}$ ,  $\lambda_\gamma = 97.253 \text{ nm}$ , etc., respectively (see appendix A for a list of the first 24 Lyman transitions and their characteristic quantities). All these transitions correspond to a single type known as *bound-bound* transitions, since the electron remains always attached to its nucleus. The exception to this, a so-called *bound-free* transitions, takes place when the atom absorbs a photon with an energy exceeding  $E_\gamma \geq 13.6 \text{ eV}$ , which corresponds to a resonant wavelength of  $\lambda_\gamma \leq 91.175 \text{ nm}$ , in which case the electron is completely removed from the nucleus. This process is known as photoionisation, and

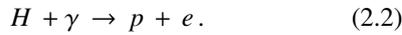
<sup>1</sup> Named after its discoverer, the U.S. physicist Theodore Lyman (\*1874, †1954)

<sup>2</sup>  $m_e = 9.105 \cdot 10^{-31} \text{ kg}$ ,  $e = 1.602 \cdot 10^{-19} \text{ C}$ ,  $\epsilon_0 = 8.854 \cdot 10^{-12} \text{ C}^2 \text{ N}^{-1} \text{ m}^{-2}$ ,  $h = 6.626 \cdot 10^{-34} \text{ J} \cdot \text{s}$



**Fig. 2.1:** Energy levels of the neutral hydrogen atom, and the first three transitions (at absorption) of the Lyman series: Lyman-Alpha ( $\text{Ly}\alpha$ ), Lyman-Beta ( $\text{Ly}\beta$ ), and Lyman-Gamma ( $\text{Ly}\gamma$ ). In general, the Lyman transitions can be denoted by  $np \rightarrow 1s$  with  $n \in \{2, 3, \dots\}$ .

corresponds to the process<sup>3</sup>



The wavelength  $\lambda_L = 91.175$  nm is commonly referred to as the Lyman edge.

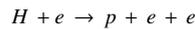
Because of the quantum nature of these interactions, not necessarily every photon interacts with an atom; rather, there is only a probability that such a process may take place. For instance, the cross-section for photoionisation, *i.e.* the probability that a hydrogen atom absorbs a photon with a wavelength  $\lambda \leq \lambda_L$  is given by<sup>4</sup>

$$\sigma_{\text{phot}}(\lambda) = g(\lambda) \cdot \sigma_{\infty} \cdot \left(\frac{\lambda}{\lambda_L}\right)^3, \quad (2.3)$$

where<sup>5</sup>

$$\sigma_{\infty} = 6.31 \cdot 10^{-18} \text{ cm}^2, \quad (2.4)$$

<sup>3</sup> Collisional ionisation, which corresponds to the process



will not be considered here.

<sup>4</sup> We adopt the convention that the  $\text{Ly}\alpha$  transition (from the ground state to the next higher energy level) be identified with  $i = 2$ , the  $\text{Ly}\beta$  transition with  $i = 3$ , etc. The photoionisation cross-section is thus consistently denoted by  $\sigma_{\infty}$ .

<sup>5</sup> For an experimental determination of  $\sigma_{\infty}$  see *e.g.* Palenius et al. (1976).

and  $g$  is the so-called Gaunt factor.<sup>6</sup>

In what follows we will drop the super-index  $\text{HI}$ , bearing in mind that throughout this work we are only interested in neutral hydrogen.

The cross-section for photons with  $\lambda > \lambda_L$ , *i.e.* for resonant transitions, is in principle not a continuous function of wavelength, like the expression (2.3), but is defined just for discrete values of  $\lambda$ , which correspond to the resonant wavelengths of the Lyman series. The general form of this cross-section for the transition  $jp \rightarrow 1s$ , with  $j \in \{2, 3, 4, \dots\}$ , is given by<sup>7</sup>

$$\sigma_j = \frac{\pi e^2}{m_e c} f_j, \quad (2.5)$$

where  $f_j$  is a dimensionless quantity, the so-called oscillator strength, which describes the strength of the corresponding transition. The discrete nature of the cross-section is precisely due to its dependence on this quantity, which is different and constant for each transition (see appendix A for a list of the oscillator strengths for the Lyman transitions).

## 2.2 ABSORPTION DUE TO HI ATOMS

An ensemble of atoms in gaseous state and in local thermodynamical equilibrium (LTE), in this case neutral hydrogen gas, may be characterised through its density  $n$  – *i.e.* the number of atoms per  $\text{cm}^3$  – and temperature  $T$ . It turns out that in the astrophysical context, instead of its density, a quantity called column (or surface) density is more suitable in order to describe the interaction between atoms in gaseous form and radiation. The column density, usually denoted by  $N$  (or  $N_{\text{HI}}$  in our case) gives the number of atoms per unit area along a line-of-sight through the gas under consideration. For a gas with a space-dependent density  $n_{\text{HI}}(\mathbf{x})$  and a given geometry of its volume, one may define

$$N_{\text{HI}} \equiv \int_{-l}^l n_{\text{HI}}(\mathbf{x}) \cdot d l', \quad (2.6)$$

where the integral is computed along the line-of-sight, and  $l$  is some characteristic measure for a linear dimension of the given geometry. For example,  $l$  would be the radius of a spherical gas cloud. In any case,  $N_{\text{HI}}$  should

<sup>6</sup> An extensive tabulation of values for the Gaunt-factor can be found in Karzas & Latter (1961).

<sup>7</sup> From now on we designate the principal quantum number  $n$  with  $j$ , in order to avoid confusion in the next section with the volume density of gas, also designated with  $n$ .

be defined in such a way that it accounts for the entire *geometric* cross-section of the atoms along an arbitrary line-of-sight through the gas.

In the special case that the atomic motion is purely of thermal nature, one may introduce a new parameter as a measure of this motion. The thermal motion is described by a Gaussian distribution of the thermal velocities, also known as Maxwell-Boltzmann (MB) distribution, which is given by

$$M(v) dv = \frac{1}{\sqrt{\pi}} e^{-(v/v_0)^2} \frac{dv}{v_0}, \quad (2.7)$$

where  $v_0$  is a parameter, which is uniquely related to the temperature of the ensemble of atoms as we shall see. The distribution (2.7) quantifies the probability of finding an atom moving in a random direction with a velocity in the range  $(v, v + dv)$ . The Boltzmann factor, as the exponential term in the above equation is called, quantifies the probability of a microstate of energy  $E_m$  to be found within a macrostate of energy  $E_M$ , according to  $\exp(-E_m/E_M)$ . In this particular case, the energy of the microstate is the kinetic energy of a particle with mass  $m$ ,  $1/2 m v^2$ , and the energy of the macrostate is the total energy of the system, which is completely defined by its kinetic temperature,  $kT$ , where  $m$  is the mass of an atom and  $k$  is Boltzmann's constant<sup>8</sup>. Thus we have

$$\frac{E_m}{E_M} = \frac{1}{2} \frac{m v^2}{k T}, \quad (2.8)$$

and by comparison with the exponent in equation (2.7) it is evident that the parameter  $v_0$  is related to the temperature via

$$v_0 = \sqrt{\frac{2kT}{m}}. \quad (2.9)$$

This relation is uniquely defined, since only the positive sign of the square-root is considered. On this basis, we choose to describe the thermal motions of the atoms in terms of  $v_0$  instead of  $T$ . For consistency with the literature, we change our notation and write  $b \equiv v_0$  and call  $b$  the Doppler parameter.<sup>9</sup>

This Doppler parameter is particularly useful when considering a gas which, in addition to thermal motions, has a Gaussian distributed turbulent velocity component

$v_{turb}$ . In this case, the Doppler parameter  $b$  is related to the temperature and this turbulent component via

$$b^2 = \frac{2kT}{m} + v_{turb}^2. \quad (2.10)$$

This result follows naturally from the fact that the composite velocity distribution is given by the convolution of two Gaussian distributions – the Maxwell-Boltzmann distribution for the thermal and a Gaussian distribution for the turbulent component –, and that the convolution of two such functions is again a Gaussian distribution, with a dispersion (variance) which is just the sum of the individual dispersions, *i.e.*  $\sigma^2 = \sigma_1^2 + \sigma_2^2$ .

Once the concepts of the column density and the Doppler parameter are defined, the interaction between an ensemble of atoms (gas) and photons can be quantitatively described.

### 2.2.1 The absorption coefficient of H I

Let  $f(\lambda)$  be the monochromatic photon flux, *i.e.* the number of photons with a given wavelength (energy)  $\lambda$ , which intercepts a gas with column density  $N_{\text{HI}}$  and Doppler parameter  $b$  along a particular line-of-sight in the direction of  $\hat{r}$ . Then, the fraction  $df$  of photons, which are scattered along the path  $dl = |d\vec{l}| \equiv |dl\hat{r}|$ , *i.e.* the difference between the flux after and before the scattering process is given by

$$df = -N_{\text{HI}} \cdot f \cdot \omega_\lambda dl, \quad (2.11)$$

if one supposes that  $\omega_\lambda dl$  is just the probability that one single photon is scattered by a single atom along the path  $dl$ . The quantity  $N_{\text{HI}} \cdot f \cdot \omega_\lambda dl$  is thus the number of photons that are scattered off the line-of-sight—hence the negative sign in equation (2.11). This missing photons are the reason for the absorption feature—mostly an absorption line—in the observed flux. After integration along the line-of-sight between the source and the observer it follows that the observed flux is given by

$$f = f_0 \cdot e^{-\tau}, \quad (2.12)$$

where  $f_0$  is the total number of atoms before the absorption process and

$$\tau(\lambda) \equiv N_{\text{HI}} \cdot \int_0^L \omega_\lambda dl \quad (2.13)$$

is the so-called absorption coefficient of the gas. In this last equation, the quantity  $L$  is the absorption path in

<sup>8</sup>  $k = 1.3803 \cdot 10^{-23} \text{J} \cdot \text{K}^{-1}$

<sup>9</sup> This term is based on the fact that  $b$  is a characteristic velocity for the motion of the atoms in the gas, and as such it is a measure of the Doppler broadening of the resonant wavelength in absorption processes (see below).

the gas in the direction  $\hat{r}$ , *i.e.* along the line-of-sight. In this context, the absorption coefficient is a measure for the probability of photons with a given energy to be scattered by gas atoms. In other words,  $\tau$  is nothing but the total cross-section of all atoms in a gas, which again is given by the product of the cross-section of a single atom and the total, geometrical cross-section, *i.e.* the column density of the gas.

According to equation (2.3), the absorption coefficient for photoionisation, *i.e.* for photons with energies corresponding to  $\lambda \leq \lambda_L$ , is given by<sup>10</sup>

$$\tau_{phot}(\lambda) = N_{\text{HI}} \cdot \sigma_{\infty} \cdot \left(\frac{\lambda}{\lambda_L}\right)^3. \quad (2.14)$$

As noted before, the form of the absorption coefficient for resonant transitions is more complicated than the previous equation, due to different mechanism which alter the frequency of the photons that are absorbed. In other words, the absorption that causes a given transition does not take place at a given photon frequency, but rather in a range of frequencies around the central (nominal) frequency, and even this frequency may be shifted from its nominal value. Usually, these broadening and shifting mechanisms are divided into local and non-local. Local mechanisms assume that the absorbing medium can be treated as being in LTE. Under this conditions, the main broadening mechanisms are the natural damping broadening, thermal or Doppler broadening, and pressure broadening.

Natural damping broadening is due to the finite lifetime  $\Delta t$  of atomic transitions. Thus, according to Heisenberg's Uncertainty principle, these transitions take place for photons with energies in a narrow, but finite energy range  $\Delta E \sim 1/\Delta t$ . This damping broadening can be classically accounted for, if one assumes that the interaction between a photon and an electron can be described in terms of a damped, harmonic oscillator. It turns out that a quantum-mechanical treatment of this effect yields the same result as the classical approach. In either case, the strength or probability of the absorption process as a function of the photon's frequency is well-described for a particular transition in terms of a Lorentzian distribution

$$L(\nu) = \frac{1}{\pi} \frac{\Gamma/4\pi}{(\nu - \nu_j)^2 + (\Gamma/4\pi)^2}. \quad (2.15)$$

In this equation,  $\nu_j$  is the resonant frequency of the transition  $jp \rightarrow 1s$  and  $\Gamma^{-1}$  is its mean lifetime. The

<sup>10</sup> The Gaunt factor introduced before is here approximated to unity.

quantity  $\Gamma$  is commonly referred to as the damping constant (see appendix A for a list of  $\Gamma$  values for the first 24 Lyman transitions). Pressure broadening takes place when the emitting or absorbing particle interacts with neighbouring particles, either by collisions or electromagnetically. In either case, the interactions modify the lifetime of the transitions, thus changing the energy of the emitted or absorbed radiation. Pressure broadening depends both on the density and the temperature of the gas, and it can be described in terms of a Lorentzian distribution of the form of equation (2.15) as well.

Both the absorption cross-section  $\sigma_i$  and the Lorentz profile  $L(\nu)$  quantify the probability of statistically independent absorption processes leading to the same transition, and thus the total probability of a photon to be absorbed by a single, neutral hydrogen atom, *i.e.* the absorption coefficient is given by the product of both probabilities,

$$\tau'_j(\nu) = \sigma_j \cdot L(\nu). \quad (2.16)$$

Note that this expression is valid only in the rest-frame of the atom, denoted by the prime. In the more general case of an atom moving with respect to an observer, one has to consider the Doppler shift of the photon frequency in the rest-frame of the atom, which in turn affects the energy range of photons that may be absorbed. If  $\mathbf{v}$  is the velocity of the atom (observer's rest-frame), and  $\nu$  the projection of  $\mathbf{v}$  along the line-of-sight<sup>11</sup>, then the Doppler shift  $\Delta\nu_j$  of the resonant frequency  $\nu_j$  is given by

$$\Delta\nu_j = \frac{\nu}{c} \nu_j, \quad (2.17)$$

if we assume that  $\nu/c \ll 1$ , so that the non-relativistic approximation can be used. The absorption coefficient in the observer's rest-frame is thus  $\tau'_j(\nu - \Delta\nu_j)$ . Hence, it is clear that for an ensemble of atoms, *i.e.* a gas, the motion of each individual atom will give rise to a particular shift in the absorption coefficient, and will hence broaden the energy range of photons that may be absorbed, a phenomenon known as Doppler broadening. If the motion of the atoms is thermal in nature, it is described in terms of a Maxwellian distribution (equation 2.7).

We may now easily derive an expression for the absorption coefficient of an ensemble of atoms, *i.e.* a gas with a Doppler parameter  $b$  and column density  $N_{\text{HI}}$ .

<sup>11</sup> The speed  $\nu$  is taken to be negative if the atom is moving away from the observer.

First, the total absorption coefficient can be computed by convolving the MB distribution with equation (2.16), and multiplying the result with the column density

$$\tau_j(\nu) = N_{\text{HI}} \cdot \int_{-\infty}^{+\infty} \tau'_j(\nu - \Delta\nu_D) M(\nu) d\nu. \quad (2.18)$$

Inserting equations (2.7), (2.15), and (2.16), and defining  $\Delta\nu_D \equiv \nu_j b/c$ ,  $a \equiv \Gamma/(4\pi\Delta\nu_D)$ , and the new variables  $x \equiv (\nu - \nu_j)/\Delta\nu_D$  and  $y \equiv \nu/b$ , equation (2.18) can be rewritten as<sup>12</sup>

$$\tau_j(\nu) = N_{\text{HI}} \cdot \frac{\sqrt{\pi} e^2}{m_e c} f_j \frac{H(a, x)}{\Delta\nu_D}, \quad (2.19)$$

where the Voigt-Hjerting function (Hjerting 1938) is defined by

$$H(a, x) \equiv \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(x-y)^2 + a^2} dy. \quad (2.20)$$

Now, replacing frequency by wavelength we get

$$\tau_j(\lambda) = N_{\text{HI}} \cdot \frac{\sqrt{\pi} e^2}{m_e c^2} \frac{\lambda_j^2}{\Delta\lambda_D} f_j H(a, x), \quad (2.21)$$

where  $\lambda_j = c/\nu_j$  and  $\Delta\lambda_D = \Delta\nu_D \lambda_j/\nu_j = \lambda_j b/c$ , and in this case,  $a \equiv \lambda_j^2 \Gamma/(4\pi\Delta\lambda_D)$  and  $x \equiv (\lambda - \lambda_j)/\Delta\lambda_D$ .

The functional form of the absorption coefficient eq. (2.21) gives rise to a particular form of the absorption line profile, known as Voigt profile<sup>13</sup>. As shown above, the Voigt profile is a natural consequence of the fact that the absorbing medium is uniquely determined by its (column) density and kinetic temperature (through its Doppler parameter), which is the case when the gas is under LTE. Hence, any deviation from a pure Voigt profile indicates deviations from those equilibrium conditions.

### 2.2.2 The Curve-of-Growth for H<sub>I</sub>

The line profile contains a wealth of information about the physical state of the absorbing medium. In particular, the strength, *i.e.* the depth and width of the profile encode information about the density and internal kinematics of the gas where the absorption takes place.

Qualitatively, it can be said that for a given column density the line profile becomes shallower and wider with increasing Doppler parameter. Conversely, it is deeper with increasing column density, for a fixed  $b$ . A quantitative measure of the strength of an absorption line can be gained in terms of

$$r(\lambda) \equiv \frac{f_c(\lambda) - f(\lambda)}{f_c(\lambda)}, \quad (2.22)$$

where  $f_c$  is the continuum level and  $f(\lambda)$  is given by equations (2.12) and (2.13). Plugging the latter into equation (2.22) yields

$$r(\lambda) = 1 - e^{-\tau(\lambda)}. \quad (2.23)$$

This quantity, even though being useful, depends on the emitted flux, and it is therefore necessary to introduce a new quantity which is not. The natural choice to get rid of this dependency is to integrate over wavelength and thus to define

$$W_\lambda \equiv \int_0^\infty r(\lambda') d\lambda' = \int_0^\infty (1 - e^{-\tau(\lambda')}) d\lambda'. \quad (2.24)$$

$W_\lambda$  is called the equivalent width.

All the information about the physical properties of the absorbing medium are encoded in the absorption coefficient  $\tau(\lambda)$  as given by equation (2.21), and more precisely in the line profile. As already noted, absorption line profiles are well described by the Voigt-Hjerting function, and the main pieces of information for this type of profile are the central wavelength of the resonant transition (and all the corresponding atomic quantities), the Doppler parameter and the column density. The equivalent width for a given transition can thus be considered as a function of  $N_{\text{HI}}$  and  $b$ , and its behaviour can be analysed for a given range of values of these quantities. This is best achieved through a graphical aid, the so-called curve-of-growth. This construction makes it possible to simultaneously analyse the dependence of the line strength as a function of one of the above parameters, while keeping the other fixed. The curve-of-growth consists of a log-log plot which shows the growth of a given absorption line in terms of the increasing column density of the absorbing medium, for different Doppler parameters.

Due to the complex dependence of the optical depth  $\tau$  on the different parameters  $\{N_{\text{HI}}, b, \lambda_j\}$ , it is not possible in general to compute the values of  $W_\lambda$  analytically. This has to be achieved by numerical integration. There are, however, limiting cases for which an

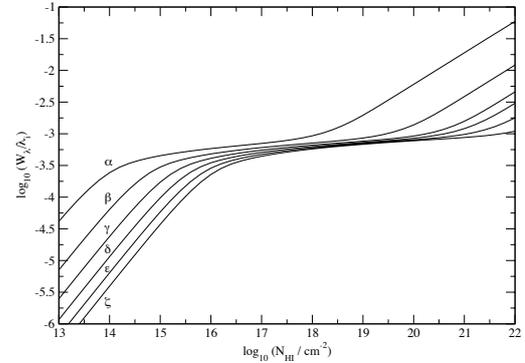
<sup>12</sup> For simplicity, we leave out the index  $i$  in the definition of  $a$ , but one has to keep in mind that its value certainly depends on the parameters of the corresponding transition.

<sup>13</sup> Named after the German physicist Woldemar Voigt (\*1850, †1919).

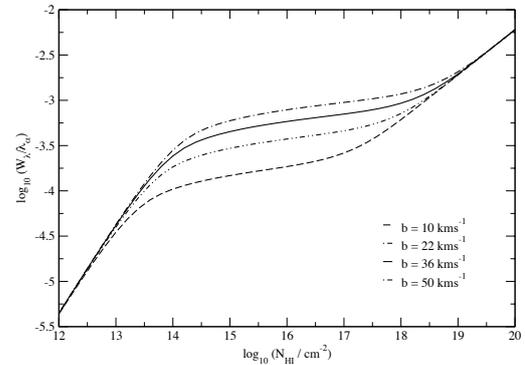
analytical approximation to the equivalent width for a given transition as a function of column density and Doppler parameter can be obtained, assuming that the line profile is given. In Fig. 2.2, the curves-of-growth for the first six Lyman transitions are shown. They have been computed by numerical integration, assuming a Voigt profile and using the approximation to the Voigt-Hjerting function by [Tepper-García \(2006, cf. Chapter 3\)](#). For a small number of absorbing particles, *i.e.* at low column densities, the behaviour of the equivalent width is to a good approximation linear with  $N_{\text{HI}}$ , *ie*  $W_\lambda \propto N_{\text{HI}}$ . This behaviour defines the linear regime of the curve-of-growth, which can clearly be seen in Fig. 2.2 for column densities  $N_{\text{HI}} \lesssim 10^{14.5} \text{ cm}^{-2}$ , for the Ly $\alpha$  transition. As can be seen in Fig. 2.3, the equivalent width is nearly independent of the Doppler parameter for a given transition in the linear regime. When the number of absorbing particles increases, the equivalent width also increases until there is no flux left to absorb, and the absorption at the central wavelength reaches its lowest possible value. The line is said to be saturated. In this regime, the equivalent width changes very slowly with  $N_{\text{HI}}$  as  $\sqrt{\ln N_{\text{HI}}}$ , and defines the flat part of the curve-of-growth. If the number of absorbers continues to grow, their column density correspondingly increases, and their interactions/collisions are so frequent that radiation damping becomes significant. Hence, even though more absorption at the line centre is not possible due to saturation, the equivalent width of the line increases again due to the absorption at wavelengths slightly shorter or larger than the central wavelength according to a Lorentzian distribution of the photon frequencies (equation 2.15). This is the mechanism that gives rise to the damping wings, which can be seen in Fig. 2.4(b). In this regime, *i.e.* column density range—called *square-root* regime—the equivalent width changes as  $W_\lambda \propto \sqrt{N_{\text{HI}}}$ , and is nearly independent of the Doppler broadening.

It should be noted, as can be seen in Fig. 2.2, that the behaviour of the curve-of-growth is qualitatively the same for any transition, even though not quantitatively. For instance, the Ly $\alpha$  absorption line saturates at column densities  $N_{\text{HI}} \approx 10^{14.5} \text{ cm}^{-2}$ , while the Ly $\zeta$  absorption line only saturates at column densities around  $10^{17} \text{ cm}^{-2}$ .

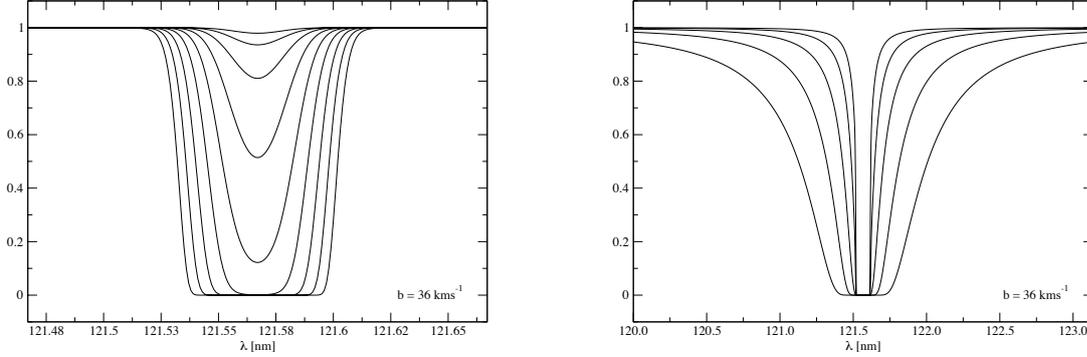
The behaviour of the equivalent width as given by the curve-of-growth can be directly observed in the line profile, as it should be. Ly $\alpha$  absorption Voigt profiles for different column densities and a Doppler parameter  $b = 36 \text{ km s}^{-1}$  are shown in Fig. 2.4. Note in



**Fig. 2.2:** Curve-of-growth for the first six Lyman transitions of neutral hydrogen for  $b = 36 \text{ km s}^{-1}$ , which approximately corresponds to the observed mean for intergalactic H I (see chapter 4). All three characteristic regimes described in the text can be clearly seen. Note that the evolution of the equivalent width with  $N_{\text{HI}}$  is qualitatively the same for all transitions, even though the demarcation column densities for each regime are different for different transitions. The curves-of-growth were computed by numerical integration assuming a Voigt profile and using the approximation to the Voigt-Hjerting function by [Tepper-García \(2006\)](#).



**Fig. 2.3:** Effect of the Doppler parameter on the equivalent width. The curve-of-growth of the Ly $\alpha$  line is shown here as an example, even though the behaviour is qualitatively the same for any absorption line. Note the insensitivity of the equivalent width to the Doppler parameter in the linear and square-root regimes, in contrast to the strong dependence in the flat part.



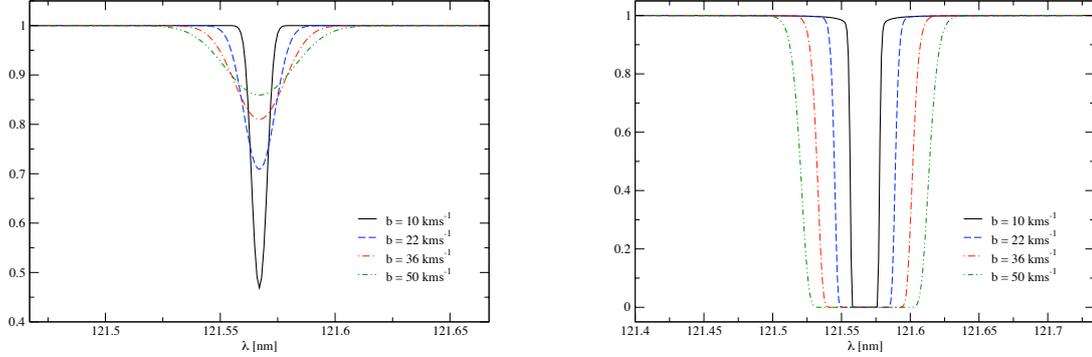
**Fig. 2.4:** Ly $\alpha$  Voigt profiles for different column densities and a Doppler parameter  $b = 36 \text{ km s}^{-1}$ . Shown are absorption line profiles for: (a)  $\log N_{\text{HI}} \in \{12.0, 12.5, 13.0, 13.5, 14.0, 14.5, 15.0, 15.5, 16.0\}$ , and (b)  $\log N_{\text{HI}} \in \{18.5, 19.0, 19.5, 20.0, 20.5\}$ . Note the difference in scale in the  $x$ -axis.

Fig. 2.4(a) the rapid increase in the line strength with column density for  $\log N_{\text{HI}} \in [12.0, 14.5]$  (linear part of the curve-of-growth), and how the equivalent width of the line does not change significantly after the line saturates, *i.e.* for  $\log N_{\text{HI}} \in [14.5, 16.0]$ . As can be seen in Fig. 2.4(b), the line strength increases again dramatically for  $N_{\text{HI}} \gtrsim 10^{18.5} \text{ cm}^{-2}$  (square-root regime), and the damping wings due to the high column density of the absorbing medium become observable.

The accurate determination of the column density and Doppler parameter of an observed absorption line is essential in order to draw conclusions about the physical state of the absorbing medium. This is of particular importance when analysing the absorption lines observed in quasar spectra, which are caused by the intergalactic neutral hydrogen encountered along the line-of-sight. In this respect, it is worth discussing the degeneracy of the curve-of-growth in the linear and square-root regime for a given transition and different Doppler parameters, as seen in Fig. 2.3. In the case of the Ly $\alpha$  absorption line, the curve-of-growth for the Doppler parameters  $b \in \{22, 36, 50\} \text{ km s}^{-1}$  are practically indistinguishable from each other in the linear regime, *i.e.* for  $N_{\text{HI}} \lesssim 10^{14.5} \text{ cm}^{-2}$ , as well as in the square-root regime. Note also that the lower the Doppler parameter, the smaller the column density at which the line saturates. Only for very low (and in the case of intergalactic H I unrealistically low) values of  $b$  can the different curves be distinguished from each other at  $N_{\text{HI}} \gtrsim 10^{13} \text{ cm}^{-2}$ . In the saturation regime, the

Doppler parameter has the effect of splitting the curves-of-growth, opening the possibility to distinguish between them, assuming that the column density is known and the equivalent width can be measured with high accuracy. This degeneracy is not welcome from the observational point of view, since it places a constraint on the accuracy with which the Doppler parameter and the column density can independently be measured.

A way out of this problem is offered by the accurate determination of the line profile, which is uniquely defined for each transition at given column density and given Doppler parameter. This can readily be seen on the shape of the line profile (equation 2.21). Hence, even though lines with different Doppler parameters and the same column density cannot be distinguished from each other by measuring their equivalent widths in the linear part, it should be possible to disentangle them on the basis of their profiles. In order to illustrate this, we show in Fig. 2.5(a) four different Ly $\alpha$  line profiles corresponding to  $b \in \{10, 22, 36, 50\} \text{ km s}^{-1}$  and  $N_{\text{HI}} = 10^{13} \text{ cm}^{-2}$ . All correspond to the linear regime of the curve-of-growth, and have the same equivalent width, but their profiles are clearly very different. As expected, the larger the Doppler parameter, the broader and shallower is the line profile. It becomes now easy to understand why lines arising in a medium with low temperatures, corresponding to low values of  $b$  (see equation 2.9), saturate at lower column densities. In the saturation regime, though, the profiles are qualitatively very similar, and they may be distinguished solely on the ba-



**Fig. 2.5:** Ly $\alpha$  line profiles for  $b \in \{10, 22, 36, 50\} \text{ km s}^{-1}$  and two different column densities: (a)  $\log N_{\text{HI}} = 13.0$  (linear regime), and (b)  $\log N_{\text{HI}} = 16.0$  (square-root regime).

sis of their widths. This can be seen in Fig. 2.5(b), where four different Ly $\alpha$  Voigt profiles corresponding to  $b \in \{10, 22, 36, 50\} \text{ km s}^{-1}$  and  $N_{\text{HI}} = 10^{16} \text{ cm}^{-2}$  are shown. A method widely used to determine the line parameters consists in using line fitting algorithms which on the basis of a  $\chi^2$  minimization technique explore a wide parameters space of the variables ( $N_{\text{HI}}$ ,  $b$ ) looking for the solution that best fits the given profile. This technique certainly demands an extremely high resolution of the order of  $\lambda/\Delta\lambda \leq 10^3$  at  $\lambda = 1 \text{ nm}$ , which is not easily achievable. An approach to a rough determination of line parameters even with lower resolution is to look for obvious observable features such as signs of saturation or damping wings. Making use of a curve-of-growth analysis, it should then be possible in principle to give a rough estimate of the line parameters. As will be mentioned in the following chapters, an accurate measurements of these quantities is required for the determination of the evolution of the intergalactic neutral hydrogen content and of the Ly $\alpha$  absorbers. Conversely, the uncertainty in the independent determination of  $b$  and  $N_{\text{HI}}$  is one of the main obstacles in our understanding of the physical state of the intergalactic medium. As an effort to contribute in some way to solve this problem, and especially in order facilitate the synthesis of line profiles, an approximation to the Voigt-Hjerting function was derived which will be presented in the following chapter.

---

# Voigt Profile Fitting to Quasar Absorption Lines: An Analytic Approximation to the Voigt-Hjerting Function

The contents of this chapter are originally published as [Tepper-García, T. 2006, MNRAS, 369, 2025](#).

---

## ABSTRACT

The Voigt-Hjerting function is fundamental in order to correctly model the profiles of absorption lines imprinted in the spectra of bright background sources by intervening absorbing systems. In this work we present a simple analytic approximation to this function in the context of absorption line profiles of intergalactic H<sub>I</sub> absorbers. Using basic calculus tools, we derive an analytic expression for the Voigt-Hjerting function that contains only fourth order polynomial and Gaussian functions. In connection with the absorption coefficient of intergalactic neutral hydrogen, this approximation is suitable for modeling Voigt profiles with an accuracy of  $10^{-4}$  or better for an arbitrary wavelength baseline, for column densities up to  $N_{\text{HI}} = 10^{22} \text{ cm}^{-2}$ , and for damping parameters  $a \lesssim 10^{-4}$ , *i.e.* the entire range of parameters characteristic to all Lyman transitions arising in a variety of H<sub>I</sub> absorbing systems such as Ly $\alpha$  Forest clouds, Lyman Limit systems and Damped Ly $\alpha$  systems. We hence present an approximation to the Voigt-Hjerting function that is both accurate and flexible to implement in various types of programming languages and machines, and with which Voigt profiles can be calculated in a reliable and very simple manner.

**keywords:** methods: analytic, quasars: absorption lines, line: formation, line: profiles, line: identification

## 3.1 INTRODUCTION

Absorption processes and their signatures (absorption lines) imprinted on the spectra of bright background sources (quasars, Gamma-ray bursts, etc.) are one of the main sources of information about the physical and chemical properties of intervening systems. It is well known that information about their temperature, density, chemical abundances, and kinematics can be extracted from the analysis of these absorption lines. For instance, a detailed insight into the physical state of the intergalactic medium (IGM) is provided by the analysis of the absorption lines found in the spectra of distant quasars (QSOs) (see *e.g.* [Hu et al. 1995](#), [Kim et al. 1997, 2001, 2002a,b](#)). These lines are due mainly to absorption by neutral hydrogen (H<sub>I</sub>) present in a class of low column density absorbers generally known as Ly $\alpha$  Forest, and due to other elements in low ionisation stages (CII, CIV, SiII, MgII, FeII, OII, etc.), which arise in higher column densities absorbing systems associated with galaxies, such as the Lyman Limit Systems (LLSs)

and Damped Ly $\alpha$  Absorbers (DLAs). A wealth of information about the distribution, density, temperature, metal content, etc. of these systems is now available as a result of exhaustive and extensive studies of QSO absorption lines (see *e.g.* Rauch 1998, Rao 2005, for a review on Ly $\alpha$  absorbers, and on metal systems and DLAs, respectively).

In this type of analysis, and within the realm of a given cosmological model, the number and observed central wavelength of the absorption lines provide information on the spatial distribution of the absorbing systems. Furthermore, knowledge about the actual physical state of these systems can be obtained basically from the line profiles. Both line counting and line profile measurement are tricky tasks though, since the accuracy with which they can be performed highly depends on the resolution of the observed spectra. For instance, depending on the spatial distribution of the absorbing systems, lines can appear very close to each other or even superpose (line blending), and a low spectral resolution may lead to the misidentification of the resulting composite profile as being a single, complex one. Because of this same reason, the determination of the exact shape of each individual absorption profile is far from being trivial, and misidentified profiles may lead to wrong conclusions about the properties of the absorbing systems.

If one assumes that the physical state of the absorbing medium is uniquely defined by its temperature and column density, single absorption line profiles are ideally described by Voigt profiles. Mathematically, a Voigt profile is given in terms of the convolution of a Gaussian and a Lorentzian distribution function, known as Voigt-Hjerting function (Hjerting 1938), and a constant factor that contains information about the relevant physical properties of the absorbing medium (cf. Sect. 3.2.1). Any departure from a pure Voigt profile in the observed lines is expected to yield information about the kinematic properties (non-thermal broadening, rotational or turbulent macroscopic motions), as well as spatial information (clustering) of the absorbing systems.

The Voigt-Hjerting function has long been known and, consequently, various numerical methods to estimate and tabulate this function have been developed and presented (Hjerting 1938, Harris 1948, Finn & Mugglestone 1965). Also, a great effort has been done in order to derive a semi-analytic approximation to this function that reproduces its behavior with high accuracy (see *e.g.* Whiting 1968, Kielkopf 1973, Monaghan 1971), being the latter by far the one with the highest accuracy. With

the help of these methods, computational subroutines have been developed that make it nowadays possible to numerically integrate this function for a wide parameter space (see *e.g.* Humlíček 1982).

The aim of this work is to make a further contribution to the practical handling of the Voigt-Hjerting function in order to compute Voigt profiles. Starting with an exact expression for this function in terms of Harris' infinite series, we argue why this series may be truncated to first order in  $a$  in the context of intergalactic H $\alpha$  absorption lines. We then show that the second term of this series can be approximated with a non-algebraic polynomial function, which is mathematically simple to handle in the sense that it does not contain singularities. Such an analytic, 'well-behaved' expression in terms of simple functions as presented here is very attractive, since it allows one to replace the many steps and operations needed for numerical integration, or to read from look-up tables of values, by a single line with simple operations. It is also extremely flexible to implement in various types of codes and machines, and is particularly useful for computational routines in higher-level programming languages (*e.g.* IDL, *Mathematica*, *Maple*, etc.), in which numerical integration or look-up table reading is cumbersome, especially if absorption line profiles have to be calculated many times with moderate precision and relative high speed. For instance, such an analytic expression should be very useful to synthesise Ly $\alpha$  absorption spectra as in *e.g.* Zhang *et al.* (1997), Richter *et al.* (2006), or in line-fitting algorithms like AUTOVP (Davé *et al.* 1997) or FITLYMAN (Fontana & Ballester 1995), used to obtain line parameters such as redshift, column density, and Doppler width from absorption Voigt profiles imprinted on observed spectra.

In the next section we briefly outline the origin of the Voigt-Hjerting function in the Physics of absorption processes, and define the context in which the desired approximation of this function is of interest to us. In Sections 3.3 and 3.4 we derive this approximation, and in Section 3.5 we compare the accuracy and speed of a numerical method for computing Voigt profiles based on our approximation to other existing methods. In Section 3.6 we present an application of our method to model Voigt profiles, and in Section 3.7 we summarise our main results.

### 3.2 THE VOIGT-HJERTING FUNCTION IN THE CONTEXT OF H<sub>I</sub> ABSORPTION LINES

#### 3.2.1 The Absorption Coefficient

The probability of a photon with an energy  $E = hc/\lambda$  to be absorbed within a gas with column density  $N$  and kinetic temperature  $T$ , also known as absorption coefficient, is given by

$$\tau_i(\lambda) = (C_i \cdot N_{\text{HI}} \cdot a) \cdot H[a, x(\lambda)], \quad (3.1)$$

where

$$C_i \equiv \frac{4 \sqrt{\pi^3} e^2 f_i}{m_e c \Gamma_i}$$

is a constant for the  $i$ th electronic transition caused by the photon absorption. Here  $m_e$  is the electron mass,  $f_i$  is the oscillator strength, and  $\Gamma_i$  the damping constant or reciprocal of the mean lifetime of the transition. The function  $H$  is the so-called **Voigt-Hjerting function** and is given by

$$H(a, x) \equiv \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(x-y)^2 + a^2} dy. \quad (3.2)$$

Let  $\lambda_i = hc/E_i$  be the resonant wavelength of the corresponding transition, and  $\Delta\lambda_D \equiv \frac{b}{c}\lambda_i$  the thermal or Doppler broadening, which defines a Doppler unit. Here, the Doppler parameter  $b$  is related to the kinetic temperature of the gas via  $b = \sqrt{2kT/m_p}$ , where  $k$  is the Boltzmann constant and  $m_p$  is the proton mass. It follows from these definitions that the damping parameter

$$a \equiv \frac{\lambda_i^2 \Gamma_i}{4\pi c \Delta\lambda_D}$$

quantifies the relative strength of damping broadening to thermal broadening, and that the variables  $x \equiv \frac{(\lambda - \lambda_i)}{\Delta\lambda_D}$  and  $y \equiv \frac{v}{b}$  are just the wavelength difference relative to the resonant wavelength in Doppler units, and the particle velocity in units of the Doppler parameter, respectively.

The particular form of the absorption coefficient (3.1) induces a characteristic absorption feature known as Voigt profile. Hence, the Voigt profile, and consequently the Voigt-Hjerting function, naturally arise in the process of absorption-line formation, when one assumes that the physical state of the absorbing medium is uniquely defined by its density and kinetic temperature. Generally speaking, it is the physical conditions what determines the shape of the absorbing features,

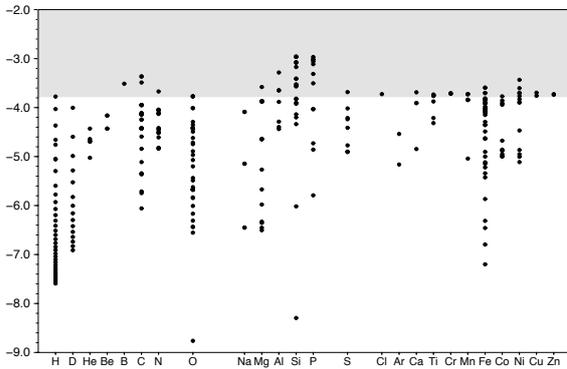
*i.e.* the line profiles. Conversely, it is true that line profiles give information about the physical state of the absorbing medium. In particular, line profiles other than Voigt profiles give information about the departure of the physical conditions assumed here.

The class of H<sub>I</sub> absorbers present in the IGM (Ly $\alpha$  Forest Clouds) and associated with galaxies and larger structures (LLSs, DLAs) can be characterised by their column density  $N_{\text{HI}}$  and kinetic temperature, and consequently their absorption features observed on *e.g.* QSO spectra are well described by Voigt profiles. Their observed column densities span a range of ten orders of magnitude, approximately from  $10^{12} - 10^{22} \text{ cm}^{-2}$ , and have temperatures that correspond to Doppler parameters approximately in the range  $10 - 100 \text{ km s}^{-1}$ , with a median value around  $b_m = 36 \text{ km s}^{-1}$  that decreases with redshift (Kim et al. 1997). For such a range in  $b$ , the damping parameter  $a$  for the Lyman transitions of intergalactic H<sub>I</sub> spans a range of  $9.3 \cdot 10^{-9} - 6.05 \cdot 10^{-4}$ . In this case, high values of  $a$  correspond to Ly $\alpha$ , while lower values are typical for higher order Lyman transitions. Transitions of other elements, such as C, Si, Mg, Fe, O, etc., and their various ionisation stages, are also typically found on QSO spectra, and their damping parameters cover a range which strongly overlaps with that of H<sub>I</sub>. This can be seen in Figure 3.1, where we show the distribution of  $a$  for different elements (including H<sub>I</sub>) in different ionisation stages, for a Doppler parameter  $b_{\text{HI}} = 36 \text{ km s}^{-1}$  for H<sub>I</sub>, and assuming that the doppler parameter for other elements is related to  $b_{\text{HI}}$  via  $b_X = b_{\text{HI}} \cdot \sqrt{m_{\text{HI}}/m_X}$ , where  $m_X$  is the mass of element  $X$ . For clarity, we have grouped all different ionisation stages of a given element under a single label. The values of the atomic constants (central wavelength of the transition and  $\Gamma$ -value) have been taken from Morton (2003)<sup>1</sup>. The shaded area contains the values of  $a$  above the range characteristic to intergalactic H<sub>I</sub> for which our approximation to the Voigt-Hjerting function—derived in the next sections—cannot be applied or should be applied with caution. Note, however, that the region spanned by  $a$  for intergalactic H<sub>I</sub>, *i.e.* the region underneath the shaded area, contains most of the  $a$ -ranges spanned by all other elements, especially those corresponding to C, O, Mg, and Fe. For other elements,

<sup>1</sup>A list containing the values of the damping parameters for the elements and their different ionisation stages shown in Figure 3.1, is available in plain-text format at [www.astro.physik.uni-goettingen.de/~tepper/hjerting/damping.dat](http://www.astro.physik.uni-goettingen.de/~tepper/hjerting/damping.dat). Please consult this list in order to know the exact value of  $a$  for a given element in a given ionisation stage.

such as Cr or Zn, the damping parameter has values right at the upper limit of this range.

For this reason and for the sake of simplicity, in the following we will constrain our discussion to the Lyman absorption lines of intergalactic H<sub>I</sub>, but the reader shall bear in mind that the the discussion and method to synthesise Voigt profiles presented in this work can be directly applied to the transitions of other elements associated with H<sub>I</sub> absorbing systems, according to Figure 3.1.



**Fig. 3.1:** Value of the damping parameter  $a$ , assuming  $b_{\text{HI}} = 36 \text{ km s}^{-1}$ , for different elements in different ionisation stages typically found in QSO spectra. The doppler parameter for other elements is assumed to be given by  $b_X = b_{\text{HI}} \cdot \sqrt{m_{\text{HI}}/m_X}$ , where  $m_X$  is the mass of element  $X$ . Elements are listed by increasing atomic mass on the  $x$ -axis, and the logarithmic value of  $a$  is given on the  $y$ -axis. For clarity, different ionisation stages of a given element have been grouped under a single label. The shaded area marks the range above the largest value of  $a$  for the intergalactic H<sub>I</sub> Lyman transitions. Note that for the sake of completeness, other elements than those found to date in QSO spectra have been included as well.

### 3.2.2 The Absorption Coefficient of H<sub>I</sub> at Low Column Densities

Following Harris (1948), it is true that for  $a < 1$ , *i.e.* when Doppler broadening dominates over damping broadening, the Voigt-Hjerting function (3.2) can be expressed as

$$H(a, x) = \sum_{n=0}^{\infty} H_n(x) a^n, \quad (3.3)$$

where the functions  $H_n(x)$  are defined by

$$H_n(x) \equiv \frac{(-1)^n}{\sqrt{\pi} n!} \int_0^{\infty} v^n e^{-(v/2)^2} \cos(xv) dv. \quad (3.4)$$

These functions are bounded with respect to  $n$  and  $x$ , and they have values of the order of unity. Indeed, taking the absolute value of the integral, neglecting the cosine, which takes values of the order of unity, and computing the resulting integral one can show that

$$|H_n(x)| \leq \frac{2}{\sqrt{\pi}} \approx 1.123, \quad (3.5)$$

for all  $n \in \{0, 1, 2, \dots\}$  and  $x \in \mathbb{R}$ . From this it follows that if  $a \ll 1$ , the Voigt-Hjerting function can very well be approximated to zeroth order in  $a$  by the first term of the series (3.3), *i.e.*  $H(a, x) \approx H_0(x)$ , as first noted by Walshaw (1955). Note that this result is *exact* in the limit  $a \rightarrow 0$ . Taking the definition (3.4), it follows that

$$H_0(x) = e^{-x^2}, \quad (3.6)$$

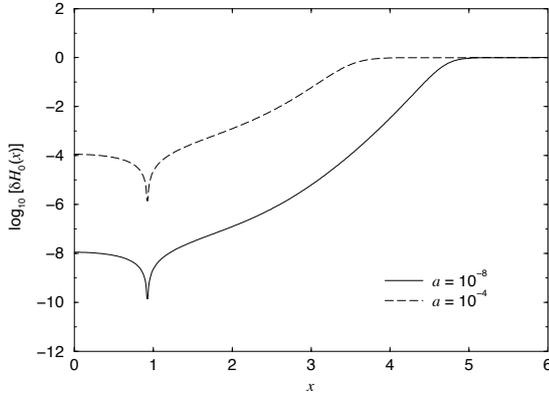
and thus  $H(a, x) \approx e^{-x^2}$ , for  $a \ll 1$ . We call this the Voigt-Hjerting function to *zeroth order*.<sup>2</sup>

But what actually means that the condition  $a \ll 1$  be satisfied, so that this zeroth order approximation can be safely used to model absorption line profiles? In order to address this, we compute for the extreme values of  $a$  for intergalactic H<sub>I</sub>,  $a \approx 10^{-8}$  and  $a \approx 10^{-4}$ , the departure of the Voigt-Hjerting function from a pure Gaussian function. This is shown in Figure 3.2 as the logarithmic difference between<sup>3</sup>  $H$  and  $H_0$  relative to  $H$  as a function of  $x$ , *i.e.* the quantity  $\delta H_0 \equiv 1 - H_0/H$ . Note how the zeroth order approximation completely differs from the actual Voigt-Hjerting function at  $x \gtrsim 3.5$  for  $a = 10^{-4}$ , and at  $x \gtrsim 4.5$  for  $a = 10^{-8}$ . Thus, even a value of  $a$  as small as  $10^{-8}$  is not a sufficient condition for the zeroth order term to be a good approximation of  $H$  for an arbitrary range in  $x$ .

In the context of the H<sub>I</sub> absorption coefficient, the condition on  $a$  for which the approximation to zeroth order of the Voigt-Hjerting function is valid, translates into a restriction on  $N_{\text{HI}}$ , and more specifically on the quantity  $C_i \cdot a \cdot N_{\text{HI}}$ . This can be seen by taking a glance at equation (3.1): Despite the fact that the condition  $a \ll 1$  holds, the product  $C_i \cdot a \cdot N_{\text{HI}}$  can be very

<sup>2</sup>A more detailed version of this section can be found in Appendix B

<sup>3</sup>The values of the Voigt-Hjerting function were computed numerically using a routine based on Monaghan's algorithm in Murphy (2002) (*cf.* Section 3.5).



**Fig. 3.2:** Departure of the Voigt-Hjerting function from a pure Gaussian function for  $a = 10^{-8}$  (solid line) and  $a = 10^{-4}$  (dashed line) as a function of  $x$ . The departure is given as the logarithmic difference between the Voigt-Hjerting function and the zeroth order approximation in the form  $\delta H_0 \equiv 1 - H_0/H$ . The greater this quantity, the less accurate is the zeroth order approximation.

large for high enough<sup>4</sup> column densities, and since the functions  $H_n$  are of the order of unity, such terms can have a significant contribution to the absorption coefficient. Along this line of reasoning, and taking into consideration that the constant  $C_i$  in equation (3.1) is of the order of  $10^{-11} \text{ cm}^2$  for all Lyman transitions, it is clear that the departure of the Voigt-Hjerting function from its zeroth order approximation becomes significant in the wavelength ranges  $x \gtrsim 3.5$  and  $x \gtrsim 4.5$  (in Doppler units) at column densities  $N_{\text{HI}} \gtrsim 10^{15} \text{ cm}^{-2}$  for a damping parameter  $a = 10^{-4}$ , and at column densities  $N_{\text{HI}} \gtrsim 10^{19} \text{ cm}^{-2}$  for  $a = 10^{-8}$ . Since intervening H<sub>I</sub> absorbers typically have column densities in the range  $10^{12} - 10^{22} \text{ cm}^{-2}$ , a Gaussian approximation to  $H$  for modeling intergalactic H<sub>I</sub> absorption line profiles is only suitable for the low end of the column density range.

In addition to the factor  $C_i \cdot a \cdot N_{\text{HI}}$  being large and even more decisive for the uselessness of the zeroth order approximation for an arbitrary range in  $x$ , is the fact that the zeroth order term,  $e^{-x^2}$ , rapidly decreases for large values of  $x$  and is therefore overwhelmed by higher order terms in Harris' expansion, which hence dominate the behavior of the absorption coefficient, as

<sup>4</sup>'High enough' means in this case that the condition  $N_{\text{HI}} > (C_i \cdot a)^{-1}$  is satisfied.

already seen in Figure 3.2. To shed some light on this, consider the following numerical example: Out to  $x \approx 4$  (in Doppler units), and for  $a = 10^{-4}$ , the Voigt-Hjerting function is of the order of  $4.2 \cdot 10^{-6}$ . The zeroth order term in the series (3.3) satisfies  $H_0(x=4) \approx 1.1 \cdot 10^{-7}$ , whereas the first order term  $(a \cdot H_1)(x=4) \approx 3.9 \cdot 10^{-6}$ . Thus, at large enough  $x$ , the behaviour of  $\tau$  is evidently governed by the terms of order  $n \geq 1$  in the series (3.3).

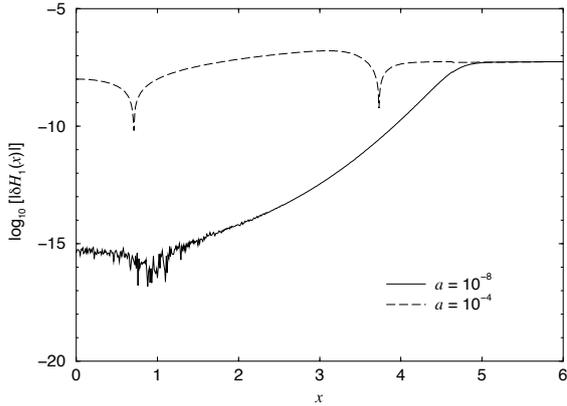
### 3.2.3 Higher Column Densities and First Order Term

Due to the arguments stated above, and even though the damping parameter satisfies  $a \ll 1$ , it is clear that the absorption coefficient of intergalactic H<sub>I</sub> cannot simply be approximated by a constant times  $e^{-x^2}$  for an arbitrary wavelength range and column densities greater than  $10^{15} \text{ cm}^{-2}$ . One actually has to take into account terms of higher order in the series (3.3), at least to first order in  $a$ , i.e.  $H(a, x) \approx (H_0 + a \cdot H_1)(x)$ . In fact, one should take into account all terms up to  $N$ th order for values of  $(C_i \cdot a \cdot N_{\text{HI}})^{-1}$  that are nearly equal or greater than the absolute difference between the sum  $\sum_{n=0}^N H_n(x) a^n$ , and the exact Voigt-Hjerting function. However, as we shall show next, the approximation to first order in  $a$  in Harris' expansion is enough to model absorption line profiles with moderate to high accuracy for the range of parameters  $(a, N_{\text{HI}}, C_i)$  characteristic to intervening H<sub>I</sub> absorbers.

In order to prove the above statement, we look at the relative contribution of the zeroth and first order terms to the Voigt-Hjerting function for  $a \in \{10^{-8}, 10^{-4}\}$  and an arbitrary range in  $x$ . This is achieved, for example, by computing the logarithm of the quantity  $\delta H_1 \equiv |1 - (H_0 + a \cdot H_1)/H|$  as a function of  $x$  for the two extreme values of  $a$ , as shown in Figure 3.3. Here we have advanced the function  $H_1$  which is being handled in the next section. Note that the greater the contribution from the zeroth and first order term to  $H$ , the smaller the quantity  $\delta H_1$ . In this case, as can clearly be seen,  $\delta H_1$  takes on values of the order of  $10^{-7}$  or less over the whole wavelength range shown here and for the whole range in  $a$  for intergalactic H<sub>I</sub>.

If one takes into account that  $H(x) \leq 1$  for all  $x$ , it is obvious that the relative difference is equal or greater than the *absolute* difference, i.e.  $\delta H_1 \geq |H - (H_0 + a \cdot H_1)|$ . Thus, in the case of  $a \approx 10^{-4}$  and  $\delta H_1 = 10^{-7}$ , the departure of  $H$  from its first order approximation becomes significant at column densities  $N_{\text{HI}} > (a \cdot \delta H_1 \cdot C_i)^{-1} = 10^{22} \text{ cm}^{-2}$ , and at even larger  $N_{\text{HI}}$  for  $a < 10^{-8}$  and/or smaller  $\delta H_1$ . Hence, the first

two terms of Harris' expansion dominate the behaviour of  $H$  over the whole wavelength range shown here, with an accuracy of  $10^{-7}$  or greater, for the range in  $a$  characteristic to intergalactic H $\alpha$ . On this basis, we consider that an approximation to first order in  $a$  of the Voigt-Hjerting function in terms of the functions  $H_0$  and  $H_1$  is suitable to model Voigt profiles. Since the function  $H_0$



**Fig. 3.3:** Contribution from the zeroth and first order terms of the series (3.3) to the absorption coefficient, when  $(C_l \cdot a \cdot N_{\text{HI}})$  is of the order of unity. The curves show the logarithmic difference between  $H$  and  $H_0 + a \cdot H_1$ , relative to  $H$  as a function of  $x$  for  $a = 10^{-8}$  (solid line) and  $a = 10^{-4}$  (dashed line). The smaller this difference, the greater the contribution from the zeroth and first order term to  $H$ . Note that the behavior of  $H$  is indeed dominated with a difference of seven or more orders of magnitude by the first two terms of Harris' expansion for the whole range in  $x$  shown here. The values for  $H_1$  were calculated according to equation (3.7), Section 3.3, using numerical integration to compute the function  $F$ .

is known and simple *per se*, we now turn to the task of finding an approximation to the function  $H_1$  in terms of a simple, analytic expression.

### 3.3 THE DAWSON FUNCTION REVISITED

According to definition (3.4) we have

$$H_1(x) = \frac{-4}{\sqrt{\pi}} \int_0^{\infty} v e^{-v^2} \cos(2xv) dv.$$

Integrating this equation partially, and computing the resulting Sinus transform of a Gaussian it follows (Mihalas 1970)

$$H_1(x) = \frac{-2}{\sqrt{\pi}} [1 - 2x F(x)], \quad (3.7)$$

where we adopt the notation first introduced by Miller & Gordon (1931)

$$F(x) \equiv e^{-x^2} \int_0^x e^{v^2} dv. \quad (3.8)$$

$F$  is known as the Dawson function (Dawson 1898).

It is evident that finding an approximation for  $H_1$  translates into the same problem for  $F$ . We thus want to show that a simple analytic expression can be found which approximates the Dawson function, and hence the function  $H_1$ , accurately enough, in order for an approximation of the Voigt-Hjerting function in terms of these functions to be useful for synthesising Voigt profiles.

#### 3.3.1 Properties of the Dawson Function

We briefly want to state some important properties of the Dawson-Function. First, this function is antisymmetric, *i.e.*  $F(-x) = -F(x)$  for all  $x \in \mathbb{R}$ . Because of this, from now on we restrict our analysis to  $x \geq 0$ . Besides, it has no roots in the positive semi-axis, and  $F(0) = 0$ , as can easily be seen from its definition (Fundamental Theorem of Calculus). Furthermore,  $F$  is bounded, since  $H_1$  is bounded itself (cf. Sect. 3.2.2). Indeed, differentiation with respect to  $x$  gives

$$\frac{d}{dx} F(x) = 1 - 2x F(x). \quad (3.9)$$

Hence, the upper bound is given by  $F(x_0) = (2x_0)^{-1}$ , where  $x_0$  is defined by equating  $F$ 's derivative to zero and solving for  $x$ . Actually,  $F$  has its maximum at  $x_0 = 0.92413$  with  $F(0.92413) = 0.54104$  (see *e.g.* Abramowitz & Stegun 1965). From this last equation it follows also that equation (3.7) can be rewritten as

$$H_1(x) = \frac{-2}{\sqrt{\pi}} \frac{d}{dx} F(x). \quad (3.10)$$

We want to know how the function  $F$  behaves asymptotically, *i.e.* for  $x \ll 1$  as well as for  $x \gg 1$ . Using the power series of the exponential we get

$$F(x) = e^{-x^2} \int_0^x \sum_{n=0}^{\infty} \frac{1}{n!} v^{2n} dv = e^{-x^2} \cdot \sum_{n=0}^{\infty} \frac{1}{n!} \frac{x^{2n+1}}{2n+1}. \quad (3.11)$$

In this case, the sum and the integral operator commute, since the power series of the exponential converges uniformly in any interval  $[a, b]$ , particularly for  $v \in [0, x]$  (see *e.g.* Forster 1983). Expressing the term  $e^{-x^2}$  by its

corresponding power series and rearranging terms, this last equation reads in explicit form

$$F(x) = x \cdot \left(1 - \frac{2}{3}x^2 + \frac{4}{15}x^4 + \dots\right).$$

Thus, for  $x \ll 1$  the Dawson function behaves asymptotically up to third order as

$$F(x) \approx x \cdot \left(1 - \frac{2}{3}x^2\right), \quad x \ll 1. \quad (3.12)$$

In order to investigate how  $F$  behaves for  $x \gg 1$ , we first rewrite the function (3.8) with the replacement  $v' = x - v$  and the aid of the power series of the exponential as

$$F(x) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(x), \quad (3.13)$$

with the definition

$$I_n(x) \equiv \int_0^x v^{2n} e^{-2xv} dv. \quad (3.14)$$

It is not hard to see that every term of the series (3.13) is separately bounded with respect to  $x$  as well as  $n$ . Making the replacement  $v' = 2vx$  in eq. (3.14), and integrating for  $x \neq 0$  we get for  $n \in \mathbb{N}_0$

$$I_n(x) = \frac{1}{(2x)^{2n+1}} (2n)! (1 - e^{-2x^2}) - e^{-2x^2} R_n(x), \quad (3.15)$$

where  $R_n(x)$  is a rather cumbersome polynomial function. Now, for  $x \gg 1$ , we may drop all terms which contain an exponential factor and in this way we get the asymptotic form

$$F(x) \approx \sum_{n=0}^{\infty} \frac{1}{(2x)^{2n+1}} (2n)!, \quad x \gg 1. \quad (3.16)$$

It can be seen from this expression that  $F$  vanishes as  $(2x)^{-1}$  for  $x \rightarrow \infty$ , and that the first derivative (3.9), and thus the function  $H_1$ , also vanish in this limit as  $(2x^2)^{-1}$ . From equation (3.9) it is also true that  $F'$  converges to unity for  $x \rightarrow 0$  and that  $H_1$  converges to  $(-2/\sqrt{\pi})$  in this limit. Since we want our approximation to  $F$ , and consequently to  $H_1$  and  $H$ , to be valid in the whole range  $x \in [0, \infty)$ , we require it to fulfill both these conditions as well.

### 3.4 THE ANALYTIC APPROXIMATION $D_1$

Equation (3.13), together with eq. (3.14), represent indeed an exact expression for the Dawson function.

However, these expressions are not suitable for practical computation. We therefore explore the possibility of finding an analytic expression which is easy to handle and which can be used to compute the value of  $F(x)$  for  $x \in [0, \infty)$ . In particular, we shall see if it is possible to truncate the series (3.13) in order to find an approximation to  $F$ , which has all its properties (antisymmetry, boundedness, etc.), which converges for  $x \rightarrow \infty$  as well as for  $x \rightarrow 0$ , and which is well defined in the whole range  $[0, \infty)$ . For instance, equations (3.12) and (3.16) do not fulfill these requirements. Nevertheless, they show us how our desired function has to behave asymptotically.

Let us define

$$D_N(x) \equiv \sum_{n=0}^N \frac{1}{n!} I_n(x). \quad (3.17)$$

where the  $I_n$ 's are given by eq. (3.14). Using this definition we get

$$D_1(x) = (1 - e^{-2x^2}) \cdot \left[ \frac{1}{2x} + \frac{1}{4x^3} \right] - e^{-2x^2} \cdot \left[ \frac{x}{2} + \frac{1}{2x} \right]. \quad (3.18)$$

It is easy to show that this function behaves qualitatively in the same way as  $F$  does, *i.e.* it is antisymmetric, bounded, and has no roots in the positive semi-axis. Furthermore, both these functions have the same asymptotical behavior. Indeed, up to third order we have for  $x \ll 1$

$$D_1(x) \approx x \cdot \left(1 - \frac{2}{3}x^2\right), \quad x \ll 1. \quad (3.19)$$

as can be shown by expanding the exponentials in eq. (3.18) in terms of their power series. A glance at eq. (3.12) makes the similarity between  $F$  and  $D_1$  evident in this limit. For  $x \gg 1$  we get from eq. (3.18), neglecting the exponentials,

$$D_1(x) \approx \frac{1}{2x} + \frac{1}{2^2 x^3}, \quad x \gg 1,$$

that is the same as for  $F$  (see eq. 3.16). Furthermore, the first derivative of  $D_1$  with respect to  $x$  is unity at  $x = 0$  and vanishes as  $(2x^2)^{-1}$  for  $x \rightarrow \infty$ . However, the maximum of  $D_1$  is at  $x = 0.87269$  with  $D_1(0.87269) = 0.52212$ , *i.e.* at slightly different values from those of  $F$ .

The magnitude and range of the error in approximating the Dawson function by the function  $D_1$  can be estimated, at least qualitatively, in the following way: From

equations (3.13) and (3.17) it follows that

$$F(x) = D_1(x) + \sum_{n=2}^{\infty} \frac{1}{n!} I_n(x). \quad (3.20)$$

The sum in this last equation, *i.e.* the absolute error in our approximation  $F(x) \approx D_1(x)$ , is a positive semi-definite quantity, since each term of the sum has this property. Hence, it is true that  $0 \leq D_1(x) \leq F(x)$  for  $x \in [0, \infty)$ . Furthermore, since both  $F$  and  $D_1$  converge to zero for  $x \rightarrow 0$  as well as for  $x \rightarrow \infty$ , and both these functions are bounded, it follows from eq. (3.20), that the error also vanishes for  $x \rightarrow 0$  as well as for  $x \rightarrow \infty$ , that it is also bounded, and that its maximum value is less equal than  $\max\{F(x) - D_1(x)\}_{x>0}$ . From this, and since their respective maxima are slightly shifted with respect to each other, one is led to the conclusion, that the error is constrained to a narrow range in  $x$  and that the maximum error in our approximation to  $F$  occurs near the maxima of these functions, *i.e.* in the vicinity of  $x = 1$ .

We went further try to quantify the actual error in our approximation to  $F$ . It shall be enough to know, for our purposes of finding an approximation to the Voigt-Hjerting function, that the error in the approximation  $F(x) \approx D_1(x)$  is bounded and constrained to a narrow wavelength interval around  $x = 1$ . Besides, we will indirectly estimate the error in this approximation when quantifying the error in our approximation to  $H$  in Section 3.5.

### 3.4.1 The Voigt-Hjerting function to First Order

Once we have found an approximation to the Dawson function, we can use it to give the desired expression for the Voigt-Hjerting function using Harris' expansion to first order in  $a$ . Replacing in equation (3.10) the function  $F$  by our approximation  $D_1$  (eq. 3.18), taking the corresponding derivative, and rearranging terms we get

$$H_1(x) \approx \frac{-2}{\sqrt{\pi}} K(x) e^{-x^2}, \quad (3.21)$$

where we have defined

$$K(x) = \frac{1}{2x^2} \left[ (4x^2 + 3)(x^2 + 1) e^{-x^2} - \frac{1}{x^2} (2x^2 + 3) \sinh x^2 \right]. \quad (3.22)$$

We want to highlight the fact that equation (3.21) is well defined, *i.e.* it has no singularities in the whole interval

$[0, \infty)$ . Furthermore, it converges to the correct value in the limits  $x \rightarrow 0$  and  $x \rightarrow \infty$ . Indeed, it is easy to show that  $\lim_{x \rightarrow 0} H_1(x) = \frac{-2}{\sqrt{\pi}}$ , and  $\lim_{x \rightarrow \infty} H_1(x) = 0$ , whether one uses for  $H_1$  the exact expression (3.7) or the approximation (3.21). In contrast, in approximations to the Voigt-Hjerting function to model Voigt profiles often used in the literature (see *e.g.* Spitzer 1978, Zhang et al. 1997) and given in the form  $c_1 \cdot e^{-x^2} + c_2 \cdot \frac{1}{x^2}$ , where the  $c_i$ 's are constants, the second term which represents the Lorentzian damping clearly diverges for  $x \rightarrow 0$ , and one has to artificially define the wavelength range in which this second term is used. It is customary to neglect this term for low column densities and in the vicinity of  $x = 0$ , but how to exactly choose the radius of the vicinity is not clear and completely arbitrary. However, with an expression like equation (3.21) at hand, no such assumptions have to be made.

Taking into account that  $a \ll 1$  in order to neglect terms of order  $n \geq 2$  in the series (3.3), we get, using the expressions for  $H_0$  (eq. 3.6) and  $H_1$  (eq. 3.21), that the Voigt-Hjerting function to first order in  $a$  is given by

$$H(a, x) \approx e^{-x^2} \left[ 1 - a \frac{2}{\sqrt{\pi}} K(x) \right]. \quad (3.23)$$

This expression is symmetric in  $x$ , as it should be, and thus it is valid for  $x \in \mathbb{R}$  and  $a \ll 1$ . According to this equation, the Voigt-Hjerting function can be regarded as a "corrected" Gaussian function, where the correction term depends on the parameter  $a$ . In the context of the absorption coefficient of H<sub>I</sub>, this correction term also depends on the column density  $N_{\text{HI}}$ , of course, via the quantity  $a \cdot N_{\text{HI}}$ .

## 3.5 ANALYSIS

In order to quantify the quality of our approximation to  $H$ , we perform a test on speed as well as on precision, comparing a numerical method to compute  $H$ , based on our approximation, to other standard, available methods to numerically compute this function. For this purpose, we use the approach and corresponding computational routine developed by Murphy (2002), which consists of the numerical implementation in FORTRAN of four different methods to compute  $H$ : Harris'  $H1$  and  $H2$ , Humlíček's, and Monaghan's. In Murphy's notation, Harris'  $H1$  and  $H2$  correspond to the Voigt-Hjerting function approximated by the first three and five terms of the series expansion (3.3), respectively. Humlíček's optimized algorithm and Monaghan's differential approach to approximate  $H$  are explained in

detail in Humlíček (1982) and Monaghan (1971), respectively. Our method to compute  $H$  consists simply in the numerical implementation in FORTRAN<sup>5</sup> of equations (3.22) and (3.23).

### 3.5.1 Speed

Following Murphy's approach, the relative speed of all five methods are determined by calculating the time that a routine based on each method requires to compute the Voigt-Hjerting function for  $x \in [0, 10]$  and damping parameters  $a$  in the range  $10^{-8} - 10^{-4}$  for a total of  $1.5 \cdot 10^7$  runs. In this way, we get that the relative speed<sup>6</sup> of each method in the order  $H1$  : (this work) : Humlíček :  $H2$  : Monaghan corresponds to 1 : 4.2 : 5.8 : 6.8 : 66.6, independent of  $a$ . According to this result, our method is second fastest.

### 3.5.2 Precision

The precision of our method relative to the other methods mentioned above is determined in the following way: First, the values of the Voigt-Hjerting function computed using Monaghan's algorithm for  $x \in [0, 10]$  and  $a \in [10^{-8}, 10^{-4}]$  are taken as fiducial. Then, the value of  $H$  for the same range in  $x$  and  $a$  is computed using each of these methods, and the logarithmic difference between each method and its fiducial value, *i.e.* the quantity  $\log_{10} \delta H$  with  $\delta H \equiv 1 - H_{\text{method}}/H_{\text{Mon}}$ , is calculated as a function of  $x$  for each different  $a$ . The result is shown in Figure 3.4. It can be seen from this figure that Harris'  $H2$  is the second most precise method to compute  $H$ , if one takes Monaghan's algorithm as fiducial, but is six times slower than Harris'  $H1$ , and one-and-a-half times slower than our method, as stated in the previous section. Our method has a precision of  $10^{-4}$  or better for  $x \lesssim 4$ . For all values of  $a$ , the difference peaks around  $x = 4$  to a value of the order of 0.01, and the precision increases again for values of  $x > 4$ . The precision is better for smaller  $a$ , as expected, since the zeroth order term gains in importance in our approximation for decreasing  $a$ . For  $a \lesssim 10^{-6}$  and  $x < 1.5$ , our

<sup>5</sup>The rearrangement of equations (3.22) and (3.23) that leads to the smallest number of operations reads, in code syntax,

$$H(a, x) = H_0 - a / \sqrt{\pi} / x^2 \cdot [H_0 \cdot H_0 \cdot (4 \cdot x^2 \cdot x^2 + 7 \cdot x^2 + 4 + Q) - Q - 1]$$

where the terms  $x^2$ ,  $H_0 \equiv e^{-x^2}$  and  $Q \equiv 1.5 \cdot x^{-2}$  have to be computed just once.

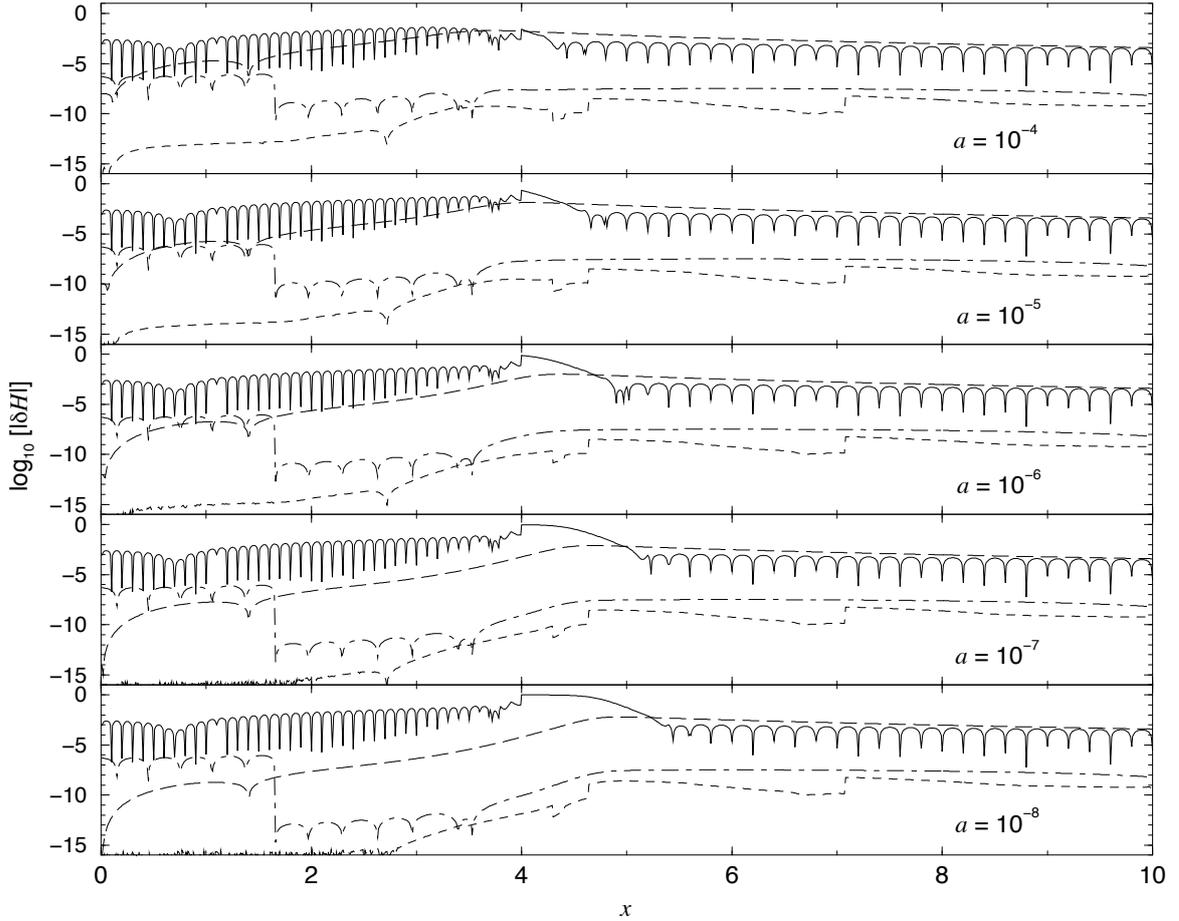
<sup>6</sup>This calculations were performed on an Intel Xeon 3.2 GHz processor.

method is more precise than Harris'  $H1$  or Humlíček's, and, as seen above, 1.5 times faster than the latter.

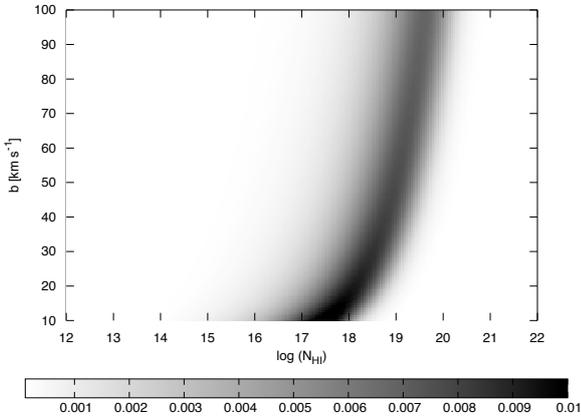
### 3.5.3 Modeling of H $\alpha$ absorption profiles

We now turn to analyse how accurate is our method in order to model H $\alpha$  absorption profiles. Taking the whole range in column density  $\log N_{\text{HI}} \in [12.0, 22.0]$  dex and Doppler parameters  $b \in [10.0, 100.0]$  km s<sup>-1</sup> characteristic to intergalactic H $\alpha$ , we synthesise for each pair  $(N_{\text{HI}}, b)$  (with a resolution of  $\Delta \log N_{\text{HI}} = 0.05$  dex, and  $\Delta b = 0.5$  km s<sup>-1</sup>) a single Ly $\alpha$  absorption profile in the range  $\lambda \in [100, 2300]$  Å with a resolution of  $\Delta \lambda = 0.01$  Å. The absorption profile is synthesised according to equations (3.1) and (3.2), using both our method and Monaghan's to compute  $H$ . We then compute for each pair  $(N_{\text{HI}}, b)$  the absolute value of the difference between the profiles generated using these two methods relative to Monaghan's, *i.e.*  $\delta V \equiv |1 - e^{-\Delta \tau}|$ , with  $\Delta \tau \equiv \tau_{\text{our}} - \tau_{\text{Mon}}$ , as a function of wavelength for the whole wavelength range, and pick the *maximum* value of this difference in the *entire* range. We choose to do so in order to pick up the worst cases possible, *i.e.* those with the lowest accuracy, and put in this way a stringent lower limit to the accuracy of our method. The result is shown as a contrast diagram on the  $(\log N_{\text{HI}}, b)$ -plane in Figure 3.5. Note that, in this case, it is not the logarithmic, but the *linear* difference which is shown. The highest precision is of the order of  $10^{-16}$  or even better. However, for the sake of simplicity, any value below  $10^{-4}$  has been coded as zero. The largest discrepancy between both methods, *i.e.* the lower limit in the precision of our method if one takes Monaghan's as fiducial, amounts to 0.01, in agreement with the result shown in Figure 3.4. As can clearly be seen, the (lower limit in the) precision of our method depends both on  $b$  and  $N_{\text{HI}}$ . While the dependence on  $b$  extends to the whole range  $10.0 - 100.0$  km s<sup>-1</sup>, the dependence on  $N_{\text{HI}}$  is limited to the range  $\log N_{\text{HI}} = 16 - 20$  dex. For a fixed  $b$ , there is a regime of values around 0.01 with a width of nearly 2 dex, which gives as a result a narrow region of low accuracy all across the plane. Along this stripe, the difference reaches its highest value of the order of 0.01 or smaller for the combinations (low  $b$ , low  $N_{\text{HI}}$ ) or (high  $b$ , high  $N_{\text{HI}}$ ), in the  $b$ - and  $N_{\text{HI}}$ -ranges stated above. Outside this stripe, the accuracy increases dramatically to values of the order of  $10^{-4}$  or even better.

The origin of any inaccuracy in our method is obviously the fact that terms of order  $n \geq 2$  have been neglected in the series (3.3), and furthermore, that the sec-



**Fig. 3.4:** Precision of different methods to compute the Voigt-Hjerting function, relative to Monaghan's differential method. Shown here is the logarithmic difference as a function of  $x$ , *i.e.* the quantity  $\log_{10} \delta H$ , with  $\delta H \equiv 1 - H_{method}/H_{Mon}$ , where  $H_{Mon}$  and  $H_{method}$  are, respectively, the Voigt-Hjerting function computed using Monaghan's algorithm and each of the methods mentioned in the text: H1 (solid line), H2 (short-dashed line), Humlíček (dot-dashed line), this work (long-dashed line). Each panel corresponds to a different damping parameter. Here we chose the range of  $a$  characteristic to intergalactic HI, *i.e.*  $a \in [10^{-8}, 10^{-4}]$ . This graph was created using the approach and corresponding routine developed by [Murphy \(2002\)](#), and is adapted from Figure A.1 of the same work.



**Fig. 3.5:** *Worst Scenario:* Lower limit to the precision of our method to synthesise Voigt profiles, given as the maximum value of the difference  $\delta V$  (cf. text for definition) between a Ly $\alpha$  absorption profile computed according to equation (3.1) using our approximation to the Voigt-Hjerting function (eq. 3.22 and 3.23) and Monaghan’s algorithm, for the whole range of values for the parameters  $(\log N_{\text{HI}}, b)$  characteristic to intergalactic HI. The value corresponding to each pair  $(\log N_{\text{HI}}, b)$  is the *maximum* value of the quantity  $\delta V$  in the *entire* wavelength range  $\lambda \in [100, 2300]$  Å. For clarity, any value below  $10^{-4}$  has been coded as zero.

ond term is this series has been approximated as well. In particular, the origin of the ‘low-accuracy’ stripe on the  $(\log N_{\text{HI}}, b)$ -plane can qualitatively be understood in terms of the functional dependence of  $\delta V$  on  $b$  and  $N_{\text{HI}}$ . Considering that both our method and Monaghan’s take the zeroth order term exactly into account, it is legitimate to state that the quantity  $\Delta H \equiv H_{\text{our}} - H_{\text{Mon}}$  has a least a dependence of first order on  $a$ , *i.e.*  $\Delta H(a, x) = a \cdot h(a, x)$  where  $h$  is a function that may be of zeroth order in  $a$ . Hence, using equation (3.1) and this last expression, it follows that

$$\Delta\tau(\lambda) = C'_i \cdot \frac{N_{\text{HI}}}{b^2} \cdot h(a, x),$$

where

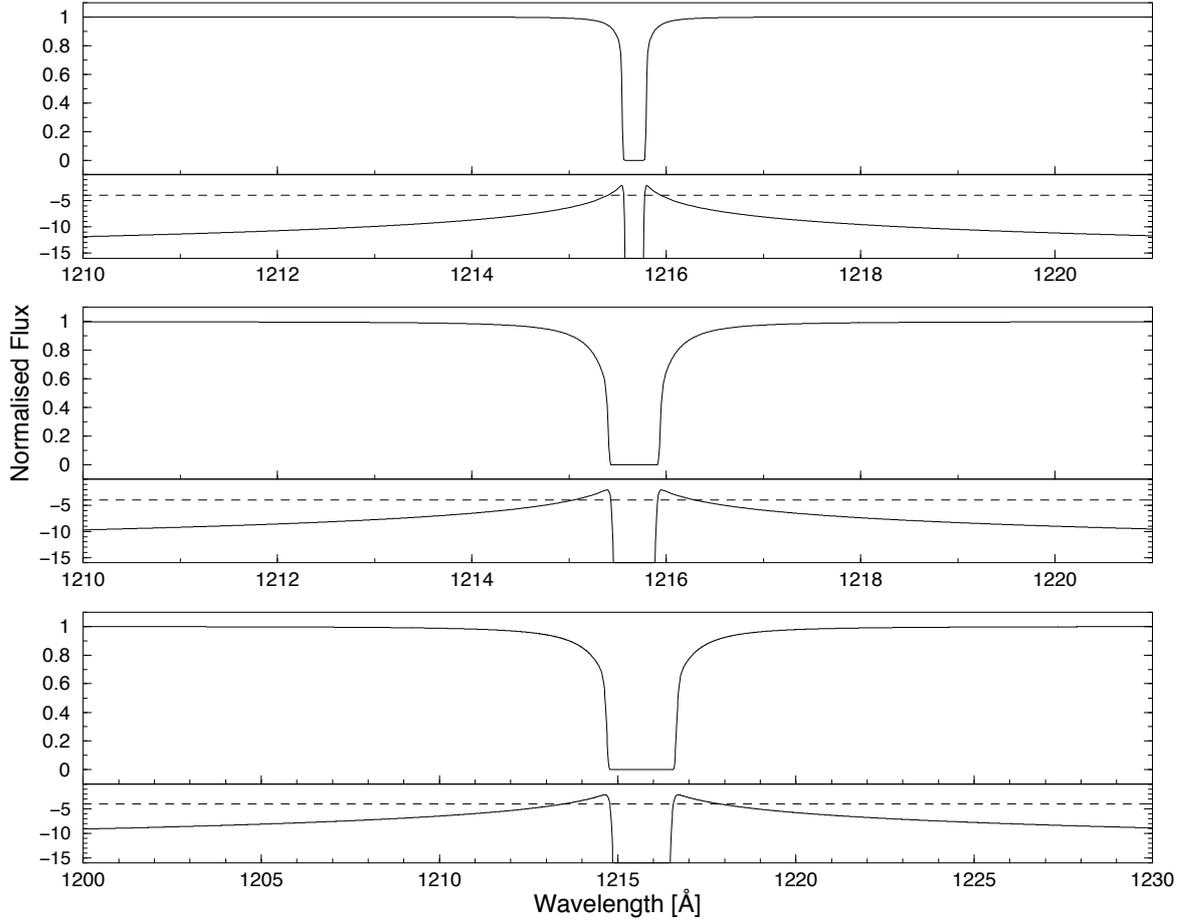
$$C'_i \equiv \frac{e^2}{4\sqrt{\pi}m_e c} \lambda_i^2 f_i \Gamma_i.$$

As can be seen,  $\tau$  strongly depends on the ratio  $\frac{N_{\text{HI}}}{b^2}$ . Therefore, an increase in  $b$  of one order of magnitude (from 10 to 100 km s $^{-1}$ ) is nearly compensated (in the sense that the value of  $\tau$  remains nearly constant) by an increase in  $N_{\text{HI}}$  of two orders of magnitude, which

accounts for the stripe of 2 dex in column density seen on the  $(\log N_{\text{HI}}, b)$ -plane. Why this happens precisely between  $\log N_{\text{HI}} = 16 - 20$  dex, as well as the shape of this stripe, is non-trivially related to the exact value of the constant  $C'_i$ , the fact that  $h$  may depend also on  $b$  through the damping parameter  $a$ , and the fact that the  $\Delta V$  depends effectively not on  $\Delta\tau$ , but on  $e^{-\Delta\tau}$ .

It is worth mentioning that, for higher-order Lyman transitions, the precision of our method to compute Voigt profile has the same behaviour on the  $(\log N_{\text{HI}}, b)$ -plane, and is the same as or even better than the precision of the Ly $\alpha$  transition shown here. The reasons for this are, first of all, that the functional form of the absorption coefficient is obviously the same, irrespective of the transition. In addition, the lowest precision possible of 0.01 is the same for the whole range in  $a$  spanned by the Lyman transitions, according to Figure 3.4. Furthermore, higher transitions have lower damping parameters and our approximation is better the lower  $a$ , as already mentioned. Finally, since the constant  $C'_i$  is smaller the higher the order of the transition, the critical range of lowest precision is shifted to higher column densities and higher Doppler parameters. Since the ranges in  $N_{\text{HI}}$  and  $b$  are fixed for intergalactic HI, this has the net effect of increasing the high-precision region on the  $(\log N_{\text{HI}}, b)$ -plane—*i.e.* the region to the left of the stripe—for higher order transitions. In other words, the precision of our method improves from Ly $\alpha$  to higher Lyman transitions.

Even though a discrepancy of the order of 0.01 when using our method to model Voigt profiles may seem large, we want to emphasise again that this is merely an *lower* limit for the precision in the *entire* wavelength range  $\lambda \in [100, 2300]$  Å, in the case of Ly $\alpha$ . It turns out that the range in wavelength for which the accuracy is lowest is negligible for practical purposes. In order to show this, we first choose three points on the  $(\log N_{\text{HI}}, b)$ -plane along the stripe of lowest precision, *i.e.* for which the maximum difference is largest. Using these parameters, we synthesise Ly $\alpha$  absorption profiles using our method and Monaghan’s, and compute again for each of these ‘‘worst scenarios’’ the quantity  $\delta V$  for the whole wavelength range. The result is shown in Figure 3.6. Each row corresponds, from top to bottom, to the parameter pairs  $(\log N_{\text{HI}}, b) = (17.0, 10.0)$ ,  $(18.0, 20.0)$ , and  $(19.0, 70.0)$  chosen along the low-precision stripe. The upper panel of each row shows the Ly $\alpha$  absorption profile synthesised using our method, and the corresponding lower panel gives the logarithmic difference  $\log_{10} \delta V$  between our method and Mon-



**Fig. 3.6:** Examples of *Worst Scenarios*: Ly $\alpha$  absorption line profile synthesised using our method (upper panel in each row) and corresponding logarithmic difference with respect to our fiducial function (lower panel in each row) for the parameter pairs  $(\log N_{\text{HI}}, b) = (17.0, 10.0)$ ,  $(18.0, 20.0)$ , and  $(19.0, 70.0)$ , from top to bottom, for which the maximum discrepancy with respect to the fiducial value of  $H$  in Figure 3.5 is largest. The maximum difference of the order of 0.01 shown in Figure 3.5 amounts to an extremely narrow wavelength range between Gaussian core and Lorentzian wings. Away from these ranges, the accuracy improves significantly. For reference, and for comparison with Figure 3.5, a line corresponding to a constant logarithmic difference of -4 dex has been included in each of the lower panels.

Monaghan's as a function of wavelength. As can be seen, the smallest discrepancies are given at the line cores, as expected, since in this regime the zeroth order term dominates and both our approach and Monaghan's exactly take this term into account. The largest discrepancies, of the order of 0.01, are present in an extremely narrow range of  $\Delta\lambda \approx 0.06 \text{ \AA}$  for  $(\log N_{\text{HI}}, b) = (17.0, 10.0)$ , of  $\Delta\lambda \approx 0.17 \text{ \AA}$  for  $(\log N_{\text{HI}}, b) = (18.0, 20.0)$ , and of  $\Delta\lambda \approx 0.53 \text{ \AA}$  for  $(\log N_{\text{HI}}, b) = (17.0, 20.0)$ . This discrepancy is found at the boundaries between Gaussian core and Lorentzian wings, due to the fact that our method neglects terms of order  $n \geq 2$ , which dominate the behaviour of  $H$  in that regime. Note, however, that the difference rapidly drops with increasing distance (in  $\text{\AA}$ ) from the line center to values of the order of *e.g.*  $10^{-10}$  at a distance  $\Delta\lambda \approx 5 \text{ \AA}$  for the first two rows, and  $\Delta\lambda \approx 15 \text{ \AA}$  for last row. Hence, the effective accuracy of our method is far better than 0.01 in the wavelength range shown here. For the same reason mentioned in the last paragraph, the precision for higher order Lyman transitions is of the same order as or even better than for the Ly $\alpha$  transition shown here.

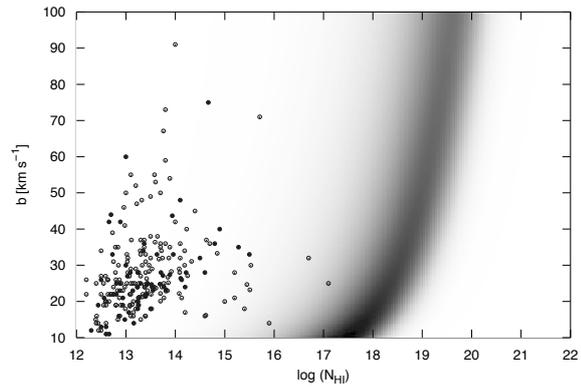
### 3.6 APPLICATION

As a further test of the quality of our approximation to model Voigt profiles, and to show its accuracy in a less academic situation as in the last section, we consider fitting a synthetic spectrum to a real quasar absorption spectrum with a population of intergalactic H I absorbers spanning a representative range in  $b$  and  $N_{\text{HI}}$  along a random line-of-sight. For this purpose we use the observed spectrum of the quasar source QSO J2233-606, a relatively bright ( $B = 17.5$ ) quasar at an intermediate redshift  $z_{\text{em}} = 2.238$ .

The spectrum of the source QSO J2233-606, centered at the HDF-S, was obtained during the Commissioning of the UVES instrument at the VLT Kueyen Telescope and reduced at the Space Telescope European Coordinating Facility. The high-resolution spectroscopy ( $R \approx 45000$ ) was carried out with the VLT UV-Visual Echelle Spectrograph (UVES). The data were reduced in the ECHELLE/UVES context available in MIDAS. The final combined spectrum has constant pixel size of  $0.05 \text{ \AA}$  and covers the wavelength range 3050-10000  $\text{\AA}$ . The S/N ratio of the final spectrum is about 50 per resolution element at 4000  $\text{\AA}$ , 90 at 5000  $\text{\AA}$ , 80 at 6000  $\text{\AA}$ , 40 at 8000  $\text{\AA}$ . The data used here are publicly available and were retrieved from

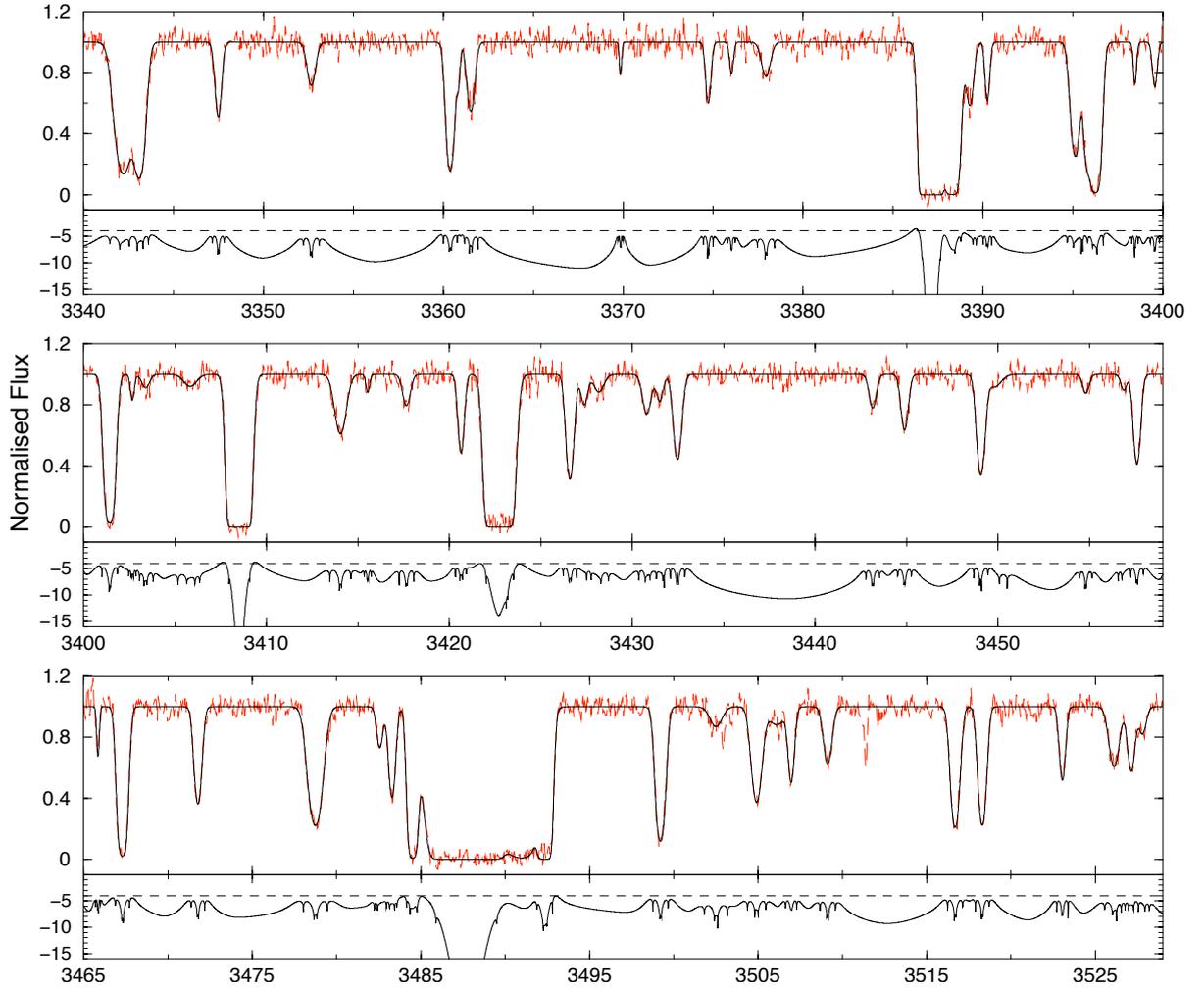
[www.stecf.org/hstprogrammes/J22/J22.html](http://www.stecf.org/hstprogrammes/J22/J22.html), in its version of November 23, 2005.

With help of the MIDAS package FITLYMAN (Fontana & Ballester 1995), which performs line fitting through  $\chi^2$  minimization of Voigt profiles, Cristiani & D'Odorico (2000) determined the redshifts, column densities, and Doppler widths of the identified absorption features imprinted in the spectra of QSO J2233-606. In this way, they found that the line of sight to QSO J2233-606 intersects a total of 270 Ly $\alpha$  Forest clouds, and identified other 24 absorption systems containing metal lines. The Ly $\alpha$  absorbers span a range in column density of  $10^{12.20} - 10^{17.10} \text{ cm}^{-2}$ , and a range in Doppler parameters of  $1.0 - 111.0 \text{ km s}^{-1}$ . The distribution of the parameter pairs  $(N_{\text{HI}}, b)$  for these systems for  $b \in [10, 100] \text{ km s}^{-1}$  is shown in Figure 3.7. Using this list of Ly $\alpha$  line parameters ( $z_{\text{abs}}, b, N_{\text{HI}}$ ),



**Fig. 3.7:** Distribution in column density and Doppler parameter of the absorbing systems along the line-of-sight towards the source QSO J2233-606. This particular LOS contains a total of 264 absorbing systems with  $b \in [10, 100] \text{ km s}^{-1}$  in the range 3050 – 10000  $\text{\AA}$  (open circles), 69 from which are in the wavelength range 3340 – 3530  $\text{\AA}$  (filled circles). The data was taken from Cristiani & D'Odorico (2000). For reference and comparison, the contrast diagram shown in Figure 3.5 has also been included.

we generate a synthetic spectrum of the QSO J2233-606 in the wavelength range 3340 – 3530  $\text{\AA}$  with a higher resolution than that of the observed spectrum of  $\Delta\lambda = 0.01 \text{ \AA}$ , using eqs. (3.1), (3.21), (3.22), and (3.23). We synthesise a second spectrum using Monaghan's algorithm, and compute the logarithmic difference  $\delta V$  between these two synthetic spectra in the same fashion as in the previous section. In this way, we test again our method against the highest-precision method



**Fig. 3.8:** Observed (dashed line) and synthetic (solid line) spectrum of the quasar HDF-S QSO J2233-606 (upper panel in each row) in the wavelength range [3340, 3530] Å (cf. text for reference). The synthetic spectrum was generated using our approximation to the Voigt-Hjerting function and the list of line parameters obtained by [Cristiani & D'Odorico \(2000\)](#). The lower panels show the logarithmic difference between our synthetic spectrum and one generated using the same list of line parameters and Monaghan's algorithm to calculate Voigt profiles. For reference, we include in the difference panels an horizontal line corresponding to a logarithmic difference of -4 dex.

available, for a typical range in column densities and Doppler parameters present in a QSO spectrum. The result is shown in Figure 3.8. The upper panels of each row show the observed spectrum and the spectrum synthesised using our method, whereas the lower panels show the logarithmic difference between both synthetic spectra. We choose to cut off the logarithmic difference at  $10^{-16}$ , since differences smaller than these are out of the range of the highest available numerical precision. It can be seen again, as in Figure 3.6, that the smallest discrepancies are given at the line cores, and the largest, of the order of  $10^{-4}$ , are given at the wings (cf. discussion of Figure 3.6, Section 3.5.3). As can be seen from the column density and Doppler parameter distribution in Figure 3.7, the largest discrepancies in this wavelength range are consistent with the maximum absolute differences shown in Figure 3.5. Note that in the spectral regions where no apparent absorption features are found, the logarithmic difference does *not* fall to  $-\infty$ , as one would naively expect. These features are present in Figure 3.6 as well. The reason for this 'valley-shaped' features is that, even though having small values away from the line center, the function  $H$  does *not* fall to zero, and thus in these regions the wings of two or more lines overlap. Strictly speaking, in these regions there is always some absorption left, *i.e.*  $\tau < 1$ , and different methods to compute  $H$  will account for this effect differently. Again, since our method neglects terms of order  $n \geq 2$ , the absorption in these regimes differs from its fiducial value. The lack of this terms in our approximation to  $H$  is also pointed out pictorially by the local maxima symmetrically placed around the deeps corresponding to the logarithmic difference at the line cores.

### 3.7 SUMMARY

The absorption lines imprinted in the spectra of background sources yield a wealth of information about the physical and chemical properties of the intervening absorbing material, as is the case of intervening neutral hydrogen (HI) systems embedded in the intergalactic medium (IGM) and associated with galaxies and larger structures. In order to extract the desired information from these absorption lines, their profiles have to be modeled in a proper way. In the case of absorption features found on QSO spectra, absorption line profiles are best modeled by Voigt profiles, which are mathematically given in terms of the Voigt-Hjerting function.

In this work, we presented a simple analytic approximation to the Voigt-Hjerting function with which Voigt profiles can be modeled for an arbitrary range in wavelength (or frequency), column densities up to  $10^{22} \text{ cm}^{-2}$ , and for damping parameters satisfying  $a \lesssim 10^{-4}$ . Starting with an exact expression for the Voigt-Hjerting function in terms of Harris' expansion that is valid for  $a < 1$ , we showed that the zeroth order term of this series, a Gaussian function, is suitable for modeling absorption line profiles emerging in a medium with low column density  $N_{\text{HI}} \lesssim 10^{15} \text{ cm}^{-2}$ . However, for higher column densities, terms of higher order have to be taken into account. A key point leading to this conclusion is the fact that it is not the damping parameter alone, but rather the factor  $a \cdot N_{\text{HI}}$  that determines to which extent terms of order higher than zeroth in Harris' expansion may or may not be neglected. We showed that the departure of the actual Voigt-Hjerting function from the first two terms in Harris' expansion is of the order of  $10^{-7}$  or less for an arbitrary wavelength range and  $a \lesssim 10^{-4}$ . Hence, we concluded that with an approximation to first order in  $a$  to the Voigt-Hjerting function Voigt profiles can be modeled with moderate to high accuracy.

On this basis, we obtained a simple analytic expression for the Voigt-Hjerting function and consequently for the absorption coefficient of intergalactic HI, in terms of an approximation for the second term  $H_1$  of Harris' expansion. The main advantages of the analytic expression we presented here are, first, that it is valid for an arbitrary wavelength range, in the sense that it has no singularities. In addition, it is simple and flexible to implement in a variety of programming languages to numerically compute Voigt profiles with moderate speed and moderate to high accuracy. As a matter of fact, our method to compute the Voigt-Hjerting function is faster with respect to other known standard methods, for instance, Humlíček's or Monaghan's algorithm. Furthermore, our approximation reaches an accuracy of  $10^{-4}$  or better in a wide wavelength range, and of the order of than  $10^{-2}$  only a negligible wavelength interval, for values of  $a$  and  $N_{\text{HI}}$  characteristic to intergalactic HI absorbers. Our method thus offers a great compromise between speed, accuracy, and flexibility in its implementation.

Even though we did not extend our discussion in this work to other transitions typically present in quasar absorption spectra and associated to HI absorbers, such as metal lines, our method to synthesise Voigt profiles can certainly be applied to most of these elements as well,

since their column densities are obviously the same, and their ranges in  $a$  strongly overlap with the range of  $a$  for intergalactic HI, for which our approximation to the Voigt-Hjerting function is valid. As a matter of fact, our approximation is valid to model absorption Voigt profiles found in any type of spectrum (stellar, solar, etc.), which arise in a medium whose damping parameter and column density satisfy  $a \lesssim 10^{-4}$  and  $N_{\text{HI}} \leq 10^{22} \text{ cm}^{-2}$ , respectively.

### 3.8 ACKNOWLEDGMENTS

I thank Uta Fritze-v.Alvensleben and the Göttingen Galaxy Evolution Group for encouraging comments. Special thanks to the referee M.T. Murphy for providing his computational routine to compute Voigt profiles, for pointing out some additional references, and for very helpful suggestions which helped to significantly improved the presentation of our results. This project was partially supported by the *Mexican Council for Science and Technology* (CONACYT), the Göttingen Graduate School of Physics (GGSP), and the *Georg-August-University of Göttingen*.

---

## The Intergalactic Medium

---

In this chapter, some fundamental cosmological concepts are introduced, which are needed to explain our current knowledge about the structure of the Universe, in particular of the intergalactic medium and its neutral hydrogen content. This in turn is necessary in order to understand the nature of the Ly $\alpha$  absorbers, which will be discussed in some detail.

### 4.1 INTRODUCTION

After the Big Bang, the baryonic matter component of the Universe was highly ionised due to the high temperature present at those epochs, leading matter and radiation to interact continuously with each other, and making the Universe opaque due to Thomson scattering<sup>1</sup> on the free electrons in the hot, dense plasma. Some  $3.8 \cdot 10^5$  years after the Big Bang—corresponding to a redshift  $z \approx 1000$ —, the Universe cooled below 3000 K (or 0.25 eV), allowing matter and radiation to decouple. From this point on, nucleons and electrons were able to recombine in order for the baryonic plasma to become neutral, and radiation was allowed to travel freely through the Universe, being redshifted due to expansion as it propagated, and becoming what can now be observed as the Cosmic Microwave Background (CMB).

After the decoupling of matter and radiation, the Universe entered a so-called "Dark Age", until the first structures and light sources (stars, galaxies, black holes)

formed due to gravitational collapse. The Universe was then slowly reheated and the neutral gas reionised probably by a first generation of stars and their corresponding supernova explosions (SNe) at the end of their lives. It is also plausible that a first generation of quasars (QSOs), *i.e.* accretion disks around massive black holes (BH) formed at this early stage in the evolution of the Universe, which generated a high-energy radiation field that reionised the baryonic matter in the intergalactic medium (IGM). With the increasing radiation field, the reionised regions around each source continuously grew, and the IGM became patchy, until the different reionised regions overlapped to form a single, highly ionised phase, completing the epoch of reionisation at a redshift  $z_{reion} \approx 10.5$  (Spergel et al. 2006).

Hence, the Universe, as is now known to us, is filled with radiation and matter. The visible component of the latter, *i.e.* the baryonic matter, shows large temporal and spatial variations in its density. In other words, in a comoving volume of space, the matter density becomes larger as we look back in time. Conversely, at a given time, matter density fluctuations are large, so large indeed that in the present Universe almost empty regions, called *voids*, alternate with regions of high density, such as galaxy clusters. According to the best of our present knowledge, the visible matter, which amounts to nearly 4% of the total matter content of the universe, traces the distribution of underlying, non-baryonic dark matter (DM), which accounts for a total of 22% of the matter content in the Universe. The remaining 74% is composed of an still unknown, mysterious component

<sup>1</sup> Named after its discoverer, the british scientist Sir Joseph John Thomson, OM, FRS (\*1856, †1940).

called dark energy (DE) (see *e.g.* Spergel et al. 2006).

## 4.2 THE BACKGROUND COSMOLOGY

The dynamics of the Universe as a whole – and the matter it contains – are governed on the large scale by the gravitational force, which is currently described by Einstein’s General Theory of Relativity (GR). The underlying idea of this theory is that the gravitational field is locally given by the geometrical structure of space-time, which is described by its metric. Conversely, the presence of any form of energy – or matter, through the relation  $E = m c^2$  – determines the metric of space-time. For a given metric, the equations of motion of the orbit  $x(\tau)$  of a free-falling test particle are given by

$$\frac{d^2 x^\lambda}{d\tau^2} = -\Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}. \quad (4.1)$$

for  $\lambda, \mu, \nu = 0, 1, 2, 3$ . The quantities  $\Gamma_{\mu\nu}^\lambda$  are the so-called Christoffel symbols which are defined by

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2} g^{\lambda\kappa} \left( \frac{\partial g_{\nu\kappa}}{\partial x^\mu} + \frac{\partial g_{\mu\kappa}}{\partial x^\nu} - \frac{\partial g_{\nu\mu}}{\partial x^\kappa} \right), \quad (4.2)$$

where  $g_{\mu\nu}$  is the metric tensor.

The equations which describe the relation between the gravitational field (metric) and its sources (energy) are Einstein’s field equations, which are given by

$$R_{\mu\nu} = -\frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right), \quad (4.3)$$

where  $R_{\mu\nu}$  is the Ricci tensor,  $T_{\mu\nu}$  is the energy-momentum tensor, and  $G$  is the Newtonian gravitational constant.<sup>2</sup> The metric tensor defines the geometry of space-time, while the energy-momentum tensor contains all the sources of the gravitational field, that in turn specify the metric of space-time, making the equations highly non-linear. Mathematically, this is reflected by the fact that the Ricci tensor is linear in second derivatives of the metric tensor, and quadratic in its first derivatives, so that equation (4.3) is a highly non-linear differential equation in  $g_{\mu\nu}$ .

There are many known exact solutions of – *i.e.* metrics that satisfy – Einstein’s field equations, one of which is of particular interest for cosmology, the so-called Friedmann-Lemaître-Robertson-Walker (FLRW)

metric.<sup>3</sup> This metric can be derived by demanding that the matter distribution, *i.e.* the energy-momentum tensor, in the Universe satisfy (only) isotropy and homogeneity when averaged over large enough volumes of space<sup>4</sup>, a requirement that is known as the Cosmological Principle<sup>5</sup> and is to some extent supported by observations. The line element  $ds$  for the FLRW metric is given by

$$ds^2 = c^2 dt^2 - a(t) \left( du^2 + f^2(u) d\Omega^2 \right), \quad (4.4)$$

where the dynamical variable  $a(t)$  is the scale factor of the Universe that quantifies its linear dimension,  $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$  is the square surface element on the two-dimensional unit sphere  $S^2$ , and  $f(u)$  is equal to  $\{\sin u, u, \sinh u\}$  for  $k$ , the time-independent curvature parameter, equal to  $\{+1, 0, -1\}$ , respectively. The values of  $k$  describe, in that order, a closed, a flat, or an open geometry.

Plugging the FLRW metric into Einstein’s field equations yields a set of differential equations in  $a$  known as Friedmann’s equations, which describe the dynamics of the Universe.<sup>6</sup> In this model, galaxies are seen as “particles” of a cosmic fluid with density  $\rho$  and pressure  $p$ , which according to equation (4.4) move along radial orbits (in comoving coordinates), onto which peculiar motions can be superposed. Friedmann’s equations, which relate various cosmological parameters, can be written as

$$H^2(t) \equiv \left( \frac{\dot{a}}{a} \right)^2 = \frac{1}{3} (8\pi G \rho + \Lambda) - k \left( \frac{c}{a} \right)^2 \quad (4.5)$$

$$3 \frac{\ddot{a}}{a} = \Lambda - 4\pi G \left( \rho + 3 \frac{p}{c^2} \right) \quad (4.6)$$

Here,  $\Lambda$  is the cosmological constant – which is thought to originate as a consequence of vacuum energy –, and  $H$  is the Hubble parameter, which measures the expansion rate of the Universe. A further quantity of interest is the dimensionless deceleration parameter  $q$ , which is defined by

$$q \equiv -\frac{1}{H^2} \left( \frac{dH}{dt} + H^2 \right) \quad (4.7)$$

<sup>3</sup> This metric was independently and almost simultaneously discovered by Alexander Friedmann, Georges Lemaître, Howard Percy Robertson, and Arthur Geoffrey Walker.

<sup>4</sup> It is of course vague what it’s meant by this, but it seems that a volume of some  $10^8$  pc across would do. This scale is what is known as the “End of Greatness”.

<sup>5</sup> One version of this principle states that: “In the Universe, all positions and directions are equivalent”.

<sup>6</sup> These equations were found in 1922 by the Russian cosmologist and mathematician Alexander A. Friedmann (\*1888, †1925). The FLRW was actually derived by Friedmann from these equations.

<sup>2</sup>  $G = 6.67300 \cdot 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^{-2}$

Friedmann's equations predict a dynamic Universe, whose fate depends on the total amount of matter (energy) of all types it contains. This quantity can be useful parametrized by introducing the density parameter  $\Omega$ , which is defined by

$$\Omega \equiv \frac{\rho}{\rho_c} = \frac{8\pi G}{3H^2}\rho. \quad (4.8)$$

Here,  $\rho_c \equiv 3H^2/(8\pi G)$  is the so-called critical density, *i.e.* the density for which the Universe would have a flat geometry. In other words,  $\Omega = 1$  would imply  $k = 0$ . Otherwise, the Universe would have an open or a closed geometry for  $\Omega < 1$  and  $k = -1$  or  $\Omega > 1$  and  $k = +1$ , respectively. The density term in Friedmann's equations includes in general the contribution from different components, which will evolve independently as the Universe expands. It is often assumed that each of these components obeys a very simple equation of state of the form (Carroll 2001)<sup>7</sup>

$$p_i = w_i c^2 \rho_i \quad (4.9)$$

where  $\rho_i$  is the density of the corresponding component, and  $w_i$  is a constant. In this case, it follows from Friedmann's equations that the density  $\rho_i$  evolves as

$$\rho_i \propto a^{-3(1+w_i)} \quad (4.10)$$

as shown in Appendix C.

The concept of the density parameter can thus be generalised to allow for all possible sources of the gravitational field (*i.e.* radiation, matter, cosmological constant, and curvature). Following Carroll (2001), the density parameter for each component may be defined in analogy to (4.8) as

$$\Omega_i \equiv \frac{8\pi G}{3H^2}\rho_i. \quad (4.11)$$

Taking all these into account, the first of Friedmann's equations can be rewritten as

$$\left(\frac{H(t)}{H_0}\right)^2 \equiv \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_\Lambda + \Omega_k a^{-2}\right]. \quad (4.12)$$

where  $\Omega_r$  is the radiation density,  $\Omega_\Lambda$  is the cosmological constant density,  $\Omega_k \equiv -kc^2/H_0^2$  is the contribution from the space curvature to the total density of

<sup>7</sup> We do not follow the convention that  $c = 1$ , and thus  $c$  appears in all equations where it should.

the Universe, and all correspond to present values.<sup>8</sup> It is a straightforward consequence of the above equation and the definition of  $H_0$  that

$$\Omega_{tot}(t=0) \equiv \Omega_r + \Omega_m + \Omega_\Lambda + \Omega_k = 1. \quad (4.13)$$

Using this last equation, one may alternatively define

$$\Omega_k \equiv 1 - \Omega_r - \Omega_m - \Omega_\Lambda. \quad (4.14)$$

A useful form of the deceleration parameter (4.7) can be obtained using equations (4.6) and (4.11). First note that equation (4.6) can be rewritten by including all forms of energy(-density) as

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G \sum_i (1+3w_i)\rho_i, \quad (4.15)$$

where  $\sum_i \rho_i \equiv \rho$ , and  $p$  has been replaced using equation (4.9). Furthermore, the deceleration parameter (4.7) can be rewritten as  $q = -\ddot{a}/(aH^2)$  (see Appendix D). Inserting the above equation into this last expression, and using equation (4.11) yields

$$q = \sum_i \frac{1+3w_i}{2} \Omega_i. \quad (4.16)$$

The different components of the total density  $\Omega_{tot}$  can be estimated from independent observations. For instance, measurements of the first peak of in the angular power spectrum of the spherical harmonic decomposition of the Cosmic Microwave Background (CMB) temperature field are consistent with a Universe with a total density  $\Omega_{tot} \approx 1$ , *i.e.* a flat Universe (see *e.g.* de Bernardis et al. 2000). This implies  $k \approx 0$ , and since the radiation density is negligible at the present epoch, it follows that  $\Omega_m + \Omega_\Lambda \approx 1$  (see equation 4.13). It follows from this and equation (4.16) that

$$q_0 = \frac{1}{2} \Omega_m - \Omega_\Lambda. \quad (4.17)$$

since  $\Omega_k = 0$  and  $\Omega_r \approx 0$ .

Observations of distant Type Ia Supernovae (SN Ia) imply that the expansion of the Universe is accelerating, *i.e.*  $q_0 < 0$ , which according to the above equation

<sup>8</sup> As a rule of convention, the subscript "0" (nought) in all cosmological quantities indicates their respective present value. For instance  $H_0$  is the value of the Hubble parameter at the present time. However, in order to avoid a confusing notation, we will leave out this subscript for the density parameters, and we will adopt the convention that they always refer to its present value, if not stated otherwise.

**Table 4.1:** State-of-the-art measurements of some cosmological parameters, obtained from the three-year WMAP Data Release (Spergel et al. 2006).

Parameter	Value	Description
$H_0$	$70.9^{+2.4}_{-3.2} \text{ km s}^{-1} \text{ Mpc}^{-1}$	Hubble parameter
$\Omega_b$	$0.0444^{+0.0042}_{-0.0035}$	Baryon density
$\Omega_m$	$0.266^{+0.025}_{-0.040}$	Total matter density (Baryons + Dark Matter)
$\Omega_\Lambda$	$0.732^{+0.040}_{-0.025}$	Dark Energy density

points to the existence of a density component with negative pressure (Perlmutter et al. 1999). These results, together with the fact that the matter density has been constrained by different measurements to  $\Omega_m \approx 0.26$  (Percival et al. 2007), imply that  $\Omega_\Lambda \approx 0.74$ . All these results are consistent with the most recent measurements of some cosmological parameters of the WMAP (Spergel et al. 2006), which are listed in table 4.1. The set of parameters  $\{H_0, \Omega_b, \Omega_m\}$ , together with the optical depth to reionisation  $\tau$ , the scalar fluctuation amplitude  $A_s$ , and the scalar spectral index  $n_s$  are called the concordance cosmology.

### 4.3 THE LARGE-SCALE STRUCTURE OF THE UNIVERSE

#### 4.3.1 The $\Lambda$ -Cold Dark Matter ( $\Lambda$ CDM) Model

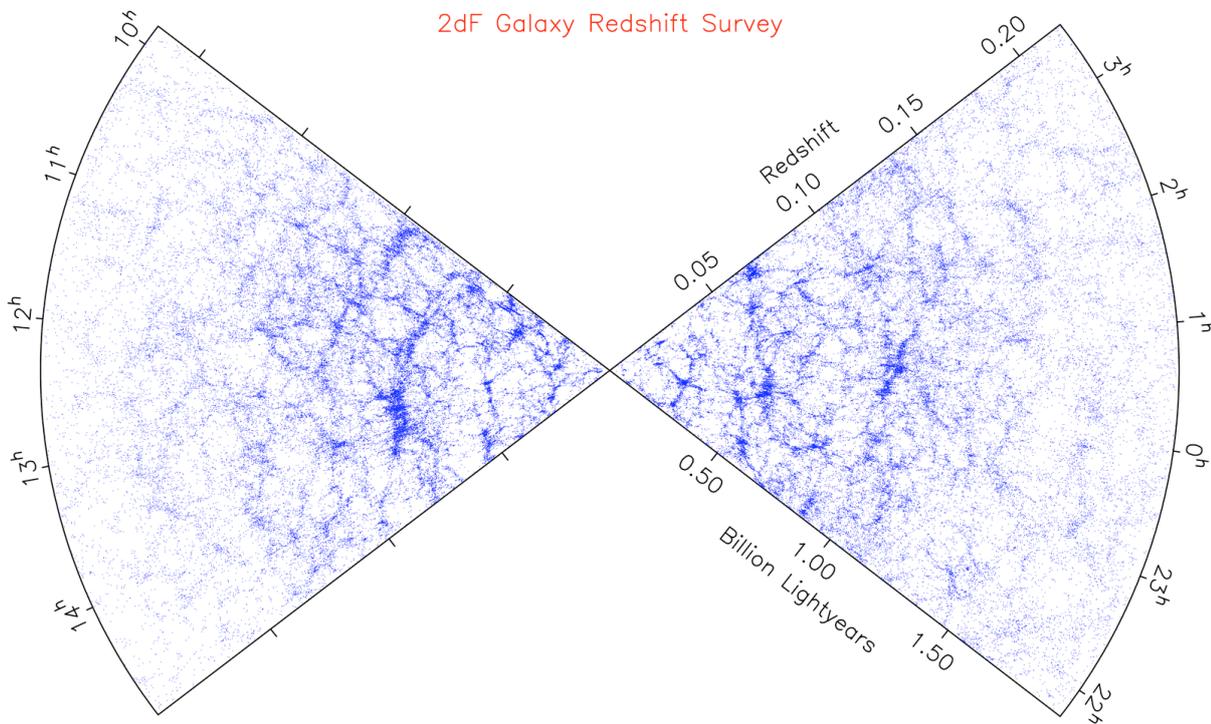
Our view of the structure and composition of the IGM has fundamentally changed in the last years motivated by refined observations and improved theoretical work. The assumption of a smooth—as opposite to lumpy—IGM, consistent of a uniformly distributed, highly ionised gas was discarded as the attempt to detect its most prominent observational signature—a continuous absorption trough on the spectra of background sources, that became known as Gunn-Peterson trough—failed. Through the observation of discrete, narrow absorption lines imprinted in the spectra of background sources over a large wavelength interval, together with the assumption that these are caused by the small fraction of neutral hydrogen present in the highly ionised intergalactic gas intersected along the line-of-sight, it soon became clear that the IGM is clumpy and not smooth or uniform. Different models were proposed to account for the distribution and the nature of observed absorption features.

The assumption of a two-phase intergalactic medium, in which a cold component containing basically neutral H I is embedded in a warmer so-called intercloud medium (ICM) as proposed by Sargent et al. (1980), Ostriker & Ikeuchi (1983), and the resulting model of pressure-confined Ly $\alpha$  clouds (Baron et al. 1989) was abandoned for the idea that the absorbing systems are dominated by their own gravity, as proposed a few years before by Melott (1980). This, in turn, got replaced by the idea that the intergalactic neutral hydrogen gas is contained in the potential wells of Dark Matter haloes (Umemura & Ikeuchi 1985). They suggested two different types of Dark Matter: On the one hand, they proposed the existence of Hot Dark Matter (HDM) (such as neutrinos) and on the other hand, the presence of Cold Dark Matter (CDM). The main argument against the existence of HDM is that this type of particles would tend to erase the structures on small scales at very early epochs in the Universe, at variance with what is observed today.

Later on, the lumpy nature of the IGM and the Universe at large scales became apparent as a result of large redshift surveys as the CfA (Davis et al. 1982), Las Campanas (Shectman et al. 1996), 2dFGRS<sup>9</sup>, SDSS<sup>10</sup>, etc. that revealed that the redshift and spatial distribution of galaxies is not uniform. Instead, as times goes on, galaxies tend to clump in clusters, and in and around the filamentary structures connecting these clusters, leaving large empty regions of space, the so called voids (see Fig. 4.1). Such observations show that the Universe is quite inhomogeneous at small scales, and nearly homogeneous and isotropic at large scales. This result was corroborated by the measurement of the Cosmic Microwave Background (CMB) first by the *Cos-*

<sup>9</sup> <http://www2.aao.gov.au/2dFGRS>

<sup>10</sup> <http://www.sdss.org>



**Fig. 4.1:** Galaxy redshift distribution as measured by the 2dF facility at the Anglo-Australian Observatory (AAO). Taken from [www2.aao.gov.au/2dFGRS](http://www2.aao.gov.au/2dFGRS)

mic Background Explorer (COBE)<sup>11</sup> and later by the *Wilkinson Microwave Anisotropy Probe* (WMAP) satellites<sup>12</sup>, which mapped the temperature distribution of the CMB on the sky, revealing a smooth background with anisotropies of the order of  $\Delta T/T = 10^{-5}$ , which trace the primordial density fluctuations that gave origin to the structure observed in the Universe today.

Obviously, the development of a consistent model explaining all these observational results and including the constraints on the cosmology implied by the measurement of different cosmological parameters was imperative. First proposed by [Ostriker & Steinhardt \(1995\)](#), the so-called Concordance or  $\Lambda$ -Cold Dark Matter model<sup>13</sup> has become the most fruitful attempt to simultaneously explain in a consistent manner the CMB radiation, the expansion of the Universe, and the large-

scale structure of the Universe. In this model, a flat Universe and the existence of a so-called Dark Energy (DE) component, parametrized by the cosmological constant  $\Lambda$ , are assumed. Furthermore, a weakly-(or non-)self-interacting, non-radiating, collisionless matter component, the so-called cold dark-matter (CDM), is thought to interact only gravitationally with ordinary (baryonic) matter.<sup>14</sup> According to this scenario, the structures in the Universe are formed hierarchically (bottom-up, as opposed to "fragmentation" or top-down) and evolve in time due to gravitational interaction. It is assumed that before the decoupling era, *i.e.* before the time when the temperature of the Universe had dropped enough as to allow for matter to recombine, and the mean free path of radiation increased to almost the size of the Universe, the CDM had already decoupled from the primordial plasma and began to collapse under the action of gravity around overdense regions generated by quantum fluctuations and enlarged during the period of inflation ([Guth](#)

<sup>11</sup> <http://lambda.gsfc.nasa.gov/product/cobe/>

<sup>12</sup> <http://map.gsfc.nasa.gov/>

<sup>13</sup> "ΛCDM says nothing about the fundamental physical origin of dark matter, dark energy and the nearly scale-invariant spectrum of primordial curvature perturbations: in that sense, it is merely a useful parametrization of ignorance."

<sup>14</sup> The difference between "cold" and "hot" dark matter refers to the particle's velocity at the epoch of decoupling: the former were non-relativistic, in contrast to the later.

1981). After matter and radiation decoupled, baryonic matter, *i.e.* the photoionised gas, began to condensate in the potential wells of the already collapsed DM, giving rise to the different structures such as galaxies and galaxies clusters, and to the gaseous structures seen in absorption against bright background sources, *e.g.* the Ly $\alpha$  absorption features (cf. Sec. 4.4). In this scenario, a fixed mass ratio for the baryons in a DM halo of a given mass is conventionally assumed. At the same time, the formation of structures at all scales under the influence of gravity lead the IGM to become clumpy and to acquire peculiar motions.

Taking the  $\Lambda$ CDM paradigm as the underlying formation scenario and with the help of state-of-the-art numerical simulations, such as the Millenium Simulation<sup>15</sup> performed using the GADGET code (Springel et al. 2005), it has indeed become possible to reproduce the absorption features and their statistical properties (see *e.g.* Demiański et al. 2000), under the assumption that they are produced by the resonant scattering of photons off the line-of-sight by neutral intergalactic hydrogen present in the (forming) structures intercepted by the line-of-sight. The results of these simulations imply that the IGM should contain 50%-90% of the baryons in the Universe out to  $z \approx 2$ . This in turn implies that the galaxies observed at high redshift such as the Distant Red Galaxies (DRGd), the Lyman Break Galaxies (LBGs), and the Damped Ly $\alpha$  Systems (DLAs, see below) trace the highest density peaks in the distribution of matter at those redshifts. Such predictions can be tested against observations by estimating the total amount of bright baryonic matter, on the one hand, and by taking advantage of the fact that the neutral, non-luminous component of the baryonic matter present in the IGM can be still detected in absorption in the form of the so-called Ly $\alpha$  absorbers on the spectra bright enough background sources. Conversely, a census of the amount and distribution of this non-luminous component is necessary in order to assess its impact on the spectra of other bright sources than quasars. In particular, it is of great interest for the development of models that account for the spectrophotometric evolution of galaxies. The effect of the absorption in the intergalactic medium on the spectra of high-redshift galaxies will be analysed in Chapter 5 and beyond. However, it is necessary first to introduce and present in some detail the properties of the intergalactic neutral hydrogen in form of the Ly $\alpha$  absorbers.

<sup>15</sup> <http://www.mpa-garching.mpg.de/galform/millennium/>

#### 4.4 THE Ly $\alpha$ ABSORBERS

It is a prediction of Big-Bang Nucleosynthesis, that hydrogen (H), helium ( $^4\text{He}$ ), deuterium (D) and lithium ( $^7\text{Li}$ ) were formed shortly after the Big Bang in a precise ratio, namely  $\text{H} : ^4\text{He} : \text{D} : ^7\text{Li} \approx 1 : 1/3 : 10^{-5} : 10^{-10}$ .<sup>16</sup> Out of this primordial gas, large, complex structures such as stars, galaxies, galaxy clusters, etc., began to form when this gas settled down and cooled down in the potential wells of the underlying dark matter, as explained above. Between redshift  $z = 15$  and  $z = 7$ , heavier and more complex elements than lithium were synthesised in the first generation of stars within proto-galaxies – so-called Population (Pop) III stars – and restored to the interstellar medium (ISM) by various mechanisms (supernovae explosions, and later stellar winds) where they would eventually be recycled for the formation of new generations of stars. Moreover, these heavier elements found its way even into the IGM, probably through galactic winds observed in many LBG, tidal stripping in galaxy interaction, and eventually also through pre-galactic Pop III stars. Thus, throughout the evolution of the Universe, the intergalactic medium has been a source for the gas that forms stars, galaxies, etc., and has served as a repository for enriched material and energy expelled from galaxies (Madau 2000).

Hydrogen is the most abundant element in the Universe, in the interstellar medium (ISM) as well as in the IGM. Due to the strong ultraviolet (UV) background, *i.e.* the UV flux of quasars (QSOs), and actively star-forming, young galaxies at high-redshift, the hydrogen present in the IGM is highly ionised. It is found, however, to be in ionisation equilibrium, with a small, but finite amount of neutral hydrogen (H $\text{I}$ ) in a ratio  $n_{\text{HI}}/n_{\text{HII}} \leq 10^{-4}$  (Rauch 1998). Even though extremely diluted, this amount of neutral hydrogen is enough to severely absorb the light of background sources along cosmological distances, via photoionisation and resonant scattering at the Lyman transitions (cf. Chapter 2).

The first attempts to assess the amount and the distribution of neutral hydrogen in the IGM go back to 1965, when Gunn & Peterson (1965) first noted that in an ex-

<sup>16</sup> The most recent measurement of the primordial helium abundance, usually denoted as  $Y_p$ , has recently been done by Peimbert et al. (2007). The quoted value is  $Y_p = 0.2474 \pm 0.0028$ . Previously, Coc et al. (2004) measured  $Y_p = 0.2479 \pm 0.0004$ , and they also estimate  $\text{D}/\text{H} = (2.60^{+0.19}_{-0.17}) \cdot 10^{-5}$ ,  $^3\text{He}/\text{H} = (1.04 \pm 0.04) \cdot 10^{-5}$ , and  $^7\text{Li}/\text{H} = (4.15^{+0.49}_{-0.45}) \cdot 10^{-10}$ . See also Izotov et al. (2007).

panding universe with a continuous gas distribution, the absorption due to H $\alpha$  should imprint in the spectra of background sources an absorption trough blueward of the Ly $\alpha$  emission line of the source, due to the continuously redshifted photons emitted by the source and the presence of intergalactic H $\alpha$  at each point along the line-of-sight. They suggested that the amount of H $\alpha$  present in the IGM could in principle be estimated by measuring the relative strength of this so-called Gunn-Peterson (GP) trough, a method that ever since has become known as the GP-test. Later that same year, Bahcall & Salpeter (1965) proposed that a clumpy gas distribution would yield a discrete set of absorption lines on the spectrum of a background source, each line corresponding to the Ly $\alpha$  line of a clump of absorbing gas. In the following years, observations of the first discrete absorption features on the spectra of high-redshift, background quasars were reported (Greenstein & Schmidt 1967, Burbidge et al. 1968).

These features are composed of hundreds of absorption lines, which under the assumption of being due to the Ly $\alpha$  transition of neutral hydrogen, as first proposed by Lynds (1971), made it possible to determine the redshift of the absorbing entities and conclude that these were not related to the quasar. Thus, between two competing scenarios on the distribution and content of neutral hydrogen in the IGM, a new view of an Universe slowly emerged, according to which the intergalactic medium contains a low-density, highly ionised gas, with a neutral component which shows large fluctuations in column density. These density fluctuations become visible against background sources as absorption lines imprinted on the observed spectra. The entities giving rise to these absorption features became to be known ever since as Ly $\alpha$  absorbers.

Hence, when observed spectroscopically, bright background sources such as QSOs, and more recently Gamma-ray bursts (GRBs) show a large number of discrete absorption features (some  $10^3$ ) in their spectra. These absorption lines are mainly due to the resonant scattering of photons off the line-of-sight due to intergalactic neutral hydrogen, with a wavelength that corresponds to the Ly $\alpha$  transition at the epoch of absorption. In addition to Ly $\alpha$ , higher-order Lyman absorption lines and absorption lines due to a variety of heavier elements in different ionisation stages, such as C, O, Mg, Si, Fe and Al are found as well. Absorption lines due to Lyman resonant scattering are observed in the rest-frame UV, more specifically in the range (91.2, 121.5) nm, and deep absorption edges due to photoionisation of H $\alpha$  are

observable at wavelength  $\lambda \leq 91.2$  nm in the rest-frame of the source (cf. chapter 2).

In the realm of the cosmological paradigm of structure formation, these Ly $\alpha$  absorbers are thought to originate in the filamentary structures of the cosmic web as the intergalactic gas cools down and falls into the potential wells of dark matter haloes, in large gas haloes around galaxies, and in the outskirts of (proto-)galactic disks. Their evolution is driven mainly by the Hubble expansion, the ionising radiation field from background quasar and star-forming galaxies, but also by the growth of structures, by morphological evolution of galaxies, and by galaxy mergers. Ly $\alpha$  absorbers span a range in column density  $10^{12} \text{ cm}^{-2} \lesssim N_{\text{HI}} \lesssim 10^{22} \text{ cm}^{-2}$ , and have temperatures of the order of  $10^4$  K. On the basis of their column densities and metal content, and most importantly on the basis of their putative origin, three types of absorbers are distinguished:

#### 4.4.1 The Ly $\alpha$ Forest

The narrow absorption lines imprinted in large numbers on the spectra of QSOs are caused by low column density absorbers, are called Ly $\alpha$  forest clouds (Weymann et al. 1981), and correspond to systems with column densities in the range  $10^{12} \text{ cm}^{-2} < N_{\text{HI}} < 10^{17.21} \text{ cm}^{-2}$ . Their measured profiles and strengths (equivalent widths) are indicative of temperatures around  $2 \cdot 10^4$  K. Besides, they show some chemical enrichment, witnessed by the presence of weak, but measurable C $\text{IV}$  lines at redshifts identical to those of the corresponding Ly $\alpha$  line. Their abundances are estimated to be around 0.1% to 1%  $Z_{\odot}$ . This kind of absorption features are thought to originate in the gas flowing out from the centre of voids towards their shells, and along the filamentary structure towards the intersections where the denser structures form. At all redshifts, some of them may also arise in galactic winds.

#### 4.4.2 Lyman Limit Systems

The absorbing systems with column densities  $10^{17.21} \leq N_{\text{HI}} < 10^{20.3} \text{ cm}^{-2}$  have optical depths  $\tau(\lambda_L) \gtrsim 1$  (see eq. 2.3) for photons with wavelengths  $\lambda \leq \lambda_L$  (rest-frame), *i.e.* shortward of the Lyman edge, thus producing a prominent discontinuity in the observed spectrum. For this reason, these absorbers are commonly referred to as Lyman limit systems (LLSs). These systems have abundances around a few percent  $Z_{\odot}$ , and are associated on the basis of their absorption redshifts with

strong MgII absorption systems with equivalent widths  $W(\text{MgII}) > 0.3 \text{ \AA}$ .

Apparently, there seems to be a one-to-one correspondence between strong MgII absorbers and luminous galaxies, even though there is still some controversy with respect to this so-called *absorber-galaxy connection* (see e.g. Williams et al. 2005, Part 1). Nevertheless, it is plausible that they originate in the gaseous haloes of bright galaxies, or at the disk/halo interface (Galactic corona) of star-forming galaxies.

There is a class of weak MgII systems known as sub-LLS since no corresponding Lyman break is observed. These systems typically have abundances around 10%  $Z_{\odot}$ , a few of them reach a value comparable to  $Z_{\odot}$ . Some of these systems show FeII absorption, while others have CIV associated with them.

#### 4.4.3 Damped Ly $\alpha$ Absorbers (DLA)

At column densities  $N_{\text{HI}} \gtrsim 10^{20.3} \text{ cm}^{-2}$ , intergalactic neutral hydrogen experiences a kind of "self-shielding" against the ionising UV background, and remains almost completely neutral. The resonant scattering for photons with  $\lambda = 121.5 \text{ nm}$  due to neutral hydrogen with such column densities is dominated by radiation damping (cf. Sec. 2.2.2). Hence, when the line of sight to a bright background source happens to intercept an absorber with a column densities of  $N_{\text{HI}} \gtrsim 10^{20.3} \text{ cm}^{-2}$ , an absorption line with pronounced damping wings becomes observable. This type of absorbers, known as damped Ly $\alpha$  absorbers (DLAs), have relatively high metallicities (0.01 - 1)  $Z_{\odot}$ , and contain a variety of chemical elements, such as Zn, Cr, Fe, Mg, N, and many others.

These column densities are typical for galactic disks, and indeed, some of these systems show clear evidence for rotation, with  $v_{\text{rot}} \approx 120 - 240 \text{ km s}^{-1}$ . Already at  $z \approx 2 - 4$  these systems are fairly massive ( $10^{10} - 10^{11} M_{\odot}$ ), but very difficult to identify by optical observations. The few systems that have low-luminosity optical detections show star-formation activity. Their estimated gas masses roughly equal the combined stellar and gaseous masses of local spiral disks. Furthermore, It is also well established that DLAs may also redden and obscure QSOs (Fall & Pei 1989) due their large content of dust. Hence, DLAs are assumed to be proto-galactic disks, in which star-formation has just set on.

Even though they are scarce, DLAs account for most of the neutral hydrogen in the Universe at all epochs.

There is a subclass of objects with column densities in the range  $10^{19} \text{ cm}^{-2} < N_{\text{HI}} \lesssim 10^{22} \text{ cm}^{-2}$  which for a long time were referred to as sub-DLAs, and were recently proposed to be referred to as super LLS, since their properties are rather different from those of the Ly $\alpha$  damped systems (Wolfe et al. 2005, Rao 2005).

An example of a QSO spectrum displaying absorption due to all three different types of absorbers is shown in Fig. 4.2.

#### 4.5 PROPERTIES OF THE Ly $\alpha$ ABSORBERS

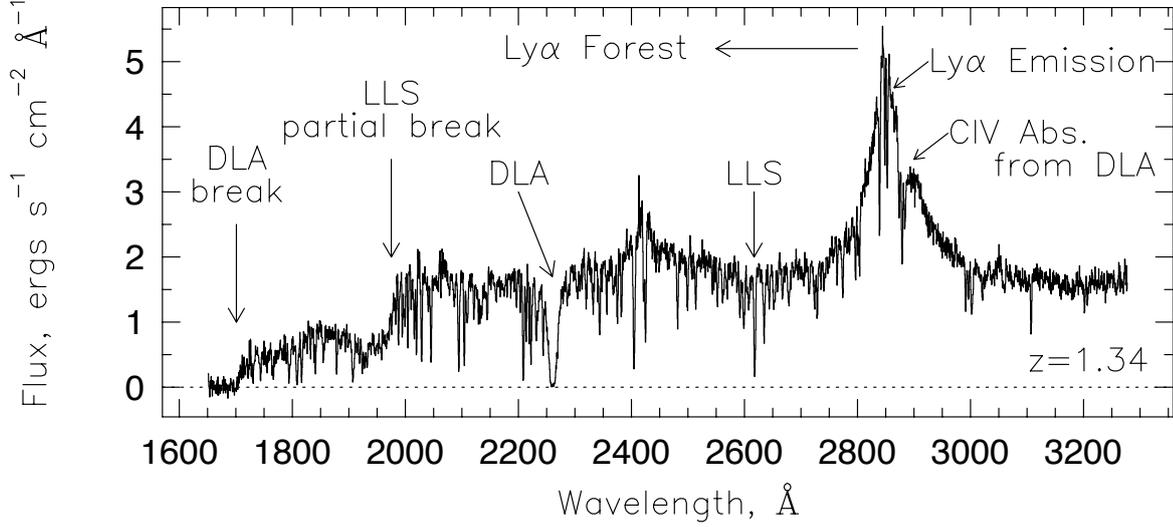
As briefly mentioned before, the intergalactic neutral hydrogen present in all types of Ly $\alpha$  absorbers, together with the baryons trapped in galaxies, trace the distribution of the underlying Dark Matter. Furthermore, since the observation and analysis of the Ly $\alpha$  absorption features are very useful to probe the evolution of cosmic gas back to the epoch of reionisation, it is important to know the baryonic content and physical conditions of such systems. Also, both Ly $\alpha$  forest clouds, LLSs and DLAs are known to significantly obscure background sources (Bechtold et al. 1987, Miralda-Escude & Ostriker 1990, Møller & Jakobsen 1990, Madau 1991, 1992, Meiksin & Madau 1993, Zuo 1993). It is therefore of interest to know the properties of these absorbers, *i.e.* the physical parameters that characterise them, such as their number densities, their (column) densities, and their temperatures in their evolution in cosmic time or redshift. As expected, the determination of all these properties is far from trivial, and results highly depend on the quality of the observations from which they are estimated. In what follows, a brief review of the formalism used to characterised the Ly $\alpha$  absorbers and the observational results that are derived from them are presented.

##### 4.5.1 The proper path length

The description of the absorption due to resonant scattering by HI in the intergalactic medium along random lines-of-sight is greatly simplified if one considers the evolution of the different properties of the absorbers in terms of the physical distance  $dl = c dt$ . We rederive here an expression for the physical length as a function of redshift which will prove very useful in the next sections.

From the definition of the cosmological redshift

$$1 + z_{\text{cosm}} \equiv \frac{a_0}{a(z)}, \quad (4.18)$$



**Fig. 4.2:** Spectrum of the quasar PKS0454+039 at  $z_{em} = 1.34$ , which shows all different types of Ly $\alpha$  absorbers blueward of the Ly $\alpha$  emission line of the QSO at  $\lambda = 2850 \text{ \AA}$ : a large number of Ly $\alpha$  forest clouds (weak absorption lines), a Lyman limit system at  $\lambda = 2614 \text{ \AA}$ , corresponding to  $z = 1.15$ , and a damped Ly $\alpha$  systems at  $\lambda = 2261 \text{ \AA}$  ( $z = 0.86$ ). Note the pronounced wings of the absorption line of the latter, and the sharp absorption edges due to this system and the LLS at  $\lambda = 1696 \text{ \AA}$  and  $\lambda = 1960 \text{ \AA}$ , respectively. This spectrum was taken with the Faint Object Spectrograph (FOS) of the Hubble Space Telescope (HST). Taken from [Charlton et al. \(2000\)](#).

it follows  $|da/a| = dz/(1+z)$ .<sup>17</sup> From this, the definition of the Hubble parameter (see equation 4.5) and the chain rule of differentiation one gets<sup>18</sup>

$$H(z) = \frac{1}{(1+z)} \left| \frac{dz(t)}{dt} \right|. \quad (4.19)$$

From both the latter equation and equation (4.12) it follows that

$$\left| \frac{dl}{dz} \right| \equiv c \left| \frac{dt}{dz} \right| = \frac{c}{H_0} \frac{1}{(1+z)} \left[ \Omega_r (1+z)^4 + \Omega_m (1+z)^3 + \Omega_\Lambda + \Omega_k (1+z)^2 \right]^{-\frac{1}{2}}. \quad (4.20)$$

In a matter-dominated universe where the total density is equal to unity, the above expression becomes (see *e.g.*

<sup>17</sup> For simplicity, let  $a_0 \equiv 1$ , which can always be achieved by a suitable choice of units.

<sup>18</sup> If we assume that  $t = t(z)$ , then strictly speaking  $H(t) \neq H(z)$  but rather  $H'(z) \equiv H[t(z)]$  where  $H$  and  $H'$  not necessarily have the same functional form. However, since this change of variable is done through all the following equations, we simply let  $H \rightarrow H'$  and avoid further notation problems.

[Barkana & Loeb 2007](#))

$$\left| \frac{dl}{dz} \right| = \frac{c}{H_0} \frac{1}{(1+z)} \left[ \Omega_m (1+z)^3 + \Omega_\Lambda + (1 - \Omega_m - \Omega_\Lambda) (1+z)^2 \right]^{-\frac{1}{2}}. \quad (4.21)$$

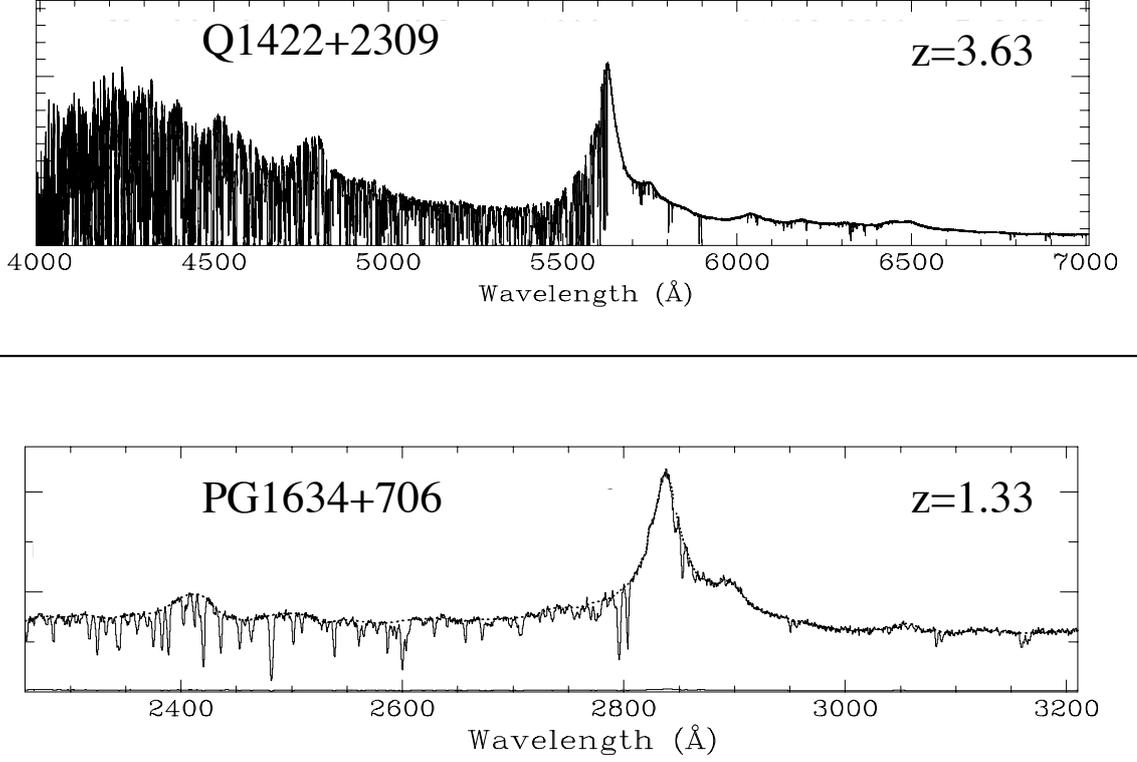
where  $\Omega_k$  has been replaced using equation (4.14).

For the special case of a closed Universe and vanishing cosmological constant  $\Omega_\Lambda = 0$ , equation (4.20) becomes ([Gunn & Peterson 1965](#))

$$\left| \frac{dl}{dz} \right| = \frac{c}{H_0} \cdot (1+z)^{-2} (1+2q_0 z)^{-\frac{1}{2}}. \quad (4.22)$$

as shown in appendix D. However, this expression cannot be valid in the light of the current cosmological paradigm, since it assumes a zero cosmological constant and a non-flat universe, *i.e.* exactly the opposite of what is observed. For a flat Universe ( $k = 0$ ) and  $\Omega_\Lambda \neq 0$ , we get for the variation of the physical path with redshift

$$\left| \frac{dl}{dz} \right| = \frac{c}{H_0} \cdot (1+z)^{-1} \left[ \Omega_m (1+z)^3 + \Omega_\Lambda \right]^{-\frac{1}{2}}. \quad (4.23)$$



**Fig. 4.3:** Comparison of the Ly $\alpha$  absorption on QSO spectra at two different redshifts. The upper panel shows the spectrum of the QSO Q1422+2309 at  $z = 3.63$ , taken with the High Resolution Échelle Spectrograph (HIRES,  $R = 25,000 - 85,000$ )<sup>a</sup> at the Keck 10m-telescope, while the lower panel shows the spectrum of the QSO PG1634+706 at  $z = 1.33$  taken with FOS ( $R = 250 - 1300$ )<sup>b</sup> at HST. Taken from Charlton et al. (2000). <sup>a</sup> <http://www2.keck.hawaii.edu/inst/hires>; <sup>b</sup> <http://quest.nasa.gov/hst/about/overview.html>

#### 4.5.2 The Number Density Evolution along the line-of-sight

The probability  $p(N_{\text{HI}}, z)$  of intercepting a single absorption system at a redshift  $z$  and with a column density in the interval  $(N_{\text{HI}}, N_{\text{HI}} + dN_{\text{HI}})$  in the physical length interval  $dl \equiv c dt$  along a random line-of-sight can be expressed as

$$p(N_{\text{HI}}, z) \equiv p_0 \cdot (n \cdot \sigma_A) dN_{\text{HI}} dl, \quad (4.24)$$

where

$$n(N_{\text{HI}}, z) = (1 + z)^3 \cdot n_0, \quad (4.25)$$

and  $n_0 = n_0(N_{\text{HI}}, z)$  is the comoving number density of absorbers with column densities in  $(N_{\text{HI}}, N_{\text{HI}} + dN_{\text{HI}})$

at  $z$ , and  $\sigma_A = \sigma_A(N_{\text{HI}}, z)$  is the geometric (effective) absorption cross-section at  $(N_{\text{HI}}, z)$ .

The normalisation constant  $p_0$  is such that

$$\int_{\Delta l} \int_{I_{N_{\text{HI}}}} p(N_{\text{HI}}, z) dN_{\text{HI}} dl \stackrel{!}{=} 1, \quad (4.26)$$

and the integral is computed over an appropriate path length  $\Delta l$  and column density interval  $I_{N_{\text{HI}}}$ . Nevertheless, the normalisation constant in equation (4.24) can be chosen in such a way that it gives the expected number of absorption systems on a given path along a random line of sight and with column densities in a given interval. This quantity is a true observable, and as such

of great value to constrain models for the evolution of the absorbers. Formally then, the number of absorbing systems with column density and redshift in the intervals  $(N_{\text{HI}}, N_{\text{HI}} + dN_{\text{HI}})$  in the path length  $dl$  along a random line-of-sight is given by (e.g. Peterson 1978, Wolfe et al. 2005)

$$\begin{aligned} dN(N_{\text{HI}}, z) &\equiv \frac{\partial^2 N}{\partial z \partial N_{\text{HI}}} dN_{\text{HI}} dz \\ &= n_0 \cdot \sigma_A \cdot (1+z)^3 \left| \frac{dl}{dz} \right| dN_{\text{HI}} dz, \end{aligned} \quad (4.27)$$

Here, the absolute value of  $dl/dz$  is needed, since  $l$  is a decreasing function of redshift, due to the expansion of the Universe. As shown in previous sections, a general expression for  $|dl/dz|$ , and hence for equation (4.27), can be obtained for the particular case of an isotropic and homogeneous Universe.

It is assumed that along a random line-of-sight, the probability of finding  $n$  absorbing systems is given by a Poisson distribution

$$P(n) = e^{-\langle N_{\text{abs}} \rangle} \frac{\langle N_{\text{abs}} \rangle^n}{n!}, \quad (4.28)$$

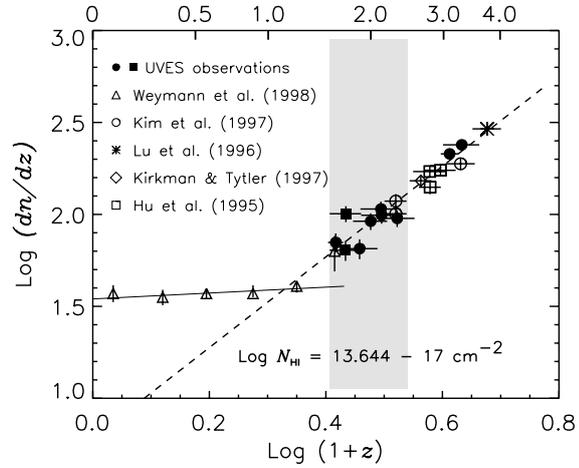
where the mean number  $\langle N_{\text{abs}} \rangle$  of systems for each line-of-sight can be computed from equation (4.27)

$$\langle N_{\text{abs}} \rangle = \iint \frac{\partial^2 N}{\partial z \partial N_{\text{HI}}} dN_{\text{HI}} dz,$$

In practice, a standard way to characterise the nature of the Ly $\alpha$  absorbers has been to account for their evolution as a function of redshift, which is usually parametrized by their line density evolution  $dN/dz$ . This quantity can be formally defined as the zeroth moment of the distribution (4.27) with respect to  $N_{\text{HI}}$ :

$$\frac{dN}{dz} \equiv \int \frac{\partial^2 N}{\partial z \partial N_{\text{HI}}} dN_{\text{HI}} \quad (4.29)$$

However, the quantity  $dN$  is not known *a priori* and thus the number density cannot be computed analytically. Rather, it has to be determined from observations. The usual method to estimate the number density evolution consists in counting the number of lines observed in the spectrum of a QSO at a certain redshift  $z_{\text{em}}$  for a given transition (in general, the Ly $\alpha$  transition, as it is the strongest) above a certain column density threshold (e.g. Peterson 1978, Sargent et al. 1980), and to average over as many lines-of-sight as possible. This technique, called line-counting, implies the highly plausible



**Fig. 4.4:** The number density evolution of the Ly $\alpha$  forest. The solid and dashed lines correspond to the maximum-likelihood fit to the data in the interval  $0 < z < 1.8$  and  $z > 1.8$ , respectively. Taken from Kim et al. (2002a, and references therein).

assumption that *each* absorber gives rise to a *single* line for a given transition.

For absorbers with no intrinsic evolution, *i.e.* whose evolution is only driven by the Hubble expansion, and for a flat Universe and  $\Omega_{\Lambda} \neq 0$  the number density evolution can be written as

$$\frac{dN}{dz} = \frac{c}{H_0} n_0 \cdot \sigma_A \frac{(1+z)^2}{[\Omega_m (1+z)^3 + \Omega_{\Lambda}]^{\frac{1}{2}}}, \quad (4.30)$$

This equations can simply be obtained by combining equations (4.23) and (4.27).

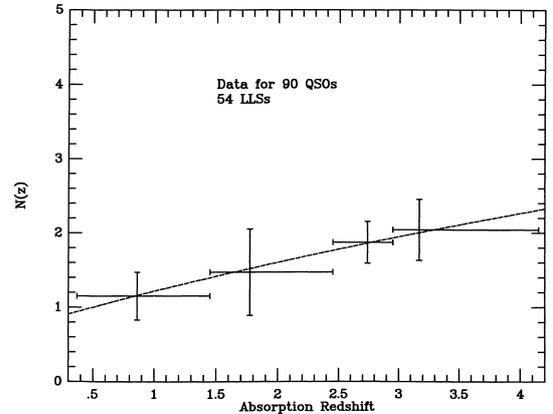
It is now widely accepted that the evolution of the absorbers is driven by the expansion of the Universe, as well as by other factors such as the UV background (see e.g. Schaye et al. 2000, Bianchi et al. 2001), at least at high redshift. Hence, the observed evolution of the number density of absorbers is expected to deviate from the above expressions. Nevertheless, it is usual to assume, in analogy to (4.30), that the number density obeys a power-law of the form

$$\frac{dN}{dz} = A \cdot (1+z)^{\gamma}, \quad (4.31)$$

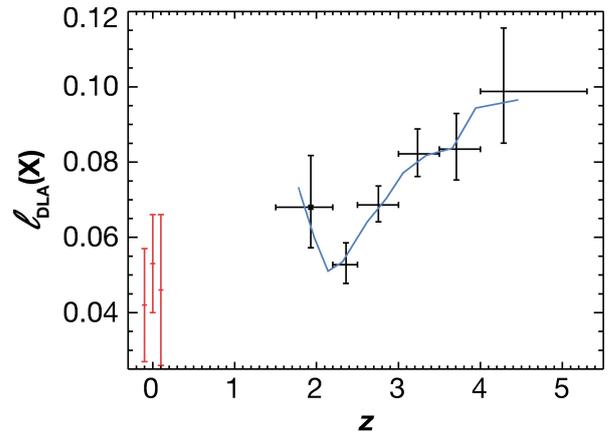
where  $A$  is a normalisation factor and the evolution parameter  $\gamma = \gamma(H_0, \Omega_m, \Omega_{\Lambda}, \dots)$  depends on the adopted cosmology.

In the literature, there are many determinations of the evolution parameter and normalisation constants performed in different redshift ranges, with different column density thresholds, which are rather confusing, and sometimes contradictory. An extensive compilation of the range of values found in the literature is given in Appendix E. Typical values of the normalisation factor  $N_0$  range from 0.1 to 180, while evolution parameters  $\gamma$  are found in the range (0.2, 5.7), depending on the redshift range chosen and the column density interval, *i.e.* type of absorber (cf. Appendix E). It has been argued that a single power-law does not well describe the evolution of the absorbers at all redshifts and column densities. Rather, power-laws with different parameters are needed for different column density and redshift ranges, which point the fact that absorbers with different column densities, *i.e.* of different types ( $\text{Ly}\alpha$  forest clouds, LLSs, DLAs) evolve differently (see Fig. 4.4). This in turn supports the idea that they also have different origins. For the  $\text{Ly}\alpha$  absorbers, it appears from observations that around  $z \approx 1.8$  there is a break in their evolution, in such a way that they show little or no intrinsic evolution in the range  $0 < z \lesssim 1.8$ , and strong evolution at higher redshifts. An illustrative example of the evolution of the low column density intergalactic H I from high to low redshift seen in QSO spectra is shown in Fig. 4.3, and the measurements of the number density evolution in the redshift range  $0 < z < 5.3$  is shown in Fig. 4.4. The LLSs, in contrast, show a mild evolution at  $0.5 < z < 4$ , as can be seen in Fig. 4.5. The optically thick DLAs evolve rapidly at  $z > 2$ , and appear to show no evolution at low redshifts (see Fig. 4.6).

The measurement of the normalisation and evolution parameters is complicated because of various reasons. First, the so-called blending, *i.e.* the overlap of two or more individual absorption lines, which may be misidentified as a single, broader line, can affect the line counts significantly for obvious reasons. Also, line counting is limited by the instrumental resolution to column densities  $N_{\text{HI}} \geq 3 \cdot 10^{12} \text{ cm}^{-2}$  (Hu et al. 1995). A further phenomenon which affects the line counting and hence the determination of the line evolution is the inverse effect, also called proximity effect, which describes the reduction in the line density towards the redshift of the quasar (Murdoch et al. 1986). The most plausible explanation of this effect is that the neutral H I density is reduced in the physical vicinity of the observed QSO due to its ionising radiation field (Bajtlik et al. 1988, Cristiani et al. 1995). There has also been reported on a "galaxy" proximity effect (Weinberg et al.



**Fig. 4.5:** The number density evolution of the  $\text{MgII}$  systems associated with Lyman limit systems. The solid line represents the maximum-likelihood fit for an assumed evolution of the form of equation (4.31), with  $A = 0.76$ , and  $\gamma = 0.68 \pm 0.54$ . Taken from Sargent et al. (1989, and references therein).



**Fig. 4.6:** The number density evolution of the damped  $\text{Ly}\alpha$  absorbers. The three data points at  $z = 0$  are local measurements from 21cm observations. Taken from Wolfe et al. (2005, and references therein).

2003), which corresponds to the observational fact that the incidence of  $\text{Ly}\alpha$  absorption lines decreases in the physical vicinity of galaxies. It has been suggested by these authors that either strong galactic winds diminish the amount of neutral hydrogen, or that the emission of  $\text{Ly}\alpha$  photons effectively replaces the absorbed flux.

### 4.5.3 Line profile fitting

As first proposed by Carswell et al. (1984), the absorption lines seen in quasar spectra can be very well modeled via one- or many-component Voigt profiles (see sections 2.2.1 and 2.2.2). Hereby, however, a high spectral resolution of  $R = 10,000$  at  $\lambda = 10$  nm or higher is required. The profile fitting is far from trivial, since it depends on the determination of the flux continuum. Furthermore, departures from pure Voigt profiles may arise under certain physical conditions, such as internal rotation, collapse, galactic outflows, bulk motions, etc., which would be hard to detect, even at the highest resolution, but which would give important clues about the physical conditions of the absorbing medium.

By fitting of a given profile, information about the column density and Doppler parameter of single absorbers can be obtained with the help of a curve-of-growth analysis (cf. Sec. 2.2.2). As in the line-counting method, the numbers of lines in a different column density-, Doppler parameter, and redshift ranges are determined and from those, the statistical properties, *i.e.* the distribution of the absorbers as a function of these parameters can be computed.<sup>19</sup>

### 4.5.4 The column density distribution function (CDDF)

The column density distribution function for Ly $\alpha$  absorbers, which is analogous to the Initial Mass Function (IMF) for star formation—or the luminosity function of galaxies (Kim et al. 2001)—can be directly computed from the differential distribution function of absorbers (eq. 4.27) by integrating over a given redshift interval<sup>20</sup>. However, in order to get rid of the explicit  $(1+z)$  dependence, it has become customary to introduce a simple change of variables, expressing the redshift  $z$  in terms of the so-called absorption distance  $\chi$  (Bahcall & Peebles 1969, Tytler 1987), which can be defined by (Wolfe et al. 2005)

$$d\chi \equiv \frac{H_0}{c} (1+z)^3 \left| \frac{dl}{dz} \right| dz. \quad (4.32)$$

With this new variable, equation (4.27) becomes

$$dN(N_{\text{HI}}, \chi) = f(N_{\text{HI}}, \chi) dN_{\text{HI}} d\chi, \quad (4.33)$$

<sup>19</sup> There are several line-fitting algorithms available in the literature, such as AUTOVP (Davé et al. 1997) or FITLYMAN (Fontana & Ballester 1995), used to obtain line parameters such as redshift, column density, and Doppler width from observed absorption Voigt profiles.

<sup>20</sup> In principle, by integrating over  $[0, \infty)$

and the column density distribution function (CDDF) is then defined as

$$f(N_{\text{HI}}, \chi) \equiv \frac{c}{H_0} n_0 [N_{\text{HI}}, \chi(z)] \sigma_A [N_{\text{HI}}, \chi(z)]. \quad (4.34)$$

An operational definition of the column-density distribution is given by (Tytler 1987)

$$f(N_{\text{HI}}) = \frac{m}{\Delta N_{\text{HI}} \sum_i \Delta\chi_i}, \quad (4.35)$$

where  $m$  is the number of absorbers observed in the column density range  $(N_{\text{HI}}, N_{\text{HI}} + \Delta N_{\text{HI}})$  obtained from a sample of  $n$  QSO with total redshift coverage  $\sum_i^n \Delta\chi_i$ , and  $\chi$  is the absorption distance defined above.

Empirically, a power-law of the form

$$f(N_{\text{HI}}) dN_{\text{HI}} d\chi = B N_{\text{HI}}^{-\beta} dN_{\text{HI}} d\chi, \quad (4.36)$$

with  $B \equiv m \cdot (1-\beta) / [(N_{\text{HI}})_{\text{max}}^{1-\beta} - (N_{\text{HI}})_{\text{min}}^{1-\beta}]$  is found to match well the column-density distribution from observations, as first noted by Carswell et al. (1984) and later confirmed by many authors (Tytler 1987, Sargent et al. 1989, Lanzetta 1991, Rauch et al. 1992, Hu et al. 1995, Kim et al. 1997, Kirkman & Tytler 1997, Janknecht et al. 2002, Kim et al. 2002b), and from simulations (Riediger et al. 1998), with  $\beta \approx 1.51$  for over about ten orders of magnitude in column density, from  $N_{\text{HI}} = 10^{12} \text{ cm}^{-2}$  to  $N_{\text{HI}} = 10^{22} \text{ cm}^{-2}$ . As mentioned in the previous section, numerical simulations of structure formation based on the  $\Lambda$ CDM cosmology predict a similar result. However, it has also been found that a featureless, single power-law is a poor fit to the data (*e.g.* Carswell et al. 1987, Cristiani et al. 1995, Wolfe et al. 2005), and measured departures from a single power-law at high column densities, *i.e.*  $N_{\text{HI}} > 10^{14.3} \text{ cm}^{-2}$  have also been reported in the literature (Kim et al. 1997, Petitjean et al. 1993). In such cases, a broken power-law  $f(N_{\text{HI}}, \chi) = B (N_{\text{HI}}/N_b)^{-\beta}$  with  $\beta = \alpha_1$  for  $N_{\text{HI}} < N_b$  and  $\beta = \alpha_2$  for  $N_{\text{HI}} \geq N_b$  (Penton et al. 2004), or in the case of DLAs a  $\Gamma$ -function  $f(N, \chi) = B (N/N_b)^\alpha \exp(-N_{\text{HI}}/N_b)$  (Wolfe et al. 2005) have been found to better fit the data. The break appears to be always in the range  $N_{\text{HI}} = 10^{14-15} \text{ cm}^{-2}$ , and, as noted by Press & Rybicki (1993), this is most possibly related to the fact that the Ly $\alpha$  lines saturate precisely at  $N_{\text{HI}} \approx 10^{14} \text{ cm}^{-2}$ , which makes the determination of their column densities quite uncertain.

A previous, alternative description of the strength of the absorption lines made use of the equivalent

width spectrum  $n(W, z)$ , defined in such a way that the  $n(W, z) dW dz$  gives the number of clouds in an equivalent width interval  $(W, W + dW)$  and redshift range  $(z, z + dz)$ . An exponential function, independent of  $z$ , of the form

$$n(W) dW = \frac{N^*}{W^*} e^{-W/W^*} dW \quad (4.37)$$

was found to match well the observations by Sargent et al. (1980), Young et al. (1982), Murdoch et al. (1986). Here,  $N^*$  and  $W^*$  are constants with typical values  $W^* = 0.362 \pm 0.021 \text{ \AA}$  and  $N^* = 154 \pm 11$  (Sargent et al. 1980). However, this approach has the drawback of lacking information about the physical state of the absorbers, which is coded in the quantities  $N_{\text{HI}}$  and  $b$ , and whose relation to  $W$  can be ambiguous.

It was first shown by Murdoch et al. (1986) that the exponential distribution (4.37) is mapped into a power-law distribution in  $N_{\text{HI}}$

$$n(N_{\text{HI}}) = B N_{\text{HI}}^{-s} dN_{\text{HI}}, \quad (4.38)$$

where  $s = 1 + k/W^*$ , for equivalent widths in the logarithmic part of the curve of growth<sup>21</sup> for which the equivalent width  $W$  is related to the column density  $N_{\text{HI}}$  for any given value of the Doppler parameter  $b$  by

$$W = k \ln(N_{\text{HI}}/N^*), \quad (4.39)$$

with  $k = k(b)$ .

This can easily be shown by requiring that  $n(W)dW \stackrel{!}{=} n(N_{\text{HI}})dN_{\text{HI}}$ , which is equivalent to

$$n(N_{\text{HI}}) dN_{\text{HI}} = n[W(N_{\text{HI}})] dN_{\text{HI}} \left| \frac{dW}{dN_{\text{HI}}} \right|. \quad (4.40)$$

Using the equation (4.39) yields

$$\frac{dW}{dN_{\text{HI}}} = \frac{k N^*}{N_{\text{HI}}}, \quad (4.41)$$

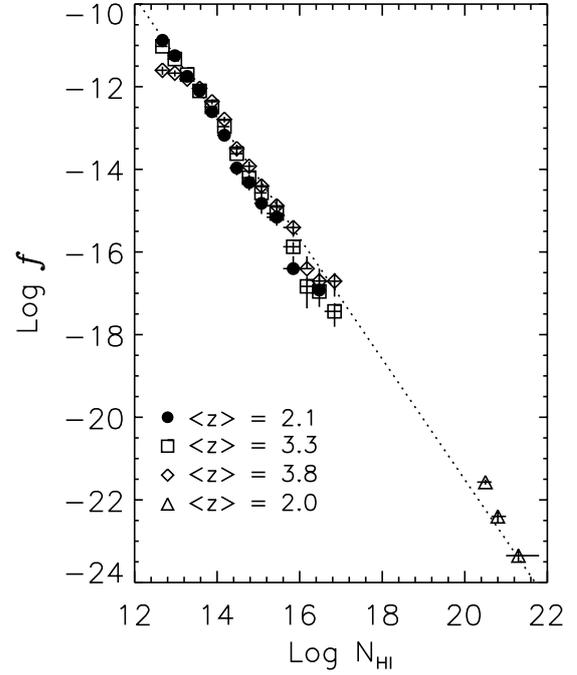
and plugging into equation (4.38) it finally follows that

$$\begin{aligned} n(N_{\text{HI}}) dN_{\text{HI}} &= \frac{k}{N_{\text{HI}}} \frac{(N^*)^2}{W^*} e^{-\frac{k}{W^*} \ln(N_{\text{HI}}/N^*)} \\ &= B N_{\text{HI}}^{-(1+k/W^*)}, \end{aligned} \quad (4.42)$$

with  $B \equiv k (N^*)^2 / W^*$ .

Thus, both approaches are to some extent equivalent. However, the description of the Ly $\alpha$  absorbers in terms

<sup>21</sup> For Ly $\alpha$ , this range corresponds to equivalent widths in the range  $W \in (0.3, 1.0) \text{ \AA}$



**Fig. 4.7:** Column density distribution function (CDDF) of the Ly $\alpha$  absorbers at different redshifts for various compilations. The dotted line represents the CDDF from Hu et al. (1995),  $f(N_{\text{HI}}) = 4.9 \cdot 10^7 N_{\text{HI}}^{-1.46}$ . Taken from Kim et al. (2002a, and references therein).

of their column–density distribution is the most suitable since it is directly related to the physical properties of the absorbers. The best estimate of the parameters  $B$  and  $\beta$  seems to be that of Hu et al. (1995):

$$f(N_{\text{HI}}) = 4.9 \cdot 10^7 N_{\text{HI}}^{-1.46}, \quad (4.43)$$

which describes quite well the observations over a large column density range, as can be seen in Fig. 4.7.

It has become the practice to include the number density evolution of the Ly $\alpha$  absorbers (equation 4.31) and the column density distribution function (equation 4.34) into one single expression of the form

$$f(N_{\text{HI}}, z) = \mathcal{N}_0 \cdot (1+z)^\gamma \cdot N_{\text{HI}}^{-\beta}. \quad (4.44)$$

Here  $\mathcal{N}_0$  is a normalization constant, chosen in such a way that

$$\frac{dN}{dz} \stackrel{!}{=} \int_{I_{N_{\text{HI}}}} f(N_{\text{HI}}, z) dN_{\text{HI}}, \quad (4.45)$$

where the left-hand-side is given by equation (4.31), and  $I_{N_{\text{HI}}}$  is an appropriate column-density interval. The function (4.44) defines the 1-dimensional distribution of the HI in the IGM probed by a random line-of-sight.

Using equation (4.33), the mean number  $\langle N_{abs} \rangle$  of systems for each line-of-sight can be expressed as

$$\langle N_{abs} \rangle = \int_{I_z} \int_{I_{N_{\text{HI}}}} f(N_{\text{HI}}, z) dN_{\text{HI}} dz, \quad (4.46)$$

where, the integral is carried out over appropriate redshift- and column density intervals  $I_z$  and  $I_{N_{\text{HI}}}$ , respectively. These intervals are usually constrained from observations. It is common however to group the absorbers into column density ranges corresponding to the empirical classification in Ly $\alpha$  forest, LLSs, and DLAs, giving for each class an evolution function of the form of equation (4.44). An example of such a set of distribution functions is shown in chapter 6, section 6.3, table 6.1.

#### 4.5.5 The Doppler parameter distribution

The Doppler parameter, as mentioned in 2, is defined as

$$b = \sqrt{\frac{2kT}{m_{\text{H}}}}. \quad (4.47)$$

where  $m_{\text{H}}$  is the mass of the hydrogen atom<sup>22</sup>. A very useful relation can be obtained by substitution of the values for  $k$  and  $m_{\text{H}}$  in the above equation, leading to

$$b [\text{km s}^{-1}] = 0.128 \cdot \sqrt{T[K]}. \quad (4.48)$$

The values of the Doppler parameter  $b$  for intergalactic neutral hydrogen are usually obtained from line profile fitting, in which the pair of values ( $N_{\text{HI}}, b$ ) that leads to the best fit and best reproduces the measured equivalent width of the line is estimated using curve-of-growth analysis and a minimisation technique, *e.g.*  $\chi^2$ -minimisation.<sup>23</sup> In what follows, a brief review of the range of values for  $b$  found to date is given.

Some of the first determinations of the widths of the Ly $\alpha$  lines imprinted on the spectra of QSO were performed by Carswell et al. (1984), who found typical values for  $b$  in the range (10, 45) km s $^{-1}$ , which according to the above expression correspond to temperatures in the range  $10^4 K < T < 10^5 K$ . Later on, a narrower  $b$

distribution peaking at values 17 km s $^{-1}$  and not extending beyond 30 km s $^{-1}$  was found by Pettini et al. (1990), while a couple of years later a broad distribution, peaking at values in the range 30 – 40 km s $^{-1}$  and a tail extending to higher velocities was reported by Rauch et al. (1992). In particular, these last authors found a median of  $b = 33 - 35$  km s $^{-1}$  for  $N_{\text{HI}} > 10^{13.3}$  cm $^{-2}$ . As they pointed out, the  $b$  values in excess of 20 km s $^{-1}$  may be due to blends of unresolved narrow lines with  $b \approx 20$  km s $^{-1}$ . Otherwise, those larger values may indicate non-thermal bulk motions. In this case, high-resolution, high S/N observations should show asymmetries or other departures from pure Voigt profiles. The first step towards solving this controversy was made by Press & Rybicki (1993), who proposed the  $b$  distribution to be a truncated Gaussian:

$$p(b) db \propto \begin{cases} \exp\left[-\frac{(b-b_0)^2}{2b_*^2}\right] db & , \quad b > 0 \\ 0 & , \quad b < 0 \end{cases}$$

or a gamma function of the form

$$p(b) db \propto b^{(b_0/b_*)-1} \exp\left(-\frac{b}{b_*}\right) db, \quad b > 0, \quad (4.49)$$

with  $b_0$  the (mode) mean of the (truncated) distribution and  $b_*$  the standard deviation. They found that the  $b$ -distribution derived from observations is consistent with a broad  $b$ -distribution, most probably a truncated Gaussian with a mean around 37 km s $^{-1}$ , implying that a significant component of  $b$  is due to bulk motion, or that the absorbing systems are not in equilibrium and significantly hotter than implied by their ionisation state.

Further analyses reported a peak value of  $b = 23$  km s $^{-1}$  corresponding to a thermal temperature  $T = 3.2 \cdot 10^4$  (Cristiani et al. 1995), and confirmed that the  $b$  distribution is well represented by a truncated Gaussian with a mean around 28 km s $^{-1}$  and  $\sigma_b = 10$  km s $^{-1}$  (Hu et al. 1995). These authors also find that the  $b$ -distribution is rather invariant as a function of  $N_{\text{HI}}$ . Later on, Kim et al. (1997) found that the  $b$ -distribution evolves with redshift, where the (mean)  $b$  value increases with decreasing redshift, a result that is in agreement with CDM numerical simulations (Riediger et al. 1998). Kim et al. (1997) established from observations that the distribution of Doppler parameter for the absorbing systems is well described by a truncated, Gaussian with a median around 35 km s $^{-1}$ , a dispersion  $\sigma(b) \approx 12$  km s $^{-1}$ , and a truncation below

<sup>22</sup>  $m_{\text{H}} = 1.673 \cdot 10^{-27}$  kg

<sup>23</sup> This technique, although rather common, is in the words of Press & Rybicki (1993), "[...] a somewhat controversial subject."

$b_c \approx 24 \text{ km s}^{-1}$  at  $z = 2.31$ , and they also found that this values evolve slightly with redshift. In a later work, [Kirkman & Tytler \(1997\)](#) found that the lower cut-off in the  $b$ -distribution increases with  $N_{\text{HI}}$ , while [Kim et al. \(2001\)](#) found that it increases with decreasing redshift for a fixed  $N_{\text{HI}}$ . They also found that the  $b$ -distribution is well fitted by a Gaussian with  $\langle b \rangle = 23 \text{ km s}^{-1}$  and  $\sigma_b = 14 \text{ km s}^{-1}$ . [Bershady et al. \(1999\)](#), for example, adopt from [Kim et al. \(1997\)](#) the following parametrization:  $\langle b \rangle = -3.85 z_{em} + 38.9 \text{ km s}^{-1}$ ,  $\sigma(b) = -3.85 z_{em} + 20.9 \text{ km s}^{-1}$  and  $b_c = -6.73 z_{em} + 39.5 \text{ km s}^{-1}$  for absorbers in the redshift range  $z \in (1.75, 5)$ .

However, [Hui & Rutledge \(1999\)](#) find a parametrization of the  $b$  distribution in the form

$$\frac{dN}{db} = b_0 \frac{b_\sigma^4}{b^5} \exp\left(-\frac{b_\sigma^4}{b^4}\right), \quad (4.50)$$

where  $N$  is the number of absorption systems,  $b_0$  is a normalisation constant,  $b_\sigma$  is a parameter related to the average amplitude of the fluctuations in the optical depth (see [Kim et al. \(2001\)](#) for values of these parameters derived from observations). From [Kim et al. \(2001\)](#), we compute  $\langle b_0 \rangle = 7.24$ ,  $\langle b_\sigma \rangle = 24.11$  for absorbers in the redshift range  $\langle z \rangle \in [1.61, 3.75]$ .

Interestingly, there is also an on-going debate on a possible correlation between  $N_{\text{HI}}$  and  $b$ . For example, [Cristiani et al. \(1995\)](#) find a deficit of lines with  $b < 20 \text{ km s}^{-1}$  for absorbers with  $N_{\text{HI}} > 10^{13.5} \text{ cm}^{-2}$ . However, [Rauch et al. \(1992\)](#) find no evidence for any significant  $b - N_{\text{HI}}$  correlation. Supporting the latter result, [Hu et al. \(1995\)](#) find the  $b$ -distribution to be quite invariant as a function of column density. In a recent paper, [Williger et al. \(2006\)](#) report a Doppler distribution for low- and high-column density absorbers at low redshift with mean, median, and standard deviations of  $\{44, 44, 22\} \text{ km s}^{-1}$  and  $\{47, 47, 21\} \text{ km s}^{-1}$ , respectively. Also, [Lehner et al. \(2006\)](#) find that, for Broad Line Absorbers (BLA)—which correspond to absorbers with  $b > 40 \text{ km s}^{-1}$  which are thought to compose the so-called warm-hot intergalactic mediums (WHIM)—, the  $b$  values appear to decrease with increasing  $N_{\text{HI}}$ . These results are nevertheless contradictory, and this controversy has not been clarified to date.

Assuming that the  $b$ -distribution is well defined by its mean, and dispersion values, it appears that all of the measurements mentioned above are roughly compatible with each other. There is still no consensus on the best parametrization for the  $b$ -distribution, and thus in this work, and for the sake of simplicity, we will adopt two different  $b$ -distributions for the attenuation

models presented in the next chapters: Following [Madau \(1995\)](#), we will assume a constant Doppler parameter of  $b = 35 \text{ km s}^{-1}$ ; and we will also explore the effect of assuming that the  $b$ -distribution is given by a truncated Gaussian distribution as reported by [Bershady et al. \(1999\)](#).

There are two further quantities that characterise the Ly $\alpha$  absorbers, and more precisely, the intergalactic neutral hydrogen: the mean optical depth, and the cosmic flux decrement.

#### 4.5.6 The mean optical depth or effective optical depth

The optical depth, as a function of wavelength, is given by ([Press et al. 1993](#))

$$\tau(\lambda) = -\ln\left(\frac{f_{obs}(\lambda)}{f_c(\lambda)}\right), \quad (4.51)$$

where  $f_c$  is the continuum level and  $f_{obs}$  the observed flux.

The mean optical depth is usually parameterised as

$$\bar{\tau}(z) = A (1+z)^{\gamma+1}, \quad (4.52)$$

where  $A$  is a normalisation constant and  $\gamma$  is the same evolution parameter as for the number density of lines (eq. 4.31). It has been found by [Press et al. \(1993\)](#) that  $\gamma = 2.46 \pm 0.37$  and  $A = 0.0175 - 0.0056 \gamma \pm 0.0002$  for absorbers in the redshift range  $2.5 < z < 4.3$ .

Since the mean optical depth cannot be determined for  $f_{obs} \approx 0$ , a further quantity of interest is the so-called effective optical depth, defined by

$$\tau_{\text{eff}} = -\ln(\langle e^{-\tau} \rangle), \quad (4.53)$$

where the brackets denote the mean value averaged over the wavelength range of interest (in principle, in the range  $[0, \infty)$ ). [Kim et al. \(2001\)](#) find that the effective optical depth can be parametrized in the form of equation (4.52) with  $\tau_{\text{eff}}(\lambda) = (0.0030 \pm 0.0008)(1+z)^{3.43 \pm 0.17}$  for  $1.5 < z < 4$ , using a sample of 3 QSOs observed with VLT/UVES. This result was verified by [Kim et al. \(2002b\)](#), who increased the former sample to a total of eight QSO and found that  $\tau_{\text{eff}}(\lambda) = (0.0032 \pm 0.0009)(1+z)^{3.37 \pm 0.20}$  for the same redshift interval.

The mean or effective optical depth is a measure for the amount of neutral hydrogen present at a given redshift, and its determination is of particular interest with

**Table 4.2:** Geometry of Ly $\alpha$  absorbers, inferred from observations along the lines-of-sight to QSO pairs and/or to multiple images of one and the same lensed QSO (Rauch 1998).

$N_{\text{HI}} \lesssim 10^{14} \text{ cm}^{-2}$	$10^{14} \text{ cm}^{-2} \lesssim N_{\text{HI}} \lesssim 10^{16} \text{ cm}^{-2}$	$N_{\text{HI}} \gtrsim 10^{16} \text{ cm}^{-2}$
thin and extended	filamentary	spherical
$100 h^{-1} \text{ kpc} - 1 h^{-1} \text{ Mpc}$	$40 - 100 h^{-1} \text{ kpc}$	$1 h^{-1} \text{ kpc}$
(length)	(width)	(radius)

respect to the estimate of the total amount of baryonic matter, and also in order to pin down the epoch of reionisation. The optical depth at the epoch of reionisation thus constrains models for the evolution of the Universe, in particular, it is one of the six free parameters of the  $\Lambda$ CDM model.

#### 4.5.7 Cosmic flux decrement

The cosmic flux decrement is a useful quantity, in particular if the available resolution of the observed spectra is relatively low, *i.e.*  $\Delta\lambda \approx 1 \text{ nm}$ . It is defined as (Oke & Korycansky 1982)

$$D_A \equiv \left\langle 1 - \frac{f_{\text{obs}}}{f_c} \right\rangle. \quad (4.54)$$

The average is computed *per definitionem* in the rest-frame wavelength range (102.5, 121.6) nm, *i.e.* between the Ly $\alpha$  and the Ly $\beta$  emission lines. The cosmic flux decrement can be viewed as the equivalent width of the total absorption in the above range. It effectively measures the depression of the observed flux  $f_{\text{obs}}$  with respect to the continuum  $f_c$ , as *e.g.* extrapolated from the red wing of the Ly $\alpha$  emission line to the range blueward of it. A widely adopted form for the extrapolated continuum is a power-law with spectral index  $\alpha$ . The uncertainty of the extrapolated continuum is the main drawback of this method. However, at sufficient resolution, it turns out to be very useful when estimating the amount of baryons contained in the IGM as a function of redshift.<sup>24</sup>

#### 4.5.8 Structure of the Ly $\alpha$ Forest

The "true" structure of the Ly $\alpha$  forest is still a matter of debate, even though a significant amount of knowledge has been gained in the last decades, especially

through the observation of QSO pairs and the analysis of the absorption seen along their corresponding lines-of-sight, or through observations of absorption systems along the line-of-sight to multiple images of one and the same QSO due to gravitational lensing. These studies show that Ly $\alpha$  forest clouds are clustered on scales  $30 - 50 h^{-1} \text{ Mpc}$ , and  $60 - 130 h^{-1} \text{ Mpc}$  (Rauch 1998).

Furthermore, it has been found that the geometry of the absorbers is correlated to their column density. Absorbers with column densities  $N_{\text{HI}} \gtrsim 10^{16} \text{ cm}^{-2}$  have a spherical geometry while filamentary structures typically give rise to absorption lines that correspond to column densities in the range  $10^{14} \text{ cm}^{-2} \lesssim N_{\text{HI}} \lesssim 10^{16} \text{ cm}^{-2}$ . Absorbers with column densities  $N_{\text{HI}} \lesssim 10^{14} \text{ cm}^{-2}$  are rather thin and extended. Thus, Ly $\alpha$  forest absorbers are filamentary and/or extended, typically  $30 h^{-1} \text{ kpc}$  across and with lengths around  $1 h^{-1} \text{ Mpc}$  (Rauch & Haehnelt 1995, see Table 4.2).

The distribution of Ly $\alpha$  absorbers along the line-of-sight shows large voids (*e.g.* Pierre et al. 1988, Duncan et al. 1989, Dobrzycki & Bechtold 1991), and the absorbers trace the large-scale distribution of the luminous matter (*i.e.* of galaxies) at low redshift. However, the correlation between absorbers and galaxies, especially at high redshift, is still controversial. In some cases, a correlation between tidal arms or galactic winds and absorption features has been found (see *e.g.* Williams et al. 2005, Part 1), but most of the Ly $\alpha$  forest systems seem to originate in filamentary structures or HI clouds with little or no star-formation activity. Their low abundances imply that no or, at most, very little star-formation can have taken place in them.

<sup>24</sup> See chapter 6 for a detailed analysis of the evolution of  $D_A$ .



---

# Modelling the absorption of light in the intergalactic medium: The Monte Carlo Approach

---

The intergalactic neutral hydrogen randomly distributed along the line-of-sight to a distant source, (*e.g.* quasar, galaxy) absorbs some its light and affects its spectrum in a way that is by no means trivial: on the one hand, the absorption is wavelength-dependent due to the energy dependence of the resonant transitions, and on the other, it is highly sensitive to the distribution and properties (temperature, density) of the absorbing hydrogen, as we have seen in Chapter 2 on the dependence of the absorption coefficient on the Doppler parameter and the column density. We already advanced in Chapter 4 that it is of particular importance to account for the absorbing effect of the intergalactic neutral hydrogen, also called *intergalactic attenuation*, when modelling galaxy spectra to correctly interpret observations. A brief review on the subject will be given in the introduction of the next chapter.

In this chapter, we focus on a detailed description of a model for the intergalactic attenuation that we have developed on the basis of the models of Møller & Jakobsen (1990), Madau (1995) and Bershadsky et al. (1999). The main goal here is to account for the stochastic absorption due to the random distribution of H I and its properties. In the next chapters, we will apply this model to account for the evolution and scatter of the cosmic flux decrement, and to assess its impact on the photometric properties of galaxies.

## 5.1 THE PATH OF RADIATION THROUGH THE IGM

A photon with wavelength  $\lambda_\gamma$  that is emitted at redshift  $z_{em}$  will lose energy, *i.e.* will be redshifted on its way as a consequence of the expansion of the Universe. Hence, the photon will be observed at a wavelength  $\lambda_{obs}$ , where

$$\lambda_{obs} = \lambda_\gamma \cdot (1 + z_{em}). \quad (5.1)$$

If the photon happens to encounter an H I absorber at redshift  $z_{abs}$ , and assuming no broadening mechanism in the absorbing medium of any kind, the photon will be scattered if and only if its wavelength at the redshift of the absorber happens to match the wavelength corresponding to one of the resonant transitions of H I, *i.e.* if

$$\lambda_j = \lambda_\gamma \cdot (1 + z_{abs}), \quad (5.2)$$

where the subscript  $j$  refers to the Lyman transitions  $\{\alpha, \beta, \gamma, \dots\}$  (cf. Chapter 2). Due to the effect of broadening, however, the photon will be scattered off the line-of-sight with a probability given by the absorption coefficient (2.21) evaluated at  $\lambda = \lambda_\gamma \cdot (1 + z_{abs})$ . Photons emitted with energies  $\lambda_\gamma \leq \lambda_L / (1 + z_{abs})$  will ionise the hydrogen atoms of the absorber with a probability given by the photoionisation cross-section (2.3). Thus, a source at redshift  $z_{em}$  emitting a photon flux  $f_{em}(\lambda)$  which encounters an absorber at redshift  $z_{abs}$  along the line-of-

sight will be observed to have a flux at  $z = 0$  given by

$$f_{obs}(\lambda_{obs}) = \frac{f_{em}(\lambda_{em})}{1 + z_{em}} e^{-\tau(\lambda_{abs})}, \quad (5.3)$$

where  $\lambda_{obs}$ ,  $\lambda_{em}$  and  $\lambda_{abs}$  are the observed wavelength, the emitted wavelength, and the wavelength at absorption, respectively. These are related by

$$\lambda_{obs} = \lambda_{abs} \cdot (1 + z_{abs}) = \lambda_{em} \cdot (1 + z_{em}). \quad (5.4)$$

The factor  $(1 + z_{em})^{-1}$  in equation (5.3) is a consequence of the expansion of the Universe and the requirement that the flux be conserved. The absorption coefficient  $\tau$  is given by the sum of the photionisation cross-section and the individual absorption coefficients for each transition, *i.e.*

$$\tau(\lambda) = \tau_{phot}(\lambda) + \sum_j \tau_j(\lambda), \quad (5.5)$$

where  $\tau_{phot}$  and  $\tau_j$  are given by equations (2.14) and (2.21), respectively, and the sum is over all transitions. The effect a single absorber on a constant spectrum  $f_\lambda$  normalised to unity is illustrated in Fig. 5.1. In this example, the absorber has a column density  $N_{HI} = 10^{17.22} \text{ cm}^{-2}$ , which corresponds to an optical depth slightly higher than unity at the Lyman edge, and a Doppler parameter  $b = 36 \text{ km s}^{-1}$ . We include the first 25 Lyman transitions. Worth noting is the strong flux depression at wavelengths  $913 < \lambda < 914 \text{ \AA}$  as a consequence of the cumulative absorption and the blending of the higher-order transitions. This, however, is in part due to the finite resolution of the spectrum.

It is now easy to see that the flux emitted at all wavelengths by a background source will be absorbed at a discrete set of wavelengths for each absorber along the line-of-sight. If the intergalactic neutral hydrogen were uniformly distributed along the line-of-sight, the continuously redshifted radiation would encounter a scattering atom at every point, producing an absorption trough in the observed spectrum. This is the so-called Gunn-Peterson trough, mentioned in Chapter 4. However, the absence of such a trough in QSO spectra up to redshifts  $z_{em} > 6$  points to the fact that intergalactic neutral hydrogen is **not** uniformly distributed, at least not after the reionisation epoch, as explained in previous chapters.

The effect of a population of absorbers along the line-of-sight is computed in a similar manner as (5.3). In this case, the full set of parameters  $\{z_{abs}, N_{HI}, b\}_{N_{abs}}$ , *i.e.* the

absorber redshift, the column density, and Doppler parameter characterising each absorber is required. Assuming this to be given, the flux observed at  $z = 0$  of a source at  $z_{em}$  with an intrinsic flux  $f_{em}$  is given by

$$f_{obs}(\lambda) = \frac{f_{em}[\lambda/(1 + z_{em})]}{1 + z_{em}} \Phi(\lambda), \quad (5.6)$$

where the quantity  $\Phi$  is the transmission factor and is given by

$$\Phi(\lambda) \equiv \prod_{i=1}^{N_{abs}} \exp[-\tau(\lambda/(1 + z_i))]. \quad (5.7)$$

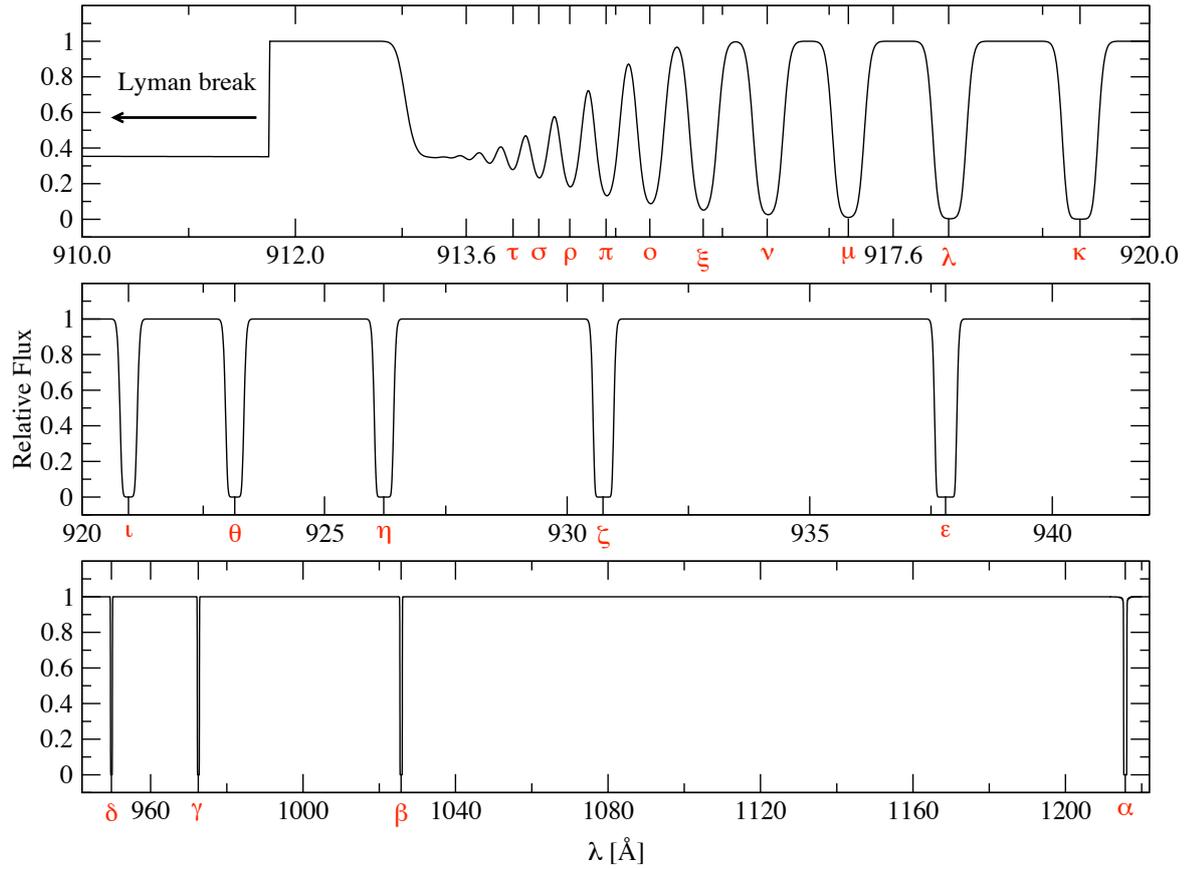
Here,  $z_i$  is the redshift at the epoch of absorption<sup>1</sup>, and  $N_{abs}$  is the number of absorbers encountered along the line-of-sight.

## 5.2 GAMBLING WITH THE Ly $\alpha$ ABSORBER PARAMETERS

In order to account for the absorption of the flux of a source at a given redshift along a random line-of-sight, a detailed model of the distribution of intergalactic HI and its properties ( $z_{abs}, N_{HI}, b$ ) is needed. More precisely, the exact number of absorbers along a particular line-of-sight and the parameters characterising each absorber need to be known. As already explained in the previous chapter, the distribution of the intergalactic neutral hydrogen and its physical properties are constrained from observations and given in a statistical sense by the density distribution functions of the form of equation (6.4). A possible approach to model the absorption along random line-of-sight is thus to generate a large number of synthetic lines-of-sight, each of which shall be characterised by a unique population of absorbers. A particular population is completely defined by the total number of absorbers  $N_{abs}$ , and the set of their parameters  $\{z_{abs,i}, N_{HI,i}, b_i\}_{i=1,\dots,N_{abs}}$ .

The natural choice for the realisation of this approach is the performance of numerical simulations which make use of the Monte Carlo method, as has been done by several authors (see *e.g.* Møller & Jakobsen 1990, Giallongo et al. 1990, Cristiani et al. 1993, Bershady et al. 1999). This method is suitable for the analysis of the behaviour of a complex random variable whose distribution function is not known. In our case,

<sup>1</sup>The reader shall bear in mind that  $\tau$  in equation (5.7) is in general different for each absorber. However, we do not write this explicitly by *e.g.* introducing a new subscript in order to avoid confusion.



**Fig. 5.1:** Effect of a single absorber on an input spectrum  $f_\lambda = \text{const.}$  normalised to unity. The absorption includes the photoionisation and the first 25 resonant Lyman transitions, but only those from Ly $\alpha$  to Ly $\tau$  are indicated for simplicity. The absorber parameters are  $(N_{\text{HI}}, b) = (10^{17.22} \text{ cm}^{-2}, 36 \text{ km s}^{-1})$ . Wavelengths are given in the absorber restframe. Note the strong absorption at wavelengths redward of the Lyman edge due to the blending and the cumulative effect of the highest order transitions. This effect is in part due to the finite resolution of the spectrum.

we want to analyse the behaviour of the transmission factor  $\Phi$  (equation 5.7). This quantity is a composite, and certainly, complex random variable which depends through the optical depth on the random variables  $N_{\text{HI}}$ ,  $z_{\text{abs}}$ , and  $b$ . The dependence of  $\Phi$  on these variables is rather complex, and an analytical expression for the distribution function for it is completely out of sight, even though there are special cases for which this is possible (see *e.g.* Madau 1995, and references therein)<sup>2</sup>

The good news is that the distribution function of each of the variables on which the transmission coefficient  $\Phi$  depends is known. Hence, the behaviour of the transmission coefficient for a given redshift  $z_{\text{em}}$  can be analysed by generating random sets of values  $\{z_{\text{abs}}, N_{\text{HI}}, b\}$  and evaluating with them equation (5.7). The ensemble of different realisations of  $\Phi$  computed in this way effectively mimics the absorption of light due to intergalactic neutral hydrogen along random line-of-sight, as long as the distribution functions (4.31), (4.36), and (5.22) (see below) are accurately determined from observations.

This procedure assumes that the random variables  $z_{\text{abs}}$ ,  $N_{\text{HI}}$ , and  $b$  are statistically independent. Apart from a possible weak correlation between  $b$  and  $N_{\text{HI}}$ , mentioned in section 4.4, it appears as if they truly are independent from each other. Furthermore, we assume that random numbers distributed according to (4.31), (4.36), and (5.22) can be generated. That this is indeed the case is briefly explained in the next section.

### 5.2.1 Some Notes on "arbitrarily"-distributed random numbers

From the theory of probability it is known that, given a random variable  $X$  distributed according to  $P_X$ , another (arbitrary) random variable  $Y$ , which is a function of the former, *i.e.*  $Y = Y(X)$ , is distributed according to<sup>3</sup>

$$P_Y(y) = P_X[x(y)] \left| \frac{dx}{dy} \right| (y). \quad (5.8)$$

In other words, the transformation  $f : x \mapsto y(x)$  invokes the functional transformation  $F : P_X \mapsto P_Y$ .

This is an inverse problem and in general it is very difficult to find the transformation  $y = y(x)$  for arbitrary functions  $P_X$  and  $P_Y$ . Nevertheless, there are special

<sup>2</sup>We will show in the next chapter that it is possible to determine analytically the form of the distribution function for  $\Phi$  using mathematical and physical arguments.

<sup>3</sup>This follows directly from the normalization condition of a probability distribution and the chain-rule.

cases of great interest in which this problem can be solved exactly.

Let  $X$  be a uniformly distributed random variable on the interval  $[0, 1]$ . The corresponding distribution function is given by

$$P_X(x) = \Theta(x) \cdot \theta(1-x), \quad (5.9)$$

where the so-called Heaviside-Function  $\Theta(x)$  is defined as

$$\Theta(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 & , \quad x \geq 0 \end{cases}$$

In this particular case, equation (5.8) becomes

$$P_Y(y) = \theta[x(y)] \cdot \theta[1-x(y)] \left| \frac{dx}{dy} \right| (y),$$

or

$$P_Y(y) = \left| \frac{dx}{dy} \right| (y) \quad (5.10)$$

for  $x(y) \in [0, 1]$ .

For a given probability distribution  $P_Y$ , all one has to do is to integrate equation (5.10) in order to find the transformation  $x = x(y)$ , so that  $Y$  is distributed according to  $P_Y$ , with  $X$  uniformly distributed on the interval  $[0, 1]$ . In other words, writing

$$F(y) = \int_{y_{\text{min}}}^y P_Y(y') dy', \quad (5.11)$$

where  $y_{\text{min}}$  is the minimum value that the variable  $Y$  takes on, it follows that

$$y(x) = F^{-1}(x). \quad (5.12)$$

Note that the range of this transformation has to be chosen in such a way that  $F(y) \in [0, 1]$  for  $y \in [y_{\text{min}}, y_{\text{max}}]$ , where  $y_{\text{max}}$  is the maximum value of  $y$ . Nevertheless, this is automatically guaranteed by equation (5.11) since  $P_Y$  should be normalised.

### Power-law distributions

Generally, a (distribution-) function of the form

$$P(x) = C \cdot x^\alpha, \quad C > 0, \alpha \in \mathbb{R}, \quad (5.13)$$

where  $C$  is a normalization constant, is called a power-law (distribution-) function.

This type of distribution very often appears in physical problems, for example, in the description of the distribution of neutral hydrogen clouds along random

lines-of-sight in the intergalactic medium, as mentioned in chapter 4. It is thus of interest for the present work to briefly show here some useful results related to this kind of distribution, such as a general expression for the normalization constant  $C$  and the way in which random numbers distributed according to (5.13) for arbitrary real values of  $\alpha$  may be generated. This last point is essential for the numerical simulations of the effect of Ly $\alpha$  absorbers on the spectra of background sources (quasars, galaxies, etc.) that we will present.

Let  $X$  be a random variable distributed according to (5.13) in the intervall  $[a, b]$ . It follows, by the requirement that the probability density be normalised, that

$$1 \stackrel{!}{=} \int_a^b P(x) dx = C \cdot \left[ \frac{1}{1+\alpha} x^{1+\alpha} \Big|_a^b \right] \quad (5.14)$$

for  $\alpha \neq -1$ , and

$$1 \stackrel{!}{=} \int_a^b P(x) dx = C \cdot \left[ \log x \Big|_a^b \right] \quad (5.15)$$

for  $\alpha = -1$  and  $0 < a, b < \infty$ . In these cases it follows that

$$C = (1 + \alpha) [b^{1+\alpha} - a^{1+\alpha}]^{-1}, \quad (5.16)$$

and

$$C = [\log b - \log a]^{-1}, \quad (5.17)$$

respectively. In what follows we will only consider this first case, since the power-law distribution functions related to the properties of intergalactic H I all satisfy the requirement  $\alpha \neq -1$  (cf. Sec. 4.4).

#### Generating power-law-distributed random numbers

According to what was presented in the last section, one can easily find a transformation  $y = y(x)$  for a random variable  $Y$  distributed according to equation (5.13).

Let us assume that  $y \geq 0$ , *i.e.* the range of values that the variable  $Y$  takes on is semi-positive. It follows, according to equation (5.11), that

$$F(y) = \int_{y_{\min}}^y P_Y(y') dy' = C \cdot \frac{1}{(1+\alpha)} [y^{1+\alpha} - y_{\min}^{1+\alpha}],$$

where  $y_{\min} \leq y \leq y_{\max}$ , and the normalization constant  $C$  is given by equation (5.16),

$$C = (1 + \alpha) [y_{\max}^{1+\alpha} - y_{\min}^{1+\alpha}]^{-1}. \quad (5.18)$$

Furthermore, we get from equation (5.12) that

$$y(x) = F^{-1}(x) = \left[ \frac{(1+\alpha)}{C} x + y_{\min}^{1+\alpha} \right]^{\frac{1}{1+\alpha}}. \quad (5.19)$$

By substituting  $C$  into this last equation one finally gets

$$y(x) = \left[ y_{\max}^{1+\alpha} x + y_{\min}^{1+\alpha} (1-x) \right]^{\frac{1}{1+\alpha}}. \quad (5.20)$$

It is easily seen from this equation that  $y(x)$  takes values in the intervall  $[y_{\min}, y_{\max}]$  when  $x$  takes values in the intervall  $[0, 1]$ . Furthermore, since  $0 \leq x \leq 1$  and  $y_{\min}$  and  $y_{\max}$  are both semi-positive, the argument of the root in expression (5.19) is always semi-positive, and therefore this expression is well defined for all  $\alpha \in \mathbb{R}^+$ . In other words, there is no restriction for  $\alpha$  except for being real and positive.<sup>4</sup> Equation (5.20) is thus the desired transformation for a random variable  $X$  *uniformly* distributed in the intervall  $[0, 1]$ , in order to generate a random variable  $Y$  distributed according to  $P_Y(y) = C \cdot y^\alpha$  in the intervall  $[y_{\min}, y_{\max}]$ .

### 5.3 NUMERICAL SIMULATIONS OF THE INTERGALACTIC ATTENUATION

We finally have all the ingredients for our model of the intergalactic attenuation, which we briefly summarize here.

The population of Ly $\alpha$  absorbers along a random line-of-sight to a distant source at a redshift  $z_{em}$  is characterized by its number density distribution  $dN/dz$ , its differential distribution function  $f(N_{\text{HI}})$ , and the distribution of its Doppler parameter (cf. Sec. 4.4). The number density distribution, which gives the number of absorbers per unit redshift, is usually parameterized as (equation 4.31)

$$\frac{dN}{dz} = A \cdot (1+z)^\gamma,$$

where  $A$  is a constant and  $\gamma$  is called the evolution parameter, since it gives the rate of evolution with redshift of the Ly $\alpha$  absorbers.

The differential distribution function gives the number of absorbers per unit redshift path and per unit column density, as a function of column density. It is empirically given by (equation 4.36)

$$f(N_{\text{HI}}) = B N_{\text{HI}}^{-\beta},$$

where  $B$  is a normalization constant and  $\chi$  is the redshift path as defined by equation (4.32). Both these equations

<sup>4</sup> The case  $\alpha = 0$  is excluded for obvious reasons, namely, the fact that a power-law distribution with  $\alpha = 0$  is trivially a uniform distribution.

are valid for any given  $b$ -distribution. Thus, for the moment, we assume the Doppler distribution  $p(b)$  to be an arbitrary probability distribution function normalised to unity, *i.e.*

$$\int_0^{\infty} p(b) db = 1. \quad (5.21)$$

For instance, the Doppler parameters may be distributed according to (cf. section 4.4)

$$p(b) = \Theta(b - b_{trunc}) \cdot \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{1}{2\sigma^2}(b-\mu)^2} \quad (5.22)$$

*i.e.* that they show a Gaussian (normal) distribution truncated below  $b_{trunc}$ .

The composite distribution function for the parameters  $(z_{abs}, N_{\text{HI}}, b)$  of the absorbers along a random line-of-sight is thus given by (equation 4.44)

$$N(z, N_{\text{HI}}) dz dN_{\text{HI}} = N_0 \cdot N_{\text{HI}}^{-\beta} (1+z)^\gamma dz dN_{\text{HI}}.$$

where the  $b$ -values are distributed according to  $p(b)$ .

The *expected* number of Ly $\alpha$  absorbers within a given column-density interval  $[N_{min}, N_{max}]$  out to  $z_{em}$  is given by (equation 4.46)

$$\langle N_{abs} \rangle(z_{em}) = \int_0^{z_{em}} \int_{N_{min}}^{N_{max}} N(z, N_{\text{HI}}) dz dN_{\text{HI}}. \quad (5.23)$$

The division into different column-density intervals is important because of the fact that the evolution of absorbers with different column densities, *i.e.* the evolution of the different types of absorbers (Ly $\alpha$  forest, LLSs, DLAs) is different with redshift (see section 4.4).

#### 5.4 COOKBOOK FOR A LINE-OF-SIGHT WITH A RANDOM POPULATION OF Ly $\alpha$ ABSORBERS

Given all the ingredients listed in the last section, all we need now to complete the model is random number uniformly distributed. There are several algorithms in the literature with which so-called *pseudo-random* numbers uniformly distributed in the interval  $[0, 1]$  can be generated. In particular, we use here the subroutines `ran1` and `ran3` from the *Numerical Recipes*. Assuming that we have a random variable  $X$  uniformly distributed on the interval  $[0, 1]$ , the absorption due to intergalactic H $\text{I}$  by applying the Monte Carlo method can be modeled following these steps:

**1** Compute the *average* number of absorbers  $\langle N_{abs} \rangle$  out to a given redshift  $z_{em}$  and in a given column density range according to equation (5.23). Using the results from last section, it follows by direct integration of this equation that

$$\langle N_{abs} \rangle(z_{em}) = N_0 \cdot (C_z \cdot C_{N_{\text{HI}}})^{-1}, \quad (5.24)$$

where

$$C_z \equiv \frac{1+\gamma}{(1+z_{em})^{1+\gamma} - 1}, \quad (5.25)$$

and

$$C_{N_{\text{HI}}} \equiv \frac{1-\beta}{(N_{\text{HI},max})^{1-\beta} - (N_{\text{HI},min})^{1-\beta}}. \quad (5.26)$$

**2** Compute the *actual* number of absorbers  $N_{abs}$  along the  $i$ th line-of-sight by drawing from a Poisson distribution with parameter  $\langle N_{abs} \rangle$ . For this purpose, we make use of the algorithm `poidev` to generate Poisson distributed random numbers given by the *Numerical Recipes*.

**3** Assign to each absorber a set of parameters  $(z_{abs}, N_{\text{HI}}, b)$ , where the column density  $N_{\text{HI}}$  and the absorption redshift  $z_{abs}$  are drawn from a distribution of the form of equation (6.4), integrated over  $[0, z_{em}]$  and the corresponding column-density interval, respectively. Again, it is straightforward from equations (5.19) and (5.18) that the redshift is given by

$$z_{abs}(x) = \left[ \frac{(1+\gamma)}{C_z} x + 1 \right]^{\frac{1}{1+\gamma}} - 1 \quad (5.27)$$

In a similar manner, the column density  $N_{\text{HI}}$  can be computed from

$$N_{\text{HI}}(x) = \left[ \frac{(1-\beta)}{C_{N_{\text{HI}}}} x + (N_{\text{HI},min})^{1-\beta} \right]^{\frac{1}{1-\beta}}. \quad (5.28)$$

The Doppler parameters are distributed among the absorbers according to  $p(b)$ .

**4** Compute the absorption on a particular model galaxy input spectrum  $f_\lambda$  according to equations (5.6) and (5.7)

5 Repeat the procedure for as many lines-of-sight as desired

The number of lines-of-sight in the ensemble is to some extent arbitrary, but in order to get reliable results, a number of lines-of-sight  $N_{los} \geq 10^3$  is a good choice. In order to improve the statistics, we choose  $N_{los} \geq 4 \cdot 10^3$  through all the work that follows.

An application of the method just described to the modelling of the cosmic flux decrement (see chapter 4) is given in the next chapter. The effect of the intergalactic absorption due to H $\gamma$  on the photometric properties of galaxies is accounted for using this model in chapter 7. Also, in this chapter, the model galaxy spectra used are described. Furthermore, in both these chapters the different sets of input distribution functions of the absorbers properties are presented.



---

# Stochastic Absorption of the Light of Background Sources due to Intergalactic Neutral Hydrogen

## I. Testing different line-number evolution models via the cosmic flux decrement

The contents of this chapter are on the review stage prior to publication. They can be found as Tepper-García, T. & Fritze, U. 2007, ArXiv e-prints, 705

---

### ABSTRACT

We test the accuracy of different models of the attenuation of light due to resonant scattering by intergalactic neutral hydrogen by comparing their predictions of the evolution of the mean cosmic flux decrement  $D_A$  to measurements of this quantity based on observations. To this end, we use data available in the literature and our own measurements of the cosmic flux decrement for 25 quasi-stellar sources in the redshift range  $2.71 < z_{em} < 5.41$  taken from the SDSS Data Release 5. In order to perform the measurements of  $D_A$ , we fit a power-law to the continuum redward of the Ly $\alpha$  emission line, and extrapolate this fit to region blueward of it, where the flux is severely affected by absorption due to intervening H I absorbers.

We compute, using numerical simulations, the redshift evolution of the mean flux depression,  $D_A(z)$ , due to the presence of Ly $\alpha$  Forest absorbers and Lyman limit systems randomly distributed along the line-of-sight, and compute its intrinsic scatter at the 1-, 2-, and 3  $\sigma$  level due to fluctuations in the absorber properties (column density, Doppler parameter, redshift) along different lines-of-sight. The numerical simulations consist of Monte Carlo realizations of distributions of the

absorber properties constrained from observations.

We compare our and previous measurements of the cosmic flux decrement to the outcomes of our simulations and find an excellent agreement between the observations and the predictions of one of the models considered in this work. Furthermore, we find that the distribution of the  $D_A$  values at a given redshift are well described by a lognormal distribution function. This implies that the effective optical depth, usually defined as the logarithm of the average flux, is necessarily Gaussian distributed, in contrast to previous studies. This result is independent to the form of the input distribution functions, and rather insensitive to the presence of high-column density absorbers, such as the Lyman limit systems.

**keywords** methods: numerical, intergalactic medium, quasars: absorption lines

### 6.1 INTRODUCTION

Since the introduction of the Gunn-Peterson (GP) test by Gunn & Peterson (1965), a detailed knowledge about the physical state of the intergalactic medium (IGM) has

been gained from the study of the absorption features identified in the spectra of quasi-stellar objects (QSOs) at restframe wavelengths  $\lambda \leq 121.5$  nm, which are now known to be mainly due to resonant scattering by intergalactic neutral hydrogen, randomly distributed along the line-of-sight, as first proposed by Lynds (1971). For instance, the null result in the search for a GP trough has been used to rule out the existence of a hot intercloud medium (ICM) (Steidel & Sargent 1987a, Giallongo et al. 1992, 1994), which was thought to confine by pressure the Ly $\alpha$  clouds (Sargent, Young, Boksenberg, & Tytler 1980, Ostriker & Ikeuchi 1983). As a result of detailed analyses of the line statistics of the absorbing material, a wealth of information on its clustering (see e.g. Ostriker et al. 1988), in particular the existence of voids (e.g. Pierre et al. 1988, Duncan et al. 1989, Dobrzycki & Bechtold 1991), and the evolution of its number densities, column densities, and Doppler parameter with redshift (e.g. Kim et al. 1997) has accumulated over the past years. These results, in combination with the use of state-of-the-art numerical simulations of structure formation based on the currently accepted paradigm of the concordance cosmology (Springel et al. 2005), show that the features seen in absorption against bright background sources arise when the line-of-sight intersects the structures that naturally emerge and evolve with time under the influence of gravitational attraction. Different types of structures such as the filaments present in the *cosmic web*, galactic haloes, and even the discs of primeval galaxies, give rise to distinct absorption features attributed to entities known historically as Ly $\alpha$  Forest clouds, Lyman limit systems (LLSs), and damped Lyman absorbers (DLAs) (see e.g. Rauch 1998, Wolfe et al. 2005, for excellent reviews, respectively). Nevertheless, the relation between the observed absorption features and the objects causing them, in particular the correlation between observed damped absorption lines, metal lines (e.g. MgII, OVI) and galaxies—the so-called Absorber-galaxy Connection—is still a matter of debate (see e.g. Williams et al. 2005, Part 1). As a consequence of numerous efforts over many years, we now have a better understanding of the origins of the different absorption features observed in QSO spectra. In particular, the notion of discrete, intervening HI absorbing systems randomly distributed along the line-of-sight has been embedded into the more general picture of an evolving *continuous* intergalactic medium with a HI density field that varies in space and time, with its evolution driven mainly by the Hubble expansion, the radiation field of

ionising UV sources, and the collapse of structures due to gravity.

### 6.1.1 Methods and Input Distributions: A Brief Review

Over decades many people have been working hard towards inferring the physical properties of the intergalactic medium such as its chemical content, density, temperature, etc. by measuring e.g. the type of transition, strength, number density, and profiles of absorption lines imprinted in the spectra of QSOs and Gamma-Ray bursts (GRBs) (e.g. Lamb & Reichart 2000). There has also been a great effort to quantify the effect of the absorption due to intergalactic neutral hydrogen on the photometric properties of background sources. As a matter of fact, several models have been developed in order to account for this so-called intergalactic attenuation, with different approaches and purposes. Møller & Jakobsen (1990) used Monte Carlo simulations to estimate the amount of absorption at wavelengths shorter than the redshifted H $\alpha$   $\lambda$ 30.4 nm line, in order to test the feasibility of the equivalent of the Gunn-Peterson test for intergalactic helium. They found that the absorption as a function of wavelength, averaged over many lines-of-sight, should display together with a characteristic stair-case profile due to the cumulative absorption at the H $\alpha$  resonant wavelengths, an additional characteristic valley-shaped feature (the “Lyman-valley”) due to the cumulative effect of the photoionisation of H $\alpha$  by photons with energies  $E_\gamma \geq hc/\lambda_L$ , where  $\lambda = 91.2$  nm. Later on, in a seminal paper Madau (1995) developed an analytical method to quantify the opacity due to intergalactic H $\alpha$  as a function of redshift, and its effect on the colors of high-redshift galaxies. The underlying assumption of this model is that the observed flux of a source at redshift  $z$  is given by the product of the *intrinsic* flux and a transmission factor that accounts for the *mean* absorption as a function of wavelength given in the form  $\exp(-\tau_{\text{eff}}) \equiv \langle \exp(-\tau) \rangle$  (see equation (3) of Madau 1995), where the brackets denote the average over an ensemble of random lines-of-sight. The most common application of this model consists in correcting the flux of a synthetic spectrum for intergalactic absorption. This correction is of particular importance at high redshift, where the absorption due to intergalactic H $\alpha$  severely absorbs the light of a background object at restframe wavelengths shorter than 121.6 nm, leading to a substantial reddening of its colour (see e.g. Bicker et al. 2004). As the numerous references in the literature attest, the Madau model has

become the most widely used attenuation model. However, in a later work [Bershady et al. \(1999\)](#) argued that it is not possible to estimate the mean change in the magnitude of a source at a given redshift due to absorption by intergalactic H<sub>I</sub> along the line-of-sight by multiplying the mean transmission curve of [Madau's](#) model with the spectrum of the source and integrating over the corresponding passband, mainly because of the existence of color terms. They suggested that the correct way of accounting for the mean effect of H<sub>I</sub> absorption on the spectrum of a background source and on its photometric properties, is to model first the absorption along many random lines-of-sight, compute the desired photometric quantities for each one of them, and *then* compute the average over all lines-of-sight. In other words, they argue that the processes of averaging over many random lines-of-sight and measuring photometric quantities are non-commutative. Indeed, they showed using a Monte Carlo technique that the average magnitudes computed following their approach substantially differ from those computed using [Madau's](#) model, even when using the same input distributions for the number of absorbers, their column densities and Doppler parameters. The approach proposed by [Bershady, Charlton, & Geoffroy \(1999\)](#) effectively mimics the measurement process that would take place if one would determine *e.g.* the mean observed brightness of a collection of galaxies with different absorber populations along their particular lines-of-sight, but otherwise identical in their intrinsic properties (spectrum, morphology, etc.), and is hence physically meaningful. It turns out that features such as the characteristic stair-case profile and the Lyman-valley cannot possibly be observed in a *single* spectrum, since they arise only by averaging over sufficient numbers of lines-of-sight, a process that has no physical meaning. Surprisingly, however, the [Bershady, Charlton, & Geoffroy \(1999\)](#) approach is rarely referenced or used in the literature.

In a more recent paper, [Meiksin \(2006\)](#) developed a method to compute the opacity due to intergalactic H<sub>I</sub> by using hydrodynamical simulations of structure formation in the framework of the concordance  $\Lambda$ CDM cosmology performed by [Meiksin & White \(2004\)](#). Applying his model to compute broad-band magnitudes for different types of objects (*e.g.* starburst galaxies, QSOs of Type I and II), [Meiksin \(2006\)](#) reports differences of 0.5 – 1.0 mag with respect to [Madau \(1995\)](#)'s model. Despite of the different results obtained, this model is similar to [Madau \(1995\)](#)'s model in the sense that it implicitly assumes that the mean opacity of the

IGM along a random line-of-sight due to the presence of H<sub>I</sub> can be accounted for by multiplying a given input spectrum with a mean attenuation curve of the form  $\exp(-\tau_{\text{eff}})$  and integrating over the corresponding filter function ([Meiksin 2006](#), equation 8).

Following [Bershady, Charlton, & Geoffroy \(1999\)](#), we state that

$$\int_0^\infty f_\lambda \langle \exp(-\tau) \rangle T(\lambda) d\lambda \neq \left\langle \int_0^\infty f_\lambda \exp(-\tau) T(\lambda) d\lambda \right\rangle, \quad (6.1)$$

where  $f_\lambda$  is the intrinsic flux,  $T(\lambda)$  is the filter transmission function, and the brackets indicate the average over all lines-of-sight. We consider that the operation denoted by the right-hand side of this expression is the correct way of estimating mean magnitudes of background objects including the effect of the absorption due to intergalactic H<sub>I</sub>. This approach is of course not restricted to the computation of mean magnitudes and colors, and can be applied to the estimate of the mean of any photometric quantity. Furthermore, it is also possible to determine not only the mean, but in principle any desired confidence interval around the mean, *e.g.*  $\pm\sigma$  range, via the computation of quantiles (see Section 6.5).

It should be clear that not only the method, but also the input physics is an (even more) crucial ingredient of a particular model that accounts for the intergalactic attenuation, as already shown by [Bershady, Charlton, & Geoffroy \(1999\)](#). It is, however, not trivial to test whether using a particular method and a set of input distributions accurately describes the observed effect of the absorption by intergalactic H<sub>I</sub> on the spectra of background sources. For example, the evolutionary synthesis models of [Bicker et al. \(2004\)](#) that include the correction for intergalactic absorption based on [Madau's](#) model match quite well the observations of galaxies in the Hubble Deep Field (see his Figure 12), since the magnitude differences reported by ([Bershady, Charlton, & Geoffroy 1999](#), Figure 7) with respect to the latter model are in this case of the order of the scatter of the observations around the predicted colors. In other words, even though these models are fundamentally different, it is difficult to test the accuracy of their predictions on the basis of a comparison to *e.g.* observed galaxy colors. A quantity that is more sensitive to the absorption due to intergalactic H<sub>I</sub> is the mean cosmic flux decrement  $D_A$  (cf. Section 6.2). The reason for this is that the restframe wavelength range over which this quantity is measured is typically 10 nm wide, and is

hence narrower than typical broadband filters. We thus consider as a primer test that any model that accounts for the absorption due to intergalactic H I should reproduce first of all the observations of this quantity. On this basis, it should be possible to discriminate between models through the comparison of their respective predictions to measurements of  $D_A$ . As we will show, the input distributions used by [Bershady, Charlton, & Geoffroy \(1999\)](#) to estimate changes in the magnitude of background galaxies do not match the observations of  $D_A$  very well, and thus the results of that model regarding galaxy colors should be taken with caution. We will show that [Madau's](#) model and [Bershady, Charlton, & Geoffroy \(1999\)](#)'s model are complementary to each other in the sense that the input physics of the first, with the approach of the second, truly match the measurements of the cosmic flux decrement.

To sum up: The main goal of this work is to compute and analyse the evolution of  $D_A$  with redshift for different evolution scenarios of the intergalactic neutral hydrogen, conveniently parametrized by input distribution functions of the form of equation (6.4, cf. Section 6.2). Thus, the models we present here differ only by the set of input distribution functions used, but they all equal in method, *i.e.* all of them take advantage of the Monte Carlo technique. We judge the goodness of a particular model by its power to reproduce the observations of  $D_A$  in a wide redshift range. Once we identify such a model, we exploit it to analyse some interesting properties of  $D_A$ . In a forthcoming paper, we will make use of such a model to assess the impact of the stochastic absorption due to intergalactic H I on the photometric properties of high-redshift galaxies.

This work is organised as follows: In Section 6.2 we briefly recall the concept of the cosmic flux decrement and discuss some issues related to its measurement. In Section 6.3 we present two different types of models for the intergalactic attenuation, which we use to compute the redshift evolution of the cosmic flux decrement. In Section 6.4, we describe our measurements of this quantity for a sample of SDSS QSO spectra. Finally, we compare these and previous measurements to the outcomes of each model, and discuss the results of this comparison as well as some other implications of the models for the evolution of  $D_A$  in Section 6.5.

## 6.2 THE COSMIC FLUX DECREMENT REVISITED

Before high-resolution (*i.e.*  $\Delta\lambda \lesssim 1$  nm), high S/N observation became feasible, the basic spectroscopic technique used to analyse the effect of the absorption due to intergalactic neutral hydrogen on the spectra of background sources was to measure the mean depression of the observed flux relative to the unabsorbed flux—also called continuum—, a quantity which became to be known as cosmic flux decrement. This quantity, first introduced by [Oke & Korycansky \(1982\)](#), can be defined as a function of redshift by

$$D_A(z) \equiv \frac{1}{\Delta\lambda} \int_{\lambda_1 \cdot (1+z)}^{\lambda_2 \cdot (1+z)} \left( 1 - \frac{f_{obs}(\lambda)}{f_c(\lambda)} \right) d\lambda, \quad (6.2)$$

where  $f_c$  and  $f_{obs}$  are the continuum and the observed fluxes, respectively, and  $\Delta\lambda \equiv (1+z) \cdot (\lambda_2 - \lambda_1)$ . Formally, the integral is computed in the restframe wavelength range [102.5, 121.6] nm, *i.e.* between the Ly $\alpha$  and the Ly $\beta$  emission lines. However, the actual estimate of  $D_A$  is usually performed between the restframe wavelengths 105 nm and 117 nm—or in an even narrower wavelength interval—in order to avoid contamination by the emission wings of the Ly $\beta$  + OVI and Ly $\alpha$  lines, respectively.

The cosmic flux decrement effectively measures the *total* equivalent width of all Ly $\alpha$  absorption lines in the chosen wavelength range, if corrected for the contribution of metal absorption lines. This idea has indeed been used by [Zuo & Lu \(1993\)](#) in order to measure this quantity by adding up the equivalent widths of lines identified as Ly $\alpha$  absorption lines in a given wavelength range, thus avoiding the contamination from metal lines. Of course, the reliability of this measurement highly depends on the correct identification of lines, a task that is not trivial at all.

Since  $D_A$  is extremely sensitive to  $f_c$ , as can be easily seen from the definition (6.2), an accurate measurement of this quantity demands a reliable estimate of the underlying continuum. Unfortunately, there is no consensus of what the best method to estimate the continuum may be. A popular choice, mainly because of the presence of emission lines in the Ly $\alpha$  forest region, consists in fitting a local continuum, most commonly using cubic splines (see *e.g.* [Lu, Sargent, Womble, & Takada-Hidai 1996](#)) or b-spline functions (see *e.g.* [Kirkman, Tytler, Suzuki, O'Meara, & Lubin 2003](#), [Tytler, Kirkman, O'Meara, Suzuki, Orin, Lubin, Paschos, Jena, Lin, Norman, & Meiksin 2004a](#),

Tytler, O’Meara, Suzuki, Kirkman, Lubin, & Orin 2004b), searching for regions apparently free of absorption blueward of the Ly $\alpha$  emission line. Other authors prefer to fit a continuum in the region redward of the red wing of the Ly $\alpha$  emission line, and extrapolate it to the region blueward of it (Steidel & Sargent 1987b, Schneider, Schmidt, & Gunn 1989, Cristiani, Giallongo, Busson, Gouiffes, & La Franca 1993). A widely adopted form for the fitted continuum in this case is a power-law with spectral index  $\alpha_\nu$ , with measured values in the range [0.28, 0.99] (Steidel & Sargent 1987b, Vanden Berk et al. 2001, and references therein). The latter method usually tends to place the intrinsic continuum level higher than it actually is, thus overestimating the measured values of  $D_A$ ; for the former method the opposite is true, in general. For either method, there is an uncertainty in the estimate of the continuum, and this is the main drawback of the mean flux depression as a technique to estimate the mean absorption due to neutral hydrogen present in the intergalactic medium (IGM). However, it turns out that reliable measurements of  $D_A$  are very useful to constrain estimates of fundamental cosmological parameters such as the mean baryon density  $\Omega_b$ , the UV background intensity (see e.g. Rauch et al. 1997), the normalization of the power spectrum  $\sigma_8$ , the vacuum-energy density  $\Omega_\Lambda$ , and the Hubble parameter  $H_0$  (Tytler et al. 2004a).

The evolution of the cosmic flux decrement has been previously modeled by different workers (see e.g. Giallongo et al. 1990, Cristiani et al. 1993, Madau 1995), usually obtaining a good agreement with observations. However, there is still a scatter in the observations of this quantity for which it has not been accounted yet in any modelling so far. Using our Monte Carlo simulations of the absorption due to intergalactic H I along a large number of lines-of-sight, we assess to which extent the observed scatter can be ascribed to the *intrinsic* scatter in  $D_A$  due to fluctuations in the absorber properties (number density, column density, Doppler parameter) along different lines-of-sight (see Section 6.5.3).

Under the assumption that the restframe equivalent width of the absorbers does not evolve with redshift, and that the number density of the absorbing systems evolves like  $\propto (1+z)^\gamma$ , It is expected that  $D_A$  should evolve with  $z$  like

$$D_A(z) \propto (1+z)^{1+\gamma} \quad (6.3)$$

where the factor  $(1+z)$  comes from the scaling of the equivalent width, as pointed out by Jenkins & Ostriker

(1991). Indeed, it has been found empirically that the redshift evolution of  $D_A$  can be described by a power law  $D_A(z) = A \cdot (1+z)^\gamma$  with  $A = 6.2 \cdot 10^{-3}$  and  $\gamma = 2.75$  (Kirkman et al. 2005), but also by an exponential  $D_A(z) = D_A^0 \cdot e^{\alpha(1+z)}$  with  $D_A^0 = 0.01$  and  $\alpha = 0.75$  (Zhang et al. 1997). More recently, Kirkman et al. (2007) showed that the observed evolution of  $D_A$  with redshift in the range  $0 < z < 3.2$  is well described by a broken power-law, even though the fit is still poor. None the less, expression like these are only valid up to a given redshift, since it diverges for  $z \rightarrow \infty$ , whilst  $D_A$  converges asymptotically to 1 in this limit, or more precisely, when  $z$  approaches the redshift  $z_{reion}$  at which reionisation sets on. Furthermore, even if the power-law form for the evolution of  $D_A$  holds, the index  $\gamma$  in the last equation should be replaced by  $\bar{\gamma}$ , where the latter index accounts for the averaged evolution of absorbers of different column densities. As we know now, different types of absorbers evolve at different rates, and an equation of the form of equation (6.3) would imply that they evolve at the same rate, independently of their column density. On the other hand, estimates of a single  $\gamma$  from  $D_A$  measurements assuming a power-law of the form of equation (6.3), as done by O’Brien et al. (1988), may give a hint on the population of absorbers *dominating* the behaviour of  $D_A$ , comparing the estimated  $\gamma$  with the power-law index of the different populations. We will show from our simulations that the predicted redshift evolution of  $D_A$  satisfies the asymptotic behaviour described above, and that, indeed, this behaviour is dominated by the absorbers with column densities  $N_{HI} \lesssim 10^{17} \text{ cm}^{-2}$ .

A compilation of  $D_A$  measurements, accumulated in the literature over the past twenty years approximately, and which includes our own measurements that extend the redshift range to  $z_{em} = 5.41$ , is shown in Figure 6.2 (cf. Section 6.4 and cited references for details on the measurements in this figure).

### 6.3 MODELLING THE INTERGALACTIC ATTENUATION

Since the observation of individual sources (galaxies, QSOs, GRBs) necessarily implies observations along different lines-of-sight, it is expected that the stochastic nature of the distribution of the Ly $\alpha$  absorbers, especially of those with the highest column densities, causes a scatter in the observed absorption, even for sources with identical intrinsic spectra. Hence, depend-

ing on the absorption along a particular line-of-sight, one would expect different observed values for each measurement of any photometric quantity, for example, the cosmic flux decrement  $D_A$ . Performing enough measurements of such a quantity for nearly "identical" sources at a fixed redshift, one could in principle estimate its mean and its scatter due to stochastic effects in the absorption by neutral hydrogen in the IGM.

The numerical realisation of this thought experiment is best achieved through Monte Carlo simulations. Following *e.g.* Møller & Jakobsen (1990), Giallongo, Gratton, & Trevese (1990), Cristiani, Giallongo, Buson, Gouiffes, & La Franca (1993), Bershady, Charlton, & Geoffroy (1999), we generate thousands ( $4 \cdot 10^3$ ) of lines-of-sight each with a random population of HI absorbers, and compute the absorption along each of them for a given input spectrum at a fixed redshift. The population of each line-of-sight consists of a random number  $N_{abs}$  of absorbing systems, each of them characterized by three parameters: its redshift  $z_{abs}$ , its column density  $N_{HI}$ , and its Doppler parameter  $b \equiv \sqrt{2kT/m_H}$ , where  $k$  is the Boltzmann constant,  $T$  is the kinetic temperature of the gas and  $m_H$  is the mass of the hydrogen atom.

The redshift and column density characterising each absorber are drawn from a distribution of the form

$$f(N_{HI}, z) = \mathcal{N}_0 \cdot (1+z)^\gamma \cdot N_{HI}^{-\beta}, \quad (6.4)$$

where  $\mathcal{N}_0$  is a normalization constant. This function defines the 1-dimensional distribution of the HI present in the IGM probed by a random line-of-sight. The number  $N_{abs}$  of systems for each line-of-sight is drawn from a Poisson distribution with parameter

$$\langle N_{abs} \rangle = \int_{I_z} \int_{I_{N_{HI}}} f(N_{HI}, z) dN_{HI} dz, \quad (6.5)$$

where the integral is carried out over appropriate redshift- and column density intervals  $I_z$  and  $I_{N_{HI}}$ , respectively.

We use different sets of input distributions that include the evolution of both low- and high density absorbers, and that give rise to the following models:

**MMC** This model relies on the input distributions from Madau (1995, equation 10) listed in Table 6.1. Here, the Doppler parameter is kept constant at a value  $b = 35.0 \text{ km s}^{-1}$ , which corresponds approximately to the mean derived by Rauch et al. (1992).

**Table 6.1:** Types of absorbers and their corresponding parameters adopted from Madau (1995, equation 10). Note, however, that he quotes as lowest column density  $N_{HI} = 2.0 \cdot 10^{12}$ , while we use  $N_{HI} = 1.0 \cdot 10^{12}$ , in order for the adopted normalisation to be consistent.

$N_{HI} [\text{cm}^{-2}]$	$\mathcal{N}_0$	$\gamma$	$\beta$
$10^{12} - 1.59 \cdot 10^{17}$	$2.40 \cdot 10^7$	2.46	1.50
$1.59 \cdot 10^{17} - 10^{20}$	$1.90 \cdot 10^8$	0.68	1.50

**Table 6.2:** Types of absorbers and their corresponding parameters. Parameters are adopted from Bershady et al. (1999, equation 10).

$N_{HI} [\text{cm}^{-2}]$	$\mathcal{N}_0$	$\gamma$	$\beta$
$10^{12} - 10^{14}$	$3.14 \cdot 10^7$	1.29	1.46
$10^{14} - 1.59 \cdot 10^{17}$	$1.70 \cdot 10^6$	3.10	1.46
$1.59 \cdot 10^{17} - 10^{20}$	$1.90 \cdot 10^8$	0.68	1.50

**BMC** This model matches the model named MC-Kim of Bershady, Charlton, & Geoffroy (1999), but we briefly describe it here for completeness. The corresponding parameters for the line-density evolution and column density distribution functions are summarized in Table 6.2. In this model, in contrast to the MMC model, the Doppler parameter for each absorber is drawn from a truncated, redshift-dependent Gaussian distribution of the form

$$P(b) \equiv \Theta(b - b_{tr}) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(b - \mu)^2\right),$$

where  $\Theta(x)$  is the Heaviside function:

$$\Theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

and the mean, standard deviation and truncation value at redshift  $z$  are given by  $\mu(z) = -3.85z + 38.9$ ,  $\sigma(z) = -3.85z + 20.9$  and  $b_{tr}(z) = -6.73z + 39.5$ , respectively. Bershady, Charlton, & Geoffroy (1999) originally used this model to analyse the impact of the intergalactic attenuation in the range  $1.75 < z < 5.0$ , but we use it in the extended range  $0.2 < z < 5.41$ . Since our highest

redshift limit is not too far away from [Bershady, Charlton, & Geoffroy](#)'s, we consider the model to be valid in our extended redshift range.

**MMC without Lyman limit systems** In order to assess the impact of the Lyman limit systems on the intergalactic absorption, we introduce this model, which consists of the same input distributions as the MMC model, excluding the systems with column densities  $N_{\text{HI}} > 1.59 \cdot 10^{17} \text{ cm}^{-2}$ .

In all the models describe above, the attenuation factor for each absorber is given by  $\exp[-\tau(\lambda)]$ , where, for the general case, the absorption coefficient  $\tau(\lambda)$  can be written as

$$\tau(\lambda) = \tau_{LL}(\lambda) + \sum_{j=2}^{N_{\text{trans}}} \tau_j(\lambda). \quad (6.6)$$

The first term on the right-hand side is the opacity due to the ionisation of neutral hydrogen by photons with wavelengths  $\lambda \leq \lambda_L \equiv 91.18 \text{ nm}$ . It is given by

$$\tau_{LL}(\lambda) = N_{\text{HI}} \cdot g(\lambda) \cdot \sigma_{\infty} \cdot \left(\frac{\lambda}{\lambda_L}\right)^3, \quad (6.7)$$

where  $\sigma_{\infty} \equiv 6.31 \cdot 10^{-18} \text{ cm}^2$  is the H $\alpha$  photoionisation cross-section, and  $g$  is the Gaunt-factor for bound-free transitions.<sup>1</sup> The second term is the sum of the opacities due to resonant scattering at each transition of the Lyman series<sup>2</sup>. In general, the absorption coefficient for the transition  $n = i \rightarrow 1$  is

$$\tau_i(\lambda) = N_{\text{HI}} \cdot \sigma_i \cdot \phi(a_i, x). \quad (6.8)$$

The cross-section  $\sigma_j$  is a function of the Doppler parameter  $b$ , the oscillator strength of the transition  $f_j$ , and the resonant wavelength  $\lambda_j$ , and is given by

$$\sigma_j = \frac{\sqrt{\pi} e^2}{m_e c^2} \frac{\lambda_j^2}{\Delta\lambda_D} f_j, \quad (6.9)$$

where  $\Delta\lambda_D = \lambda_j b/c$  is the Doppler broadening, and the variable  $x \equiv (\lambda - \lambda_j)/\Delta\lambda_D$  is the distance to the line

<sup>1</sup> An extensive tabulation of values for the Gaunt-factor can be found in [Karzas & Latter \(1961\)](#).

<sup>2</sup> We adopt the convention that the Ly $\alpha$  transition (from the ground state to the next higher energy level) be identified with  $i = 2$ , the Ly $\beta$  transition with  $i = 3$ , etc. The photoionisation cross-section is thus consistently denoted by  $\sigma_{\infty}$ .

center in Doppler units. We assume the profile function  $\phi$  of the absorption line to be given by the Voigt-Hjerting function

$$H(a_j, x) \equiv \frac{a_j}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(x-y)^2 + a_j^2} dy. \quad (6.10)$$

Here,  $a_j \equiv \lambda_j^2 \Gamma_j / (4\pi \Delta\lambda_D)$  is the relative strength of the natural broadening to Doppler broadening for the  $i$ th transition, and  $y \equiv v/b$  is the kinetic velocity in units of the Doppler parameter. In this work, we neglect the opacity due to the photoionisation term and consider only the first resonant transition, *i.e.* the Ly $\alpha$  transition, since this is the only one of interest in the wavelength range studied here. Furthermore, we use the approximation to  $H$  for values of  $a$  and column densities characteristic for intergalactic H $\alpha$  of [Tepper-García \(2006\)](#).

### 6.3.1 The transmission factor $\Phi$

The cumulative absorption along a random line-of-sight of the flux  $f_{em}$  of a source at redshift  $z_{em}$  is calculated according to expression

$$f_{obs}(\lambda_{obs}) = \frac{f_{em}(\lambda_{em})}{1 + z_{em}} \cdot \Phi(\lambda_{obs}), \quad (6.11)$$

where  $\lambda_{obs}$  and  $\lambda_{em}$  are the observed wavelength and the emitted wavelength, respectively. These are related by  $\lambda_{obs} = \lambda_{em} \cdot (1 + z_{em})$ . The quantity  $\Phi$  is the transmission factor and is given by

$$\begin{aligned} \Phi(\lambda) &\equiv \prod_{i=1}^{N_{\text{abs}}} \exp[-\tau_2(\lambda/(1+z_i))] \\ &= \exp\left[-\sum_{i=1}^{N_{\text{abs}}} \tau_2(\lambda/(1+z_i))\right], \end{aligned} \quad (6.12)$$

where  $\tau$  is given by equation (6.8), and  $z_i$  is the redshift at the epoch of absorption<sup>3</sup>.

Introducing the relation  $f_c(\lambda) = f_{em}[\lambda/(1+z)]/(1+z)$ , it follows from equations (6.2) and (6.12) that

$$1 - D_A(z) = \frac{1}{\Delta\lambda} \int_{\lambda_1(1+z)}^{\lambda_2(1+z)} \Phi(\lambda) d\lambda. \quad (6.13)$$

Note that the right-hand side of this expression is just wavelength-averaged value of  $\Phi$  at redshift  $z$ , and we

<sup>3</sup> The reader shall bear in mind that  $N_{\text{HI}}$ , and  $\sigma_j$  and  $a_j$ —through the dependence on the Doppler parameter—in equation (6.8) are different in general for each absorber. However, we do not write this explicitly by *e.g.* introducing a new subscript in order to avoid confusion.

will denote it by  $\overline{\Phi}_z$ . Note, however, that this quantity still depends on redshift, as indicated by the subscript. Since  $D_A(z)$  and  $\overline{\Phi}_z$  differ only by a constant factor, they may be considered as equivalent with respect to their statistics, which will be discussed in Section 6.5.2.

#### 6.4 OUR $D_A$ MEASUREMENTS

We want to compare the predictions for the evolution of  $D_A$  that result from the models described above to observations. For this purpose, we use previous measurements of  $D_A$  reported in the literature, and we perform ourselves new measurements of this quantity, extending the redshift range of the measurements to  $z_{em} = 5.41$ . Our own measurements are done on QSO spectra from the SDSS Data Release 5.

Our selection procedure of sources suitable for this purpose was as follows: since the wavelength range available from SDSS (DR5) is  $\lambda \in [380, 920]$  nm, and we measure the continuum depression  $D_A$  in the rest-frame wavelength interval  $\lambda \in [105, 117]$  nm, the redshift of our sample is restricted to  $z_{min} \geq 380/105 - 1 = 2.62$ . We choose  $z_{min} = 2.7$  as our lowest redshift in order to avoid the low S/N at the blue end of the spectrograph. A simple query for high- $z$  quasi-stellar objects on the SDSS SkyServer Spectroscopic Query Form with this only restriction returns around 2400 spectra. From this first selection, we rejected those objects for which the redshift was either not measured, the measurement had failed, or the measured photometric and spectroscopic redshifts were inconsistent with each other. We then binned the quasars in redshift intervals of  $\Delta z = 0.1$  and selected for each redshift bin the spectrum with the highest S/N, leaving us with 28 sources, from which we removed three further objects due to low data quality. The resulting sample is listed in Table 6.3.

##### 6.4.1 Continuum fit

The continuum of a quasi-stellar source is often assumed to be of the form  $f_\nu = f_\nu^0 \nu^{-\alpha_\nu}$  (see *e.g.* Steidel & Sargent 1987b, Laor et al. 1997, Zheng et al. 1997), or equivalently,  $f_\lambda = f_\lambda^0 \lambda^{-\alpha_\lambda}$ , where both indices are related by  $\alpha_\lambda = 2 - \alpha_\nu$ , and  $f_\lambda^0 = f_\nu^0 \cdot c^{1-\alpha_\nu}$ , and  $c$  is the speed of light. Despite of the fact that the spectral index  $\alpha_\lambda$  varies over large wavelength ranges (see *e.g.* Neugebauer et al. 1979), Zheng et al. (1997) find by constructing a composite QSO spectrum from 284 HST FOS spectra, that a single power-law describes well the continuum for wavelengths between 105 and

220 nm, but that the continuum steepens significantly for  $\lambda \leq 105$  nm. Telfer et al. (2002) report, using a sample nearly twice as large as the previous group, that the continuum in the extreme ultra-violet region between 50 nm and 120 nm is well described by a single power-law. Analogously, constructing a composite QSO spectrum from a homogeneous sample of over 2200 SDSS QSOs, Vanden Berk et al. (2001) find that the continuum in the rest-frame wavelength range  $\lambda < 500$  nm can be very well modeled by a single power-law either in wavelength or frequency. However, as already mentioned in Section 6.2, other authors prefer to fit continua locally using cubic splines (see *e.g.* Lu et al. 1996) or b-spline functions (see *e.g.* Kirkman et al. 2003, Tytler et al. 2004a,b) searching for regions apparently free of absorption blueward of the Ly $\alpha$  emission line.

Since the Ly $\alpha$  forest region is severely absorbed due to intervening H I systems, especially for high- $z$  QSOs, and a local fit to the continuum in this region is extremely difficult, we choose to estimate the continuum of our selected sources in the Ly $\alpha$  forest region by fitting a power-law to the QSO spectrum redward of the Ly $\alpha$  emission line and extrapolating it for  $\lambda \leq 121.567$  nm in the restframe of the source. As long as the assumption of the underlying power-law holds, this approach has the advantage that the continuum estimate is completely independent of the spectral resolution and S/N in the Ly $\alpha$  forest region. It is true, however, that due to the steepening blueward of  $\lambda = 105$  nm found by Zheng et al. (1997), fitting the continuum of each source redward of Ly $\alpha$  with a single power law and extrapolating it to shorter wavelengths may tend to underestimate the absorption in the Ly $\alpha$  forest region. Nevertheless, and due to a fortunate coincidence, the measurements of  $D_A$  are usually performed precisely for  $\lambda \geq 105$  nm to avoid the wings of the Ly $\beta$  + O VI emission lines. Thus, the assumption of an underlying continuum in the form of a *single* power-law, and consequently the measurements of  $D_A$  assuming such a continuum, are more than validated in the light the results mentioned above.

The power-law that we fit to each spectrum is of the form

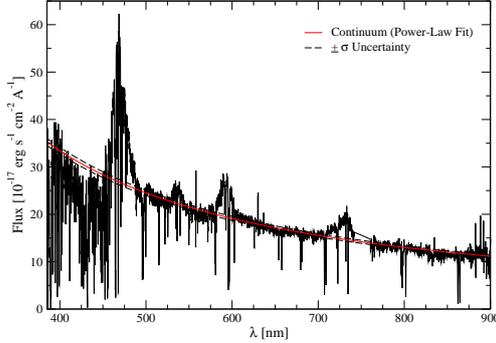
$$f_c(\lambda; \alpha_\lambda) = f_\lambda^0 (\lambda + \lambda_0)^{-\alpha_\lambda} . \quad (6.14)$$

where the flux amplitude  $f_\lambda^0$ , the wavelength off-set  $\lambda_0$ , and the spectral index  $\alpha_\lambda$  are the parameters to be determined. We fit the continuum in the wavelength range  $[130 \cdot (1 + z_{em}), 900]$  nm, in order to avoid the red emission wing of the Ly $\alpha$  line and the red end of the

**Table 6.3:** QSO sample selected from the SDSS DR5. The first four columns include the object designation<sup>a</sup>, the emission redshift as quoted in the SDSS DR5 catalog, the spectral index used for fitting the continuum, and its uncertainty, respectively. The last two columns give the measurement of  $D_A$  and its uncertainty, together with additional error estimate (see text for details.)

Object	$z_{em}$	$\alpha_\lambda$	$\sigma(\alpha_\lambda)$	$D_A$	$\sigma_{err}(D_A)$
SDSS J115538.60+053050.5	2.712	1.213	0.002	$0.390^{+0.010}_{-0.010}$	0.010
SDSS J112107.99+513005.4	2.843	1.256	0.003	$0.281^{+0.015}_{-0.016}$	0.015
SDSS J010619.24+004823.3	2.882	1.170	0.003	$0.294^{+0.018}_{-0.018}$	0.018
SDSS J075618.13+410408.5	2.956	1.320	0.007	$0.329^{+0.038}_{-0.040}$	0.039
SDSS J164219.89+445124.0	3.125	1.550	0.001	$0.349^{+0.005}_{-0.005}$	0.005
SDSS J004054.65-091526.8	3.185	1.278	0.000	$0.459^{+0.000}_{-0.000}$	0.000
SDSS J124306.55+530522.1	3.317	1.255	0.012	$0.426^{+0.057}_{-0.064}$	0.060
SDSS J083122.57+404623.4	3.365	1.230	0.003	$0.261^{+0.017}_{-0.017}$	0.017
SDSS J085343.32+370402.3	3.475	1.115	0.005	$0.527^{+0.019}_{-0.020}$	0.020
SDSS J093523.32+411518.7	3.566	1.407	0.007	$0.396^{+0.036}_{-0.039}$	0.038
SDSS J094349.65+095400.9	3.713	1.199	0.014	$0.529^{+0.057}_{-0.064}$	0.060
SDSS J023137.64-072854.5	3.750	1.269	0.008	$0.462^{+0.037}_{-0.039}$	0.038
SDSS J144717.97+040112.4	3.931	1.303	0.016	$0.525^{+0.062}_{-0.072}$	0.067
SDSS J162331.15+481842.1	3.990	1.113	0.006	$0.494^{+0.022}_{-0.023}$	0.023
SDSS J014049.18-083942.5	4.112	1.181	0.010	$0.488^{+0.041}_{-0.044}$	0.042
SDSS J234150.01+144905.9	4.155	1.094	0.010	$0.558^{+0.035}_{-0.038}$	0.036
SDSS J081240.68+320808.6	4.332	1.074	0.028	$0.517^{+0.100}_{-0.127}$	0.113
SDSS J103601.03+500831.8	4.449	1.149	0.012	$0.612^{+0.039}_{-0.043}$	0.041
SDSS J162626.50+275132.4	4.580	1.187	0.023	$0.681^{+0.057}_{-0.069}$	0.062
SDSS J005006.35+005319.2	4.663	1.204	0.018	$0.693^{+0.043}_{-0.049}$	0.046
SDSS J083914.14+485125.7	4.885	1.350	0.030	$0.737^{+0.057}_{-0.073}$	0.064
SDSS J163950.52+434003.7	4.976	1.321	0.024	$0.694^{+0.056}_{-0.069}$	0.062
SDSS J233446.40-090812.3	5.107	1.332	0.000	$0.735^{+0.000}_{-0.000}$	0.000
SDSS J101447.18+430030.1	5.275	1.220	0.029	$0.758^{+0.058}_{-0.076}$	0.066
SDSS J142123.98+463317.8	5.414	1.321	0.080	$0.825^{+0.087}_{-0.173}$	0.121

<sup>a</sup> The designation of each object meets the IAU nomenclature, as required. For details on the official SDSS designation of an object, please consult [www.sdss.org/dr5/coverage/IAU.html](http://www.sdss.org/dr5/coverage/IAU.html)



**Fig. 6.1:** Spectrum of the QSO SDSS J112107.99+513005.4 at  $z_{em} = 2.843$ . The heavy solid and heavy dashed line indicate the power-law fit to the continuum and its uncertainty, respectively. The corresponding spectral index is  $\alpha_\lambda = 1.256 \pm 0.003$ .

spectrograph, respectively. Note, however, that for redshifts  $z_{em} \gtrsim 5$  the available wavelength range is smaller than 120 nm, and this may introduce a large uncertainty in the fitted continuum. The fit parameters and their uncertainties were obtained with help of the IDL task CURVEFIT. It turns out that the continuum fit is rather insensitive to the uncertainties in the flux amplitude and the wavelength off-set, and extremely sensitive to the uncertainty in the spectral index. Because of this and for simplicity, in our further analysis we neglect the error in the first two parameters and consider only the uncertainty in the spectral index to be of relevance<sup>4</sup>. The spectral index and its uncertainty for each source is listed in Table 6.3. An example of a QSO spectrum and its corresponding fit are shown in Figure 6.1.

#### 6.4.2 Measurement of $D_A$

For each QSO spectrum, we compute the *total* absorption pixel by pixel in the range  $\lambda \in [105, 117]$  nm according to the expression

$$D_A(z_{em}; \alpha_\lambda) \equiv \frac{1}{N_{pix}} \sum_{i=1}^{N_{pix}} \left( 1 - \frac{f_{obs}(\lambda_i)}{f_c(\lambda_i; \alpha_\lambda)} \right), \quad (6.15)$$

<sup>4</sup> The full list of fit parameters for each source and their uncertainty will be available in machine-readable form at <http://astro.physik.uni-goettingen.de/tepper/da/fitparam.txt>

where  $N_{pix}$  is the total number of pixels between  $\lambda_1 = 105 \cdot (1 + z_{em})$  nm,  $\lambda_2 = 117 \cdot (1 + z_{em})$  nm. As a consistency check, we adopt two different methods to estimate the error in our measurements of  $D_A$ . The first method assumes that the  $\pm\sigma$  range for each measurement is given by  $D_A[z_{em}; \alpha_\lambda \pm \sigma(\alpha_\lambda)] - D_A[z_{em}; \alpha_\lambda]$ , respectively, for the corresponding values of  $\sigma(\alpha_\lambda)$  listed in Table 6.3. The second method is based on error propagation, according to which the error  $\sigma(f)$  in the estimate of a quantity  $f(x_i)$ , which depends on  $n$  independent random variables  $\{x_i\}$ , each with an uncertainty  $\sigma(x_i)$ , is given by

$$\sigma^2(f) = \sum_{i=1}^n \sigma^2(x_i) \cdot \left( \frac{\partial f}{\partial x_i} \right)^2. \quad (6.16)$$

In the case of our measurements of  $D_A$ , only the uncertainty in the index  $\alpha_\lambda$  is relevant, and hence equation (6.16) becomes

$$\sigma_{err}(D_A) \equiv \sigma(\alpha_\lambda) \cdot \frac{1}{\Delta\lambda} \int_{\lambda_1}^{\lambda_2} \ln(\lambda + \lambda_0) \frac{f_{obs}(\lambda)}{f_c(\lambda; \alpha_\lambda)} d\lambda, \quad (6.17)$$

where we have used equations (6.14) and (6.15). Our  $D_A$  measurements together with the errors computed according to both methods are listed in Table 6.3. Note that the errors computed using equation (6.17) approximately (in some cases exactly) correspond to the arithmetic mean of the  $\pm\sigma$  uncertainty quoted in the measurements of  $D_A$ , which were computed according to the first method.

Since we do not aim at a high-precision measurement of  $D_A$  (as done for example by Tytler et al. 2004a), we consider our error estimate satisfactory for our purposes. Because of this same reason, we do not correct our measurements for contamination of metal lines. However, this should not introduce a large error, since their contribution is small. For example, Tytler et al. (2004a) find that they contribute by  $2.3 \pm 0.5$  per cent to the total absorption at  $z = 1.9$ . The validity of this assumption will be also shown in the next section by the comparison of  $D_A$ -measurements to the results of our simulations, in which the absorption due to metal lines is not included.

## 6.5 RESULTS & DISCUSSION

### 6.5.1 Observations vs. Models

We compute the evolution of  $D_A$  in the redshift interval  $0.35 < z_{em} < 6.0$  using three different models to account

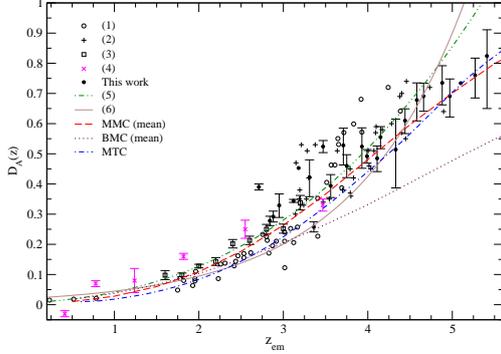
for the intergalactic absorption: the models MMC (with and without Lyman limit systems) and BMC presented in the Section 6.3, which include the effect of different populations of absorbers, *i.e.* Ly $\alpha$  forest clouds and Lyman limit absorbers. For completeness, we include the values of  $D_A$  computed using Meiksin (2006)'s model (see Section 6.1). We refer to this model as MTC. For the MMC (with and without LLS) and BMC models, we simulate an ensemble of lines-of-sight at fixed redshift, and compute for each of them the flux decrement  $D_A$  according to equation (6.13). In this way we get for each given redshift an ensemble of  $D_A$  values for each model, from which we estimate the 50 per cent quantile (median), and the  $\pm 34.13, \pm 43.32, \pm 47.72, \pm 49.38$ , and  $\pm 49.87$  per cent quantiles around the median, which correspond to the  $\pm 1, \pm 1.5, \pm 2, \pm 2.5$  and  $\pm 3\sigma$  ranges. We do not compute the mean and  $\sigma$  ranges in the usual way, since the distribution of  $D_A$  is unknown a priori. However, as will be shown in Section 6.3, the distribution of  $D_A$  is close to a lognormal or even a Gaussian distribution; thus, the identification of mean with median, of  $\pm\sigma$  range with the  $\pm 34.13$  quantile around the median, and so on, is justified. Since the MTC model does not use empirical input distributions of the form of equation (6.4), it is not possible to apply the Monte Carlo technique to this model, and hence we do not further analyse its implications for the evolution of  $D_A$ . We compute for this model a single value for  $D_A$  for each redshift by numerically integrating the transmission function for the corresponding redshift in the rest-frame wavelength range [105, 117] nm.

We compare our simulations to measurements of  $D_A$  accumulated in the literature over the past two decades, done with different methods and approaches. This compilation is by no means intended to be complete. The reason for choosing these measurements is mainly that they were performed in redshift ranges which are more or less mutually exclusive, and which all together cover the range  $0 < z_{em} < 4$ . Thus, these previous measurements together with our own cover a wide redshift range that makes the comparison to models more rigorous. Besides, by using measurements carried on using different methods, we diminish in some way the bias introduced in the  $D_A$  measurements due to the particular method chosen by each group. This makes the comparison of observation to models even more objective. The literature data and our own measurements that extend the redshift range of the observations to  $z_{em} = 5.41$ , is shown together with our model calculations in Figure 6.2 (see cited references for details of the measurements

shown).

As can be seen in Figure 6.2, the predictions for the evolution of  $D_A$  from all three models, MMC, BMC, and MTC, are practically indistinguishable from each other when compared to observations at  $z_{em} \lesssim 3$ . We do not include in this figure the predictions for  $D_A$  based on the MMC model without LLSs, since the difference between this and the full MMC model is negligible. At higher redshifts, the values of  $D_A$  obtained from the MMC and MTC models match the observations pretty well. Yet, it is not possible to discriminate between these models due to the uncertainty and the strong scatter in the observations, especially around  $z_{em} \approx 3.5$  (cf. Section 6.5.3). In contrast to the MMC and the MTC models, the predictions from the BMC model dramatically deviate from the measurements of  $D_A$  at  $z_{em} > 3$ . Since the models differ only in terms of the input distributions and not in terms of the method, it is clear that the input distributions of the BMC model are not quite accurate. Furthermore, and because of the fact that the effect of attenuation is largest at large redshifts, we conclude that the magnitude changes for high- $z$  colors reported by Bershady, Charlton, & Geoffroy (1999) are possibly erroneous, since the measurements of  $D_A$  cannot be recovered when computing its redshift evolution using their model.

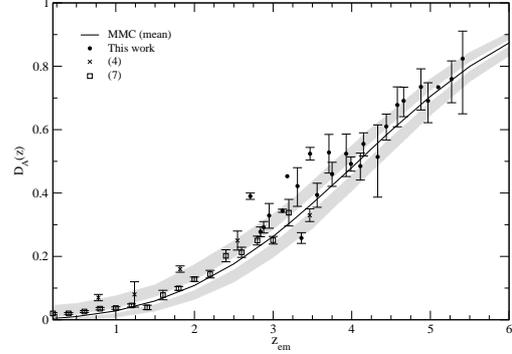
We include in Figure 6.2 the empirical fits of Zhang et al. (1997) of the form  $D_A(z) = D_A^0 e^{\alpha(1+z)}$  with  $D_A^0 = 0.01$  and  $\alpha = 0.75$  and of Kirkman et al. (2005) of the form  $D_A(z) = A(1+z)^\gamma$  with  $A = 0.0062$  and  $\gamma = 2.75$ . These empirical fits match the observations quite well at  $z_{em} \lesssim 4.5$ . However, as stated before, they should be taken with caution, especially at high redshifts. While  $D_A$  asymptotically converges to unity as the redshift approaches the epoch of reionisation,  $z_{reion}$ , these empirical fits do not. In contrast, the predicted evolution of  $D_A$  from all models describe above does satisfy the expected and observed asymptotic behaviour. However, the rate of convergence to this limiting value depends on the particular set of input distributions used: the stronger the number density evolution, the faster the convergence. In other words, different input distribution functions imply different values for  $z_{reion}$ , and this information may be used as a further constraint on the accuracy of a particular set of input distributions. Extending the computations of the evolution of  $D_A$  with the MMC and MTC models to redshifts  $z_{em} > 6.0$ , it turns out that  $D_A$  is almost unity at  $z_{em} \approx 7.0$ . This value is slightly higher than the value



**Fig. 6.2:** Evolution of  $D_A$  computed using the MMC (dashed line), BMC (dotted line), and MTC (dot-dashed line) models, compared to observations performed over the past twenty years by different groups and different methods: (1) Zuo & Lu (1993), (2) Schneider et al. (1991), (3) Kirkman et al. (2005), (4) O'Brien et al. (1988). Note that, despite of the heterogeneity of the approaches to measure  $D_A$ , its evolution as computed using the MMC and MTC models very accurately matches the observations, while they disagree strongly with the BMC model, especially at redshifts  $z_{em} > 3.0$ . Below this redshift, the models are practically indistinguishable from each other. For completeness, we include the empirical fits to the evolution of  $D_A$  from: (5) Zhang et al. (1997), and (6) Kirkman et al. (2005).

of 6.5 quoted by Fan et al. (2002) for the epoch of reionisation.

In Figure 6.3, we show the redshift evolution of  $D_A$  as computed from the MMC model. We show the 50 per cent quantile, *i.e.* the median, and the  $\pm 34.13$ , and  $\pm 49.87$  per cent quantiles, which correspond to the  $\pm 1$ , and  $\pm 3\sigma$  ranges around the median. Again, the difference between the predictions of the MMC with and without LLSs is negligible, and thus we show only the results for the full MMC model. Also shown in this figure are our measurements of  $D_A$  together with the measurements of O'Brien et al. (1988) from Figure 6.3, and the most recent measurements of  $D_A$  available in the literature from Kirkman et al. (2007). We include only these measurements, since they are the only ones that include their uncertainty, and also because these data sets fully cover the redshift range  $0.2 < z_{em} < 5.41$ . As can be seen, within the uncertainty the observations are well matched by the evolution of  $D_A$  predicted by the



**Fig. 6.3:** Evolution of the mean  $D_A$  according to the MMC model (solid line) in the range  $0.2z_{em} < 6$ . Here we show again our measurements of  $D_A$ , together with the most recent measurements from Kirkman et al. (2007) (7). The white and shaded areas around the solid line indicate the intrinsic scatter at the  $\pm 1$  and  $\pm 3\sigma$  level, respectively, due to variation in the absorption from one random line-of-sight to another.

MMC model in the redshift range  $0.2 < z_{em} < 5.41$ , with the exception of some outliers around  $z_{em} \approx 3$ , where the scatter of the observations is large, as also seen in Figure 6.2. As is apparent from Figures 6.2 and 6.3, the largest scatter in the observed values of  $D_A$  at a given redshift cannot be ascribed to the *intrinsic* scatter in the absorption due to variations from one line-of-sight to another, not even at the  $\pm 3\sigma$  level. It may be possible that this scatter is not real, but only an artifact of the method employed to measure  $D_A$ . However, if real, this scatter indicates that the models can only reproduce well the redshift evolution of the mean  $D_A$ , but cannot account for its variation among different lines-of-sight.

None the less, since the MMC reproduces the observations within their corresponding uncertainty very well, we consider it as a fiducial model for the intergalactic absorption and explore in detail its consequences regarding the evolution of  $D_A$ .

The fact that our estimates of  $D_A$  nicely fit into previous measurements and are also well described by the MMC and MTC models, as can be seen in Figures 6.2 and 6.3, supports our assumption that the underlying QSO continuum in the Ly $\alpha$  forest region can be estimated via the extrapolation of the power-law fitted in

the region redward of the Ly $\alpha$  emission line. However, we are aware of the fact that we may have systematically overestimated the continuum level, and that our measurements of  $D_A$  maybe therefore slightly overestimated as well. This would explain why our measurements are higher on average than the median value of  $D_A$  predicted by the MMC model. It is worth mentioning, however, that in contrast to previous models (e.g. Zuo 1993) we explicitly avoid normalising in any way the used distributions to match the observed  $D_A$  at some given redshift, or manipulating the models whatsoever to satisfy any other restriction. We simply take the distributions as reported in the literature, where they were determined directly from line statistics by the authors, and references therein. We mention this in order to emphasize even more the excellent agreement between the MMC model and the observations.

### 6.5.2 Distribution of $D_A$

Judging from the dependence of  $\Phi$  on  $\tau$  and  $N_{abs}$  (see equation 6.12), it is expected that the transmission factor and consequently  $D_A$  are rather complicated random variables. As is well known from statistics, a random variable  $\mathbf{x}$  that can be expressed as the product of a large number of small, statistically independent factors is distributed lognormally, *i.e.* according to the distribution

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^2 x} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2}, \quad (6.18)$$

where  $\mu$  and  $\sigma$  are the mean and the standard deviation of  $\ln \mathbf{x}$ . This expression is equivalent to the statement that a random variable  $\mathbf{x}$  is distributed lognormally, if and only if its logarithm is distributed normally.

The expected value  $\mu'$  and the standard deviation  $\sigma'$  of a lognormal distributed quantity can be expressed in terms of the parameters  $\mu$  and  $\sigma$  as

$$\mu' = e^{\mu + \frac{1}{2}\sigma^2}, \quad (6.19)$$

and

$$\sigma' = \left(e^{\sigma^2} - 1\right)^{1/2} \mu'. \quad (6.20)$$

From the form of equation (6.12) we may suspect that the transmission factor is a lognormally distributed variable, since it can be expressed as the product of a large number of statistically independent factors that take on values in the range [0, 1]. The implications of this statement are profound: if  $\Phi$  is distributed lognormally, so does  $\overline{\Phi}_z$  and consequently  $D_A(z)$ . Furthermore, due to

the property of the lognormal distribution stated above, the effective optical depth of the Ly $\alpha$  absorption, usually defined as  $\tau_{\text{eff}} \equiv -\ln(1 - D_A)$  (see e.g. Kim et al. 2001) should obey a Gaussian distribution. This result follows independently from the fact that the total optical depth can be expressed as the sum of the independent contribution of each system, as indicated in equation 6.12). Thus, for a sufficiently large number of absorbers  $N_{abs}$ , and if the optical depth for each absorber has the same mean value  $\langle\tau\rangle$  and dispersion  $\sigma(\tau)$  at each wavelength,  $\tau_{\text{eff}}$  should obey a Gaussian distribution at each redshift, centered at  $N_{abs}\langle\tau\rangle$  and dispersion  $\sqrt{N_{abs}}\sigma(\tau)$ . The most astonishing fact is that these statements are completely independent of the form of evolution of the intergalactic neutral hydrogen, as long as the transmission factor can be expressed in the form of equation (6.12). Thus, it should be a fact of nature that the distribution of the absorbed flux at any given wavelength should obey a lognormal distribution.

In order to test whether the values of  $D_A$  obey a lognormal distribution, we compute the mean and standard deviation of  $\ln D_A$  for the ensemble of  $4 \cdot 10^3$  lines-of-sight at a given redshift according to the equations

$$\mu(\ln D_A) \equiv \frac{1}{N_{\text{LOS}}} \sum_{i=1}^{N_{\text{LOS}}} \ln D_A^i, \quad (6.21)$$

and

$$\sigma^2(\ln D_A) \equiv \frac{1}{N_{\text{LOS}} - 1} \sum_{i=1}^{N_{\text{LOS}}} \left(\ln D_A^i - \langle\ln D_A\rangle\right)^2, \quad (6.22)$$

With these parameters, we generate normally distributed random numbers using the Monte Carlo technique, and compare them to the  $\ln D_A$  values from our simulations at each given redshift, in order to determine whether the latter are normal distributed. Recall that  $\ln D_A$  is distributed normally if and only if  $D_A$  obeys a lognormal probability distribution. As a first approach, we choose to compare both data sets graphically using Quantile–quantile plots (QQ plots for short), and to this end we compute for each set of values the 50 per cent quantile (median), and the  $\pm 34.13, \pm 43.32, \pm 47.72, \pm 49.38,$  and  $\pm 49.87$  per cent quantiles around the median, as done in previous sections. With this graphical aid it is possible to tell qualitatively if the data sets compared to each other are drawn from the same parent distribution. As a thumb rule, one can say that the closer the quantiles lie to the identity line in a QQ plot, the larger the probability that

both data sets obey the same distribution. In order to evaluate this statement quantitatively, we define the following measure

$$\delta^2 \equiv \sum_{i=1}^{N_q} \left( \frac{Q_i - q_i}{Q_i} \right)^2, \quad (6.23)$$

where the  $Q$ 's are the quantiles determined from random numbers distributed normally, the  $q$ 's are the corresponding quantiles of the  $\ln D_A$  values at a given redshift that result from our simulations, and  $N_q = 11$  is the number of quantiles. By definition,  $\delta$  explicitly gives the absolute deviation of one data set with respect to the other, and hence quantifies the departure of the assumed distribution. Indeed, the smaller the value of  $\delta$ , the larger the probability that both data sets belong to the same distribution. In Figure 6.4 we show the QQ-plots at each redshift and together with the corresponding value of  $\delta$ . As can be judged by visual inspection, the departure of the distribution of  $\ln D_A$  from a Gaussian distribution is negligible at all redshifts. Furthermore, the quoted values of  $\delta$  seem also vanishingly small. However, since the distribution of the random variable  $\delta$  itself is unknown, we try to determine the significance of the results by generating two sets of normally distributed random numbers with mean and standard deviations distributed uniformly, and comparing them to the values found for our simulated data. We do this for  $1.2 \cdot 10^6$  pairs of sets, and determine from these what is the fraction of realisations with a value of  $\delta$  smaller than a given value. We expect this procedure to quantify the probability that the values of  $\delta$  quoted in Figure 6.4 happen by chance. According to these estimates, there is a probability of  $\{1.4 \cdot 10^{-2}, 3.1 \cdot 10^{-3}, 1.9 \cdot 10^{-4}, 2.3 \cdot 10^{-5}\}$  that a value  $\delta \leq \{0.50, 0.25, 0.10, 0.05\}$  will be produced by chance, respectively. In Figure 6.6 we show the values for  $\delta$  as a function of redshift with the MMC model. Note that for the whole redshift range shown,  $\delta \lesssim 0.5$ , and in particular  $\delta \lesssim 0.1$  for  $z_{em} > 1.5$ . Based on these results, we may conclude that the distribution of  $\ln D_A$  at a given redshift is Gaussian with a confidence level of 99 per cent, which implies that  $D_A$  obeys a lognormal distribution with the same confidence.

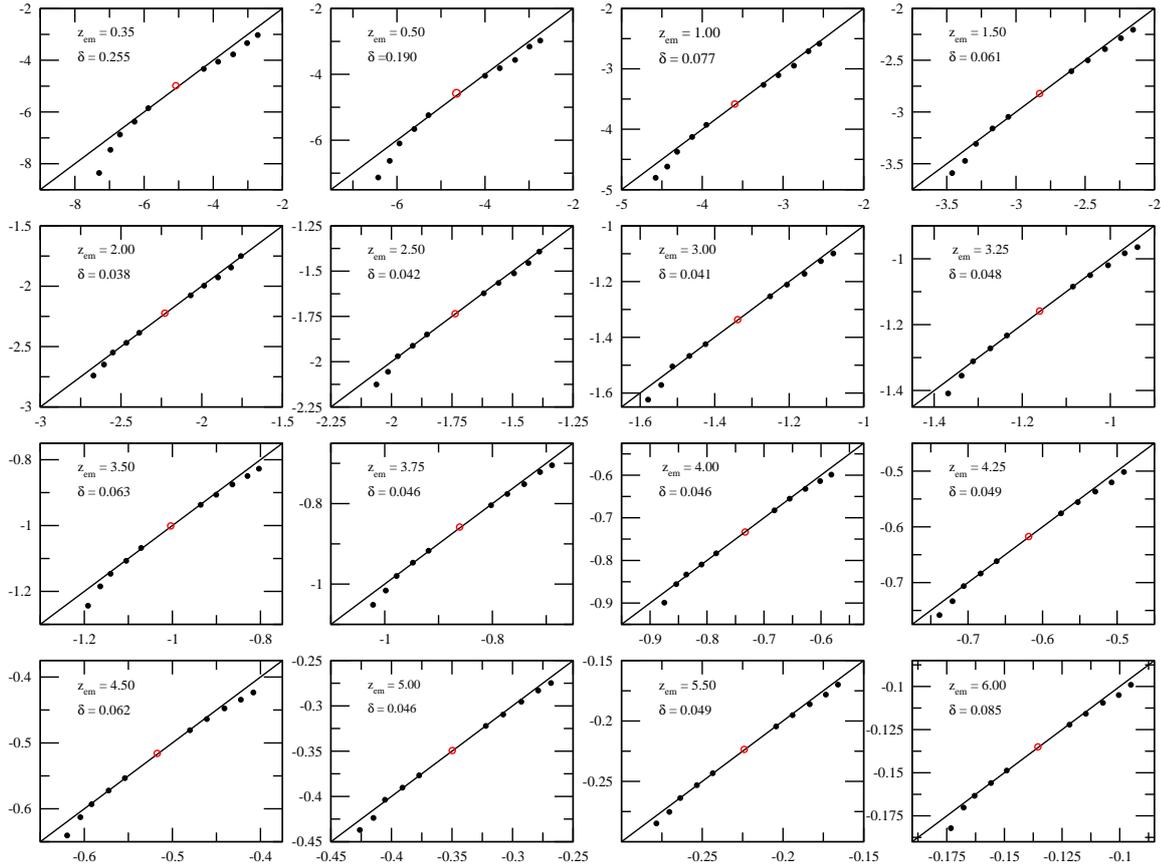
It follows from the Central Limit Theorem that the larger the number of absorbers  $N_{abs}$ , the closer the approach of the distribution of  $\Phi$  to a lognormal distribution. Since  $N_{abs}$  increases with redshift, it should be expected that the accuracy with which the distribution of  $\Phi$  approaches a lognormal distribution also increases with redshift. Furthermore, if the integral in equation

(6.13) is approximated by a sum of the form

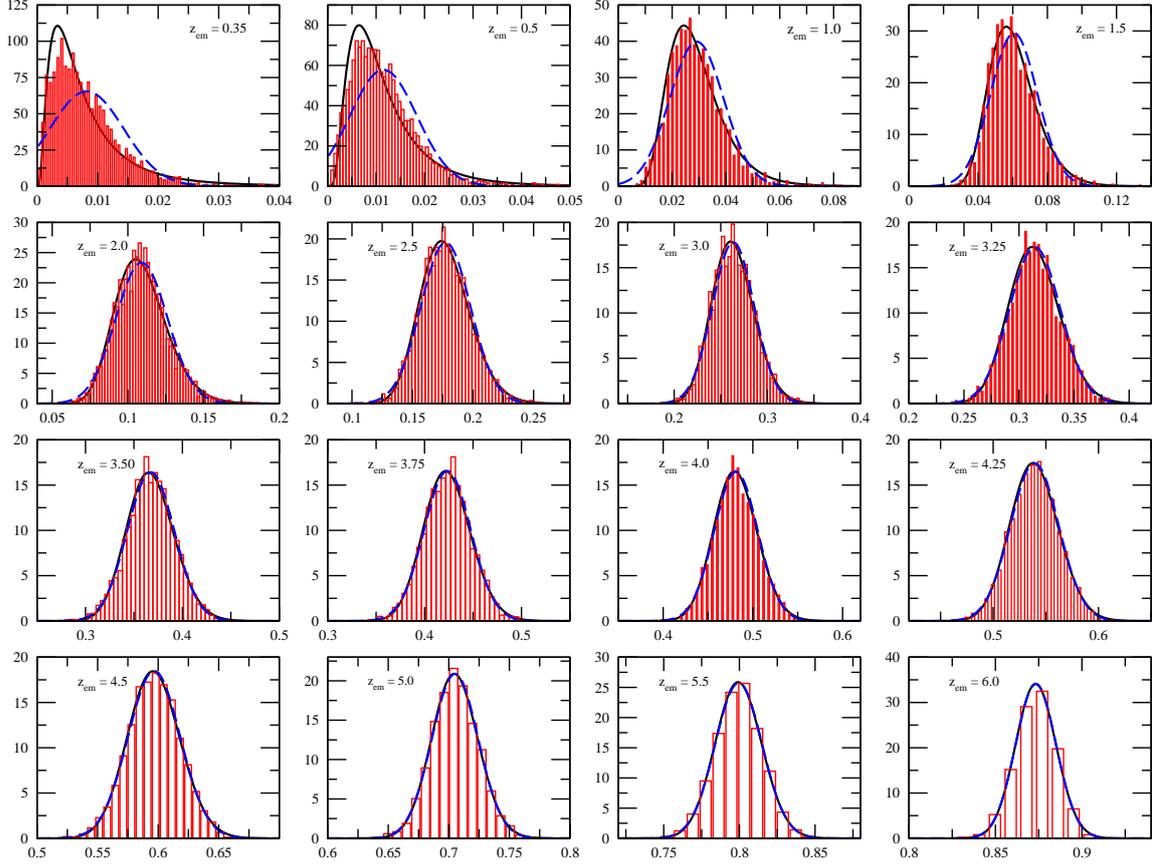
$$\bar{\Phi}_z \approx \frac{1}{N} \sum_{i=1}^N \Phi_i \Delta \lambda_i,$$

where  $N$  is the number of pixels, it becomes evident that for sufficiently large  $N$ , and assuming that  $\Phi$  has the same mean value  $\langle \Phi \rangle$  and dispersion  $\sigma(\Phi)$  at each pixel, the distribution of  $\bar{\Phi}_z$  should approach a Gaussian distribution with mean  $\langle \Phi \rangle$  and dispersion  $\sigma(\Phi)/N$ . Since the width of a given restframe wavelength range, *i.e.* the number of pixels also increases with redshift as  $(1+z)$ , it is expected that the approximation of the distribution of  $\bar{\Phi}_z$  and hence of  $D_A$  to a Gaussian becomes better with increasing redshift. We thus have the superposition of two effects: On the one hand, the distribution of  $\Phi$  at a fixed wavelength approaches a lognormal distribution at all redshifts, with the accuracy increasing with redshift. On the other hand, the distribution of  $\bar{\Phi}_z$  and  $D_A(z)$  approaches a Gaussian distribution with increasing redshift. The net result should be that  $\bar{\Phi}_z$  and  $D_A$  are distributed lognormally at low redshifts, and that their distribution approaches a Gaussian for higher redshifts. Since  $D_A$  asymptotically converges to unity for very high redshifts, the Gaussian distribution at these redshifts is expected to be highly peaked around its mean value. This is naturally given by the fact that the dispersion of  $\bar{\Phi}_z$  scales like  $\sigma(\Phi)/N$  around  $\langle \Phi \rangle$ , as stated above.

We compute the theoretical lognormal probability distribution as given by equation (6.18) with the mean and standard deviation of  $\ln D_A$  computed using equations (6.21) and (6.22). We do the same for a Gaussian distribution, using the same equations as before, but replacing  $\ln D_A$  by  $D_A$ , and compare both these theoretical distributions to the distribution of the  $D_A$  values at a given redshift that result from our simulations. This comparison is shown in Figure 6.5. Note the excellent agreement at all redshifts between the lognormal probability distribution and the distribution of the  $D_A$  values resulting from our simulations. We want to emphasize that the theoretical curves shown are not fits to the binned data, but are computed using only the mean and standard deviation of the unbinned data. The data were binned only for display purposes. At lower redshift, the agreement at the lower cut-off, *i.e.* at  $D_A = 0$  is worth mentioning. It is remarkable that this cut-off, which is physically given by the fact that  $D_A$  cannot take on values smaller than zero, arises in a natural way due solely to the fact that  $D_A$  is distributed log-



**Fig. 6.4:** Quantile-quantile Plot for the values of  $\ln D_A$  computed from the MMC model. Shown are the quantiles corresponding to the  $\pm 1, \pm 1.5, \pm 2, \pm 2.5$  and  $\pm 3 \sigma$  ranges for the logarithm of our simulated values of  $D_A$  and for a realisation of normal deviated random numbers with the same parameters, *i.e.* mean and standard deviation, as the simulated data. An open circle denotes the median, and its nearest neighbouring points the  $\pm 1 \sigma$  range, the next nearest the  $\pm 1.5 \sigma$  range, and so on. The solid line represents the ideal case where both data sets are drawn from the same parent distribution. The deviation from this ideal situation is quantified by the parameter  $\delta$ , which takes on values between 0 (both data sets belong to the same distribution with unit probability), and  $\infty$  (both data sets *do not* belong to the same distribution with unit probability).



**Fig. 6.5:** Distribution of the  $D_A$  values at a given redshift computed from simulations based on the MMC model (histogram). The distribution is normalised to unit area. Shown are also a theoretical lognormal distribution (solid curve) computed according to equations (6.18), (6.21), and (6.22), and a Gaussian distribution (dashed curve) computed using the same equations as before but replacing  $\ln D_A$  by  $D_A$ . The y-axis indicates the probability of the corresponding  $D_A$  value on the x-axis. The bin size used to compute the histogram at each redshift has been arbitrarily chosen to be given by  $\max\{D_A(z)\}/100$ , and is hence different at each redshift. However, the binning of the data is intended only for display purposes. The binned data have not been used in any form to estimate the parameters of the corresponding probability distribution. We want to emphasize that the theoretical curves shown are not fits to the binned data, but are computed using only the mean and standard deviation of the unbinned data. Note the excellent agreement between the data and the theoretical lognormal distribution, especially with respect to the skewness and the cut-off at  $D_A = 0$ . The distribution of the data is not well described by a Gaussian distribution at the lowest redshifts shown. However, the approach of the data distribution to a Gaussian distribution increases with increasing redshift (see text for details).

normally. Note, in contrast, that a Gaussian distribution does not satisfactorily describes the distribution of the data at these low redshifts. Nevertheless, the description of the data by Gaussian distribution becomes better with increasing redshift, as previously stated, and the corresponding Gaussian distribution becomes narrower at every increasing redshift. Note also that the theoretical lognormal and Gaussian distribution become visually indistinguishable from each other at high redshifts. In order to assess quantitatively the differences between these distributions with respect to the distribution of the data, we compute again the value of  $\delta$  at each redshift, assuming now that the values of  $D_A$  values obtained from the MMC model are drawn from a Gaussian parent distribution. We compare these values to the corresponding values computed before for an assumed lognormal parent distribution. This comparison is shown in Figure 6.6. As can be seen, the values of  $\delta$  for both distributions are low at all redshifts, and the difference between them at  $z_{em} \gtrsim 2$  is negligible small. This explains why the lognormal and Gaussian distribution functions shown in Figure 6.5 are practically indistinguishable from each other. However, note that the values of  $\delta$  for a assumed Gaussian parent distribution become vanishingly small with increasing redshift, and thus the probability that the data are drawn from this distribution increases with redshift, eventually becoming larger than the corresponding probability for the lognormal distribution. Besides, it can be see also that the  $\delta$  values for a assumed lognormal distribution rise again towards high redshifts, implying that the distribution of the data at these redshifts is no longer well described by a lognormal distribution. This confirms our statement made above, that  $D_A$  is expected to be distributed lognormally at low redshifts and normally at higher redshifts. The redshift at which the transition from a lognormal to a Gaussian distribution takes place may depend on the particular set of input distributions used. For redshifts where the distribution of  $D_A$  is well approximated by a lognormal distribution, It is straightforward to see that the optical depth must be distributed normally. This result is not at odds with the results from *e.g.* Madau (1995, Figure 1), Meiksin & White (2004, Figure B1), or Bernardi et al. (2003, Appendix C).

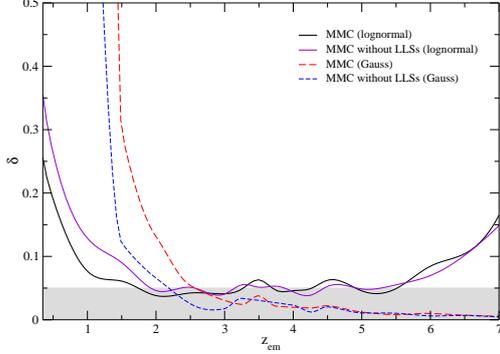
#### The Effect of the Lyman limit systems (LLSs)

It is usually assumed that the presence of the high-density Lyman limit systems tends to skew the distributions of the opacity and hence of the absorption towards

larger values, making values smaller than the mean more probable (see *e.g.* Madau 1995, Meiksin & White 2004). In order to quantify the effect that the LLSs have on the distribution of  $D_A$ , we compare, following the analysis of Section 6.5.2, the results of the simulations for the MMC model with and without the optically thick LLSs. We find that  $D_A$  is lognormally distributed as well with a high confidence for the case where the LLSs are excluded, as can in Figure 6.6, where the values for  $\delta$  computed at each given redshift are shown. Note that the values of  $\delta$  for the MMC model without LLSs are larger at redshifts  $z_{em} \lesssim 3$  than those for the full MMC model. This is due to the fact that the absence of the optically thick LLSs enhances the probability of the attenuation factor  $\exp(-\tau)$  to be closer to unity at a given wavelength along a random line-of-sight. This means effectively that the number of factors in equation (6.12) are reduced. Hence, when the LLSs are absent, the approach to a lognormal distribution should be worse with respect to the case where the LLSs are included. Furthermore, this effect should be enhanced towards lower redshifts, for which the number of factors, *i.e.* of absorbers decreases as  $\propto (1 + z_{em})^\gamma$ . Note that the trend is the opposite in the case of a assumed Gaussian parent distribution.

We have found when computing the evolution of  $D_A$  with the MMC model with and without LLSs that the predictions for the evolution of  $D_A$  with redshift for the full MMC model are practically indistinguishable from the results of the MMC model without LLSs, meaning that the effect of these systems on the total absorption is negligible. Furthermore, it turns out that the distribution of  $D_A$  values is skewed, *i.e.* it is lognormal irrespective of the presence of LLSs. All these results point to the fact that the LLS have a negligible impact on the evolution of the intergalactic absorption, as long as the input distributions used here correctly describe the number density evolution of the absorbers. This confirms the results from Desjacques et al. (2007), and the picture shall not be fundamentally different, if higher column density systems such as damped Ly $\alpha$  systems were present, as pointed out by McDonald et al. (2005).

These results certainly demand an explanation. We think that, as long as the distribution functions realistically describe the evolution of the Ly $\alpha$  absorbers, the LLSs *cannot* have a great impact neither on the absorption, nor on its statistics because of the following reasons: 1) They are scarce, and even more compared to the thinner Ly $\alpha$  forest systems (*e.g.* 5 LLSs on average along a random line-of-sight compared to approx-



**Fig. 6.6:** Values of  $\delta$  for the MMC model assuming a lognormal (heavy solid line) and Gaussian (heavy dashed line) parent distribution. The shaded area indicates the range  $\delta \leq 0.05$ , that corresponds to a probability of  $1.9 \cdot 10^{-5}$  that the data are not drawn from the same parent distribution. Note how the assumption of a Gaussian parent distribution improves with redshift, but that a lognormal distribution is a better assumption for low redshifts. The light solid and dashed lines correspond to the MMC model without Lyman limit systems (see Section 6.5.2). Note that the approximation to a lognormal distribution, is better when these systems are included (see text for discussion). The fluctuations seen with respect to a perfect smooth curve in each case are due to the random nature of the process at each redshift, and are not significant.

imately 1500 Ly $\alpha$  forest absorbers out to  $z = 3.0$ , according to the MMC model); 2) Their contribution to the absorption is due solely to their Ly $\alpha$  absorption line (not the Lyman break), and the equivalent width of the systems with column densities  $N_{\text{HI}} = 10^{17.21-18} \text{ cm}^{-2}$ , as can be seen from a curve-of-growth (Figure 6.7), is not too different from that of the systems with column densities  $N_{\text{HI}} = 10^{14-15} \text{ cm}^{-2}$ , which are by far more numerous. For example, the input distributions of the MMC model predict that the ratio of the average number of the optically thin Ly $\alpha$  forest absorbers to the number of optically thick Lyman limit systems at a redshift  $z_{em} = 4.0$  amounts to 477. If we suppose that *all* Ly $\alpha$  forest and Lyman limit systems are absorbers with the largest possible column density for their respective column density ranges, *i.e.*  $N_{\text{HI}} \approx 10^{17}$  and  $N_{\text{HI}} \approx 10^{20} \text{ cm}^{-2}$ , then the ratio of the Ly $\alpha$  equivalent width of the thin systems to that of the thick systems for a mean Doppler parameter of  $36 \text{ km s}^{-1}$  is ap-

proximately 0.12. For the case that we take the lowest column density and more numerous systems to be representative of their respective populations, the ratio of their equivalent widths is approximately 0.006. Even in this case, the *total* equivalent width of the optically thin Ly $\alpha$  forest systems dominates over that of the optically thick LLSs by a factor of  $477 \cdot 0.006 = 2.89$ , at  $z_{em} = 4.0$ . In order to see the relative contribution of each population at a given redshift more precisely, we weight the equivalent width  $W(N, b)$  with the column density distribution, and compute the ratio

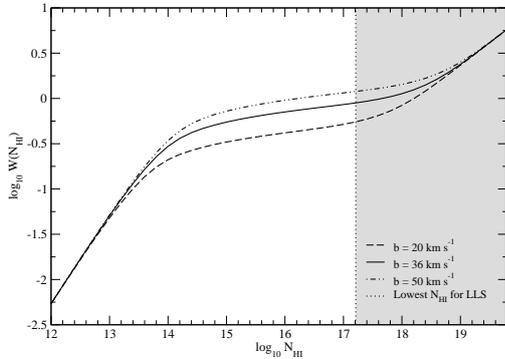
$$\varrho(z_{em}; b) \equiv \frac{\langle W \rangle_{\text{Ly}\alpha}}{\langle W \rangle_{\text{LLS}}}, \quad (6.24)$$

where

$$\langle W \rangle_i \equiv \int_0^{z_{em}} \int_{N_{min}}^{N_{max}} (1+z) \cdot W_0(N, b) f_i(N, z) dN dz, \quad (6.25)$$

Here,  $f_i$  is the distribution function of population  $i \in \{\text{Ly}\alpha, \text{LLS}\}$ ,  $W_0(N, b)$  is the rest equivalent width of a Ly $\alpha$  absorption line for a column density  $N$  and Doppler parameter  $b$ , and  $N_{min}$  and  $N_{max}$  are the column density limits that define each population, respectively. For a reasonable value for the Doppler parameter of  $36 \text{ km s}^{-1}$ , and the input distributions of the MMC model, we find that  $\varrho \gg 1$  at all redshifts. The result is qualitatively the same for the BMC model. It can be concluded from this that the Ly $\alpha$  forest systems dominate the absorption over the optically thicker Lyman limit systems at all epochs. This explains in the first place why the difference between the predictions for  $D_A$  from the MMC model with and without LLSs, is vanishingly small. Also, it is consistent with the fact that the distribution of  $D_A$  should not be far from lognormal or Gaussian, with or without LLSs, since this only depends on the fact that the absorption factor be expressed in the form of equation (6.12), and this is truly independent of the form of the input distributions, as stated previously. Finally, it also explains why the behaviour of  $\sigma(D_A)$ , as shown in the next section (cf. Figure 6.8), is qualitatively the same irrespective of the presence of these systems.

We thus conclude that the distribution of  $D_A$ , *i.e.* the flux field for a given redshift as predicted by the MMC model, is lognormally distributed with a high confidence at redshifts  $z_{em} \lesssim 6$ , and consequently, that  $\ln D_A$  and hence the optical depth of H $\alpha$  obey a Gaussian distribution, in disagreement with the results of previous analyses (*e.g.* Madau 1995, Bernardi et al. 2003, Tytler et al. 2004a, Meiksin & White 2004).



**Fig. 6.7:** Curve of growth of the Ly $\alpha$  absorption line, for three typical values of the Doppler parameter. The shaded region corresponds to the column densities characteristic to Lyman limit systems.

### 6.5.3 Scatter in $D_A$

We expect the intrinsic scatter in the absorption due to cosmic variance to be strongest at lines-of-sight of middle length. At low redshifts, both the number of thin Ly $\alpha$  forest clouds and thick Lyman limit systems is small, and the addition of a few more does not change dramatically the amount of absorption. However, the number of Ly $\alpha$  forest clouds increases rapidly with  $z$ , and thus the probability of encountering more or less systems than average increases as well. Correspondingly, the absorption increases and so does its scatter. At even higher redshifts, the number of Ly $\alpha$  forest systems increases so dramatically and the absorption is so severe that the addition of more systems does not make any difference neither to the absorption nor to the scatter. Thus, we should expect the stochastic effect, *i.e.* the scatter in absorption, to peak at some intermediate redshift  $z_{int}$ .

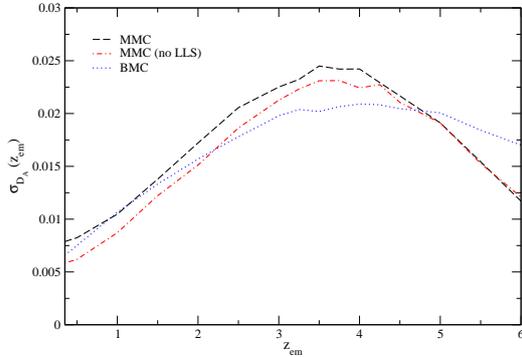
We compute the intrinsic scatter at the  $\sigma$  level for the values of  $D_A$  at each given redshift obtained from our simulations using equations (6.19), (6.20), (6.21), and (6.22). The result is shown in Figure 6.8, where we show the evolution of the scatter in  $D_A$  with redshift for the MMC model with and without LLSs. We also include for completeness the result from the BCG model. It can be seen that, irrespective of the model, the scatter peaks at a intermediate redshift between  $z_{em} \approx 3.5$

and  $z_{em} \approx 4.0$ . Note that the peak is significant, since it represents an increase in the scatter of 2.5 times with respect to its value at  $z_{em} \approx 1$ . It is interesting that this result had also been found by Zuo (1993, Figure 2), who using a semi-analytic approach and different input distributions, reported that  $\sigma(D_A)$  is largest at redshifts near 3.7. Thus, the qualitative behaviour of the intrinsic scatter of  $D_A$  shown in Figure 6.8 may be an unavoidable feature of this observable, which could explain at least in part the large scatter in the measurements of  $D_A$  seen in the same redshift interval (cf. Figure 6.2). As stated above, however, the large scatter in the observations at these redshifts cannot be accounted for by the models, not even at the  $3\sigma$  level.

When comparing models that only differ by the presence of the optically thick Lyman limit systems, we find that the scatter in  $D_A$  is larger at any given redshift when the LLSs are present. None the less, the absolute value of the scatter does not differ significantly between the situation where these systems are present and where they are absent. Thus, the contingent presence of a few less or more thicker systems does introduce a variation in the absorption from one line-of-sight to another, even though the absorption itself does not vary in a significant amount. In other words, the net effect of the LLSs is to enlarge the intrinsic scatter in the absorption. However, note that for redshifts  $z_{em} \gtrsim 4.0$ , the amount of scatter at the  $1\sigma$  level becomes indistinguishable between both models, owing to the fact that the thinner Ly $\alpha$  forest systems are overwhelmingly numerous at these redshifts.

The results above support our previous conclusion that the optically thick LLSs cannot have a great impact on the absorption and on its scatter. The fact that LLSs (and maybe DLAs) do not have a great impact on the scatter in absorption is even more apparent when assessing their effect on broad-band colors. This is mainly because their Lyman break is *always* blueward of the Lyman break of the source, and their Ly $\alpha$  lines do not contribute significantly to the absorption, especially when integrating the flux over a large wavelength range characteristic of broad-band filters. We leave this analysis to a forthcoming paper.

We want to highlight two curiosities: First, by taking a close look at Figures 6.3 and 6.8 it becomes apparent that the maximum in the evolution of  $\sigma(D_A)$  roughly coincides with the point of inflection of the curve that describes the redshift evolution of  $D_A$ . Mathematically, this implies that the intrinsic scatter of  $D_A$  is propor-



**Fig. 6.8:** Evolution of the intrinsic scatter of  $D_A$  due to the stochastic nature of the absorption in the intergalactic medium for the two competing models MMC (dashed line) and BMC (dotted line). The dot-dashed line corresponds to the MMC without Lyman limit systems (see Section 6.5.2 for details). Note that the behaviour of  $\sigma(D_A)$  is qualitatively the same for all three models. As expected, the amount of scatter for the MMC model is less in the case where the Lyman limit systems are absent.

tional to the rate of change of  $D_A$  with redshift, *i.e.*

$$\sigma(D_A) \propto \frac{\partial D_A}{\partial z}. \quad (6.26)$$

Indeed, if one computes numerically the derivative of  $D_A$  with respect to  $z$ , it turns out that it qualitatively matches the evolution of  $\sigma(D_A)$ , up to a scale transformation. A possible interpretation may be that a stronger evolution of  $D_A$ , *i.e.* a stronger change with  $z$  (where the rate of change is given by the derivative) implies a rapid change in the evolution of the absorbers. This implies a larger variation in the absorption from line-of-sight to line-of-sight, and thus a larger value of  $\sigma(D_A)$ .

Second, it is interesting to note that the redshift at which the intrinsic scatter in evolution of  $D_A$  peaks roughly coincides with the redshift at which the observations show a strong scatter. *i.e.* around  $z_{em} \approx 3.5$ . This roughly matches the redshift at which [Bernardi et al. \(2003\)](#) reported a particular feature in the evolution of the Ly $\alpha$  optical depth, which has been interpreted by the authors as a signature of the reionisation of He II. If these is a mere coincidence or if it has a more profound meaning requires further detailed analysis, which is however beyond of the scope of the present

work.

## 6.6 SUMMARY & CONCLUSIONS

1. We measured the cosmic flux decrement  $D_A$  and its uncertainty for 25 QSOs of the SDSS DR5 catalog in the redshift range  $2.71 \leq z_{em} \leq 5.41$ .
2. We modeled the redshift evolution of  $D_A$  in a Monte Carlo fashion, using two different sets of input distributions for the absorber properties, and found that the prediction of the MMC model reproduce well the observations in the range  $0.2 < z_{em} < 5.41$ , in contrast to the BMC model. We conclude from this that the underlying input distributions of the BMC model do not quite describe the evolution of the Ly $\alpha$  absorption on the IGM, and that previous estimates of the impact of the intergalactic attenuation on the photometric properties of high-redshift galaxies using this model are inaccurate. Incidentally, by comparing the simulations to the data, we show the power of the relatively simple approach used here to model the effect of intergalactic absorption, as compared to models based on hydrodynamical simulations which are by far more complex and yield similar results.
3. We found from our simulations that the distribution of the  $D_A$  values at a given redshift is well described by a lognormal distribution at low redshifts and even better by a Gaussian distribution at high redshifts, in agreement with the fact that the absorption consists of the product of statistically independent factors. This result implies that at redshifts where  $D_A$  is distributed lognormally, the effective optical depth of the intergalactic HI obeys a normal distribution, contrary to the results of previous studies. However, a larger number of accurate measurements are needed to determine whether  $D_A$  is truly distributed (log-)normally. We leave as an open question whether the result that  $D_A$  is distributed lognormally is linked in some way to the assumption that the primordial density fluctuations that gave origin to the structure formation in the Universe, and that ultimately gave rise to the Ly $\alpha$  absorbers, obey a lognormal distribution as well, as proposed by [Bi et al. \(1992\)](#).

4. We conclude that the optically thick Lyman limit systems have a negligible effect on the absorption, even though they do increase its intrinsic scatter. As expected on a theoretical basis, the presence of these systems does not significantly affect the distribution of  $D_A$ . These results are subject to the assumption that the evolution of these systems is given by the input distributions used.

## 6.7 ACKNOWLEDGMENTS

We are grateful to Avery Meiksin for computing a special set of tables of his transmission functions. We also want to thank Nico Bissantz for useful suggestions about the statistical treatment of the data. TTG acknowledges financial support from the Mexican Council for Science and Technology (CONACyT).

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England<sup>5</sup>.

The SDSS is managed by the Astrophysical Research Consortium for the participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

---

<sup>5</sup> The SDSS Web Site is <http://www.sdss.org/>.



---

# Stochastic Absorption of the Light of Background Sources due to Intergalactic Neutral Hydrogen

## II. The impact on the photometric redshifts of galaxies

---

### 7.1 INTRODUCTION

The *mean* photometric properties of field galaxies in deep surveys are nowadays well described by evolutionary synthesis models such as GISEL (Bruzual A. & Charlot 1993), PEGASE, (Fioc & Rocca-Volmerange 1997), and GALEV (Bicker et al. 2004). Nevertheless, a scatter in the observed luminosities and colors of galaxies is present which cannot be accounted for by any of these models. The origin of this observed scatter may be two-fold: On the one hand, there are intrinsic properties of the galaxies like the dust content or strong on-going starbursts which cause them to have colors and luminosities intrinsically different from those of their respective "average" counterparts or templates (E, S0, Sa, . . . , Sd). The correction for the variations in colors and luminosities due to absorption by dust has not yet consistently been included into evolutionary synthesis models, although an important step towards this goal has been achieved by Möller et al. (1999). Furthermore, Fritze-v. Alvensleben & Bicker (2006) found that some of the extremely red galaxies observed in the HDF-N with  $(m_{450} - m_{606}) \geq 1.5$  (Sawicki et al. 1997) can be explained as dust-free(-strong-)post-starburst phases of normal galaxies which experienced a major event of Star Formation (SF) in the past.

On the other hand, the light of a galaxy is partially absorbed by the intervening neutral hydrogen (HI) in the intergalactic medium (IGM) in form of discrete Ly $\alpha$  absorbers. Since the absorption primarily affects the rest-frame UV, the mean net effect of the intergalactic attenuation is to redden and dimm the spectrum. However, since the distribution of the Ly $\alpha$  absorbers along the line of sight is stochastic in nature, there is a chance that the line-of-sight to galaxies which are either bluer or redder than the average population that is well described by models accounting for an average amount of attenuation.

Due to the random distribution of neutral hydrogen in the IGM, the population of Ly $\alpha$  absorbers characterized by their redshift  $z_{abs}$ , column density  $N_{HI}$  and Doppler parameter  $b$  is unique for each line of sight. Hence, in order to correctly account for the absorption of light caused by this intervening absorbing HI systems, a detailed knowledge of their properties (number density, temperature, etc.) and evolution is required. It has been found from statistical analyses of quasar absorption lines (see *e.g.* Hu et al. 1995, Kim et al. 1997, 2002a, and references therein) that, when averaged over a large number of lines-of-sight, the number density evolution of Ly $\alpha$  absorbers is well described by a redshift- and column density-dependent power law of

the form  $dN \propto N_{\text{HI}}^{-\beta} (1+z)^\gamma$  (Peterson 1978, Carswell et al. 1984, Atwood et al. 1985, Murdoch et al. 1986, Carswell & Rees 1987).

The impact of attenuation of light by intervening HI systems on the spectra of bright background sources has until now been studied by Møller & Jakobsen (1990), Madau (1995), Bershadsky et al. (1999), and more recently by Meiksin (2006). Using different sets of parameters for this power law derived from observations, these studies have, with different purposes and approaches, made efforts to account for the *mean* attenuation due to intervening Ly $\alpha$  absorbing systems of light emitted from distant sources like QSOs and galaxies (Møller & Jakobsen 1990, Madau 1995) and for its impact on the distribution of galaxy colors (Madau 1995, Bershadsky et al. 1999, Meiksin 2006). It is our aim to complement these previous studies and to explore to which extent the observed scatter in photometric properties of field galaxies can be due to the stochastic nature of the attenuation. In this first approach, we focus on this effect only, neglecting the effects of dust and ongoing starbursts, as well as the absorption within the galaxy.

Here we present a further, improved effort to quantitatively account for the stochastic effect of the Ly $\alpha$  absorbers on the photometric properties of galaxies. Following Møller & Jakobsen (1990) and Bershadsky et al. (1999), we apply Monte Carlo techniques to simulate the effects of hundreds to thousands of absorbers along a random line of sight and model the mean attenuation of light and any desired  $\sigma$  level around this mean as a function of redshift. The advantage of our approach is that it includes the corrections concerning the calculation of the  $\sigma$  scatter pointed out by Bershadsky et al. (1999). We apply this model to address the question of how photometric redshift estimates are affected by the stochastic attenuation.

## 7.2 MODEL DESCRIPTION & INPUT PHYSICS

In Tepper-García & Fritze (2007, from now on Paper II) we presented a model of the absorption of the light from background sources due to the presence of randomly distributed intergalactic neutral hydrogen along the line-of-sight. We compute of an ensemble of lines-of-sight, each of which is characterised by a unique, random population of Ly $\alpha$  absorbers. By a population we mean a given number  $N_{\text{abs}}$  of absorbers and a set  $\{z_{\text{abs}}, N_{\text{HI}}, b\}_{i=1, \dots, N_{\text{abs}}}$ , where the absorption red-

shift  $z_{\text{abs}}$ , the column density  $N_{\text{HI}}$  and the Doppler parameter  $b$  uniquely determine the absorption properties of each systems. The population of each line-of-sight is computed in a Monte Carlo fashion, drawing the total number of absorbers and their corresponding parameters from distribution functions constrained from observations. Despite considerable efforts on the observational side, these distribution functions are not yet uniquely settled. We therefore chose to use two different sets of distribution functions and calculate for both of them the redshift evolution of the cosmic flux decrement. Contrasting the results from these computations with observations serves as a consistency check. We find that one of sets of distribution functions, when combined with the approach described, correctly reproduces the evolution of the cosmic flux decrement as derived from observations. We thus identify a particular model, called MMC (from Madau and Monte Carlo), which has successfully has passed the test against observations. We now use this model to analyse the impact of the stochastic attenuation on the photometric properties of galaxies, in particular, on their magnitudes. This allows us to estimate in how far photometric redshifts based on the measured spectral energy distributions (SEDs) are affected by this stochasticity.

### 7.2.1 Absorption masks

Since the absorption due to intergalactic HI along random lines-of-sight is different for each line-of-sight, and observations of different galaxies necessarily imply observations along different lines-of-sight, a non-intrinsic scatter in the observed photometric quantities can be expected, even for galaxies with identical intrinsic properties. This effect, as mentioned before, adds to the scatter in the observed colors and luminosities caused by variations in the intrinsic properties of the galaxies such as star-formation fluctuations, dust content, morphology, geometry, etc.

The effect of a population of absorbers along the line-of-sight is fully determined by the set of parameters  $\{z_{\text{abs}}, N_{\text{HI}}, b\}_{i=1, \dots, N_{\text{abs}}}$ . Assuming this to be given, the observed flux of a galaxy with *intrinsic* flux  $f_{\text{em}}$  at redshift  $z_{\text{gal}}$  is given by

$$f_{\text{obs}}(\lambda) = \frac{f_{\text{em}}[\lambda/(1+z_{\text{gal}})]}{1+z_{\text{gal}}} \Phi(\lambda), \quad (7.1)$$

where the quantity  $\Phi$  is the transmission factor and is

given by <sup>1</sup>

$$\Phi(\lambda) \equiv \prod_{i=1}^{N_{abs}} \exp[-\tau(\lambda/(1+z_{abs}))]. \quad (7.2)$$

where  $\tau$  is the absorption coefficient which includes the photoionisation cross-section and the absorption due to resonant scattering (Paper II).

It is important to note at this point that the emitted flux in equation (7.1) as given by our model galaxy spectra (see below), accounts for the *evolutionary* effect, *i.e.* the fact that galaxies at high redshifts are seen in younger evolutionary stages, and their intrinsic spectra are *not* equivalent to template spectra which have only been redshifted. This effect has to be taken into account when computing magnitudes in a given pass-band by including, in addition to the  $k$ -correction, an additional term, the so-called  $e$ -correction, due to this effect.

In a first approach, we consider galaxies of the same spectral type to be galaxies with identical intrinsic properties. The method used is briefly outlined: We assume an ensemble of  $N$  galaxies of the same intrinsic type to be distributed randomly across the sky, all at the same redshift  $z_{gal}$ . Note that each of these galaxies corresponds to the same model spectrum. Using the MMC model (cf. Paper II), we generate an ensemble of an equal number  $N_{LOS}$  of lines-of-sight, or equivalently, absorption masks. An absorption mask is just the attenuation factor  $\Phi$  in equation (7.2), and consists of a flux distribution as a function of wavelength, which encodes the cumulative absorption due to the population of absorbers along that particular line-of-sight (cf. Fig. 7.2). Applying each of this absorption masks<sup>2</sup> on the input spectrum, we obtain an ensemble of  $N_{LOS}$  attenuated spectra, from which we may calculate any photometric quantity. From the ensemble of values, it is then possible to compute its statistical moments, *i.e.* mean, standard deviation, etc., which may be compared to the corresponding observed values. However, since the statistics of these rather complex random variables is not known a priori, we do not compute mean values and standard deviations in the usual way, *i.e.* implicitly assuming Gaussian statistics, but rather determine

quantiles from the ensemble of simulated values (see section 7.2.2).

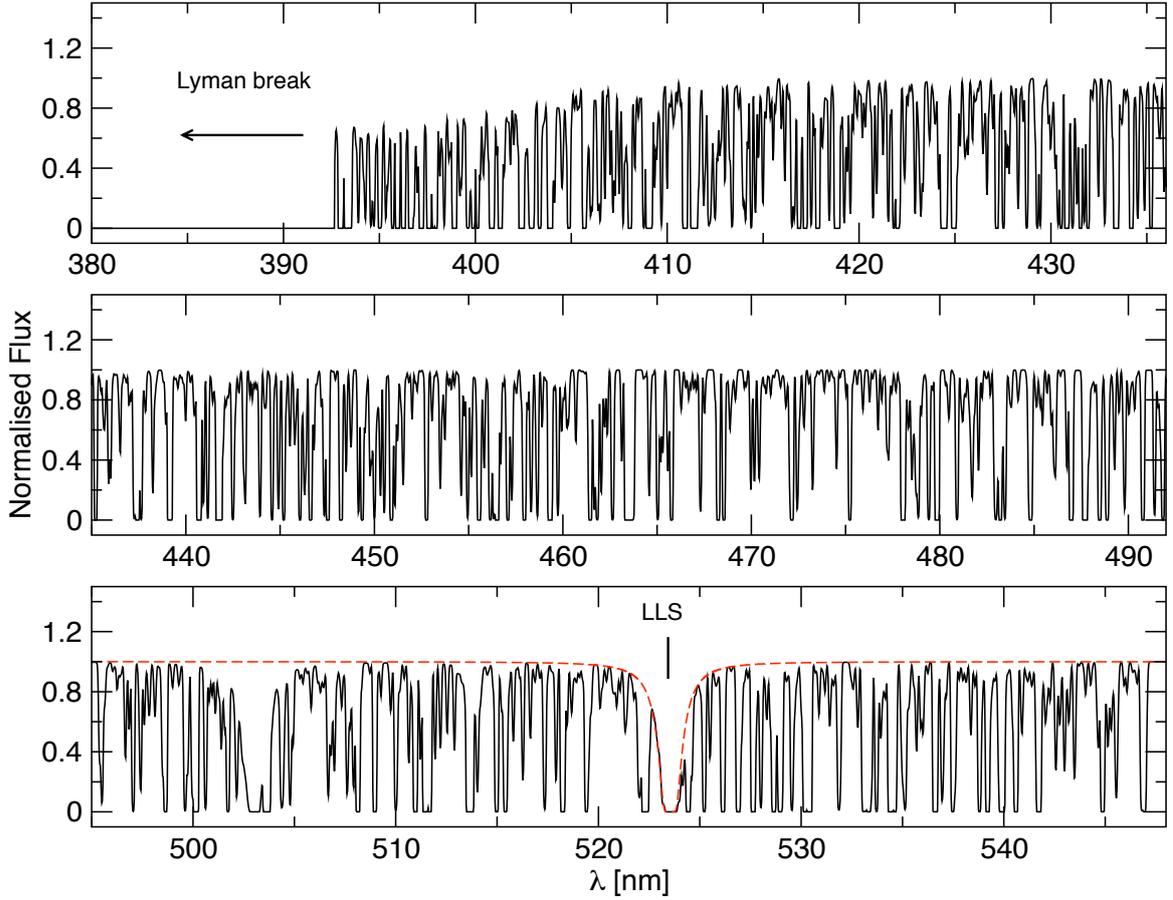
The MMC model as presented in Paper II has been extended here in order to include the absorption due to photoionisation of H I by photons with energies  $\lambda \leq 91.175$  nm, as well as due to the first ten Lyman transitions. The cosmic flux decrement  $D_A$  is per definitionem affected only by the Ly $\alpha$  transition. In contrast, the full spectrum of resonant transitions and the photoionisation may have a great impact on the luminosities and colors of an observed galaxy, depending on its redshift and the wavelength range in which it is observed. The Madau model (Madau 1995) includes the first 17 Lyman transitions, together with the absorption at the Lyman edge. Bershadsky et al. (1999) noted that including eight more transitions, *i.e.* a total of 25, results approximately in a 0.5 per cent change in the computed colors. In his model, Meiksin (2006) includes a total of 31 transitions. However, it can be estimated that the further six transitions have a negligible effect on the computed luminosities. On the one hand, the strength at line centre of the  $i$ th transition relative to Ly $\alpha$  for a fixed  $N_{HI}$  and  $b$  is proportional to  $\lambda_i \cdot f_{osc,i} \cdot [\exp(a^2) \operatorname{erfc}(a)]$ , assuming that the line profile is given by the Voigt-Hjerting function  $H(a, x)$ . The term in brackets correspond to  $H(a, 0)$  (see *e.g.* Zaghoul 2007) and  $\operatorname{erfc}(x) \equiv 1 - \operatorname{erf}(x)$  is the complementary error function. This term is rather insensitive to  $a$ , and hence the strength of the line of order  $i$  relative to the strength of the Ly $\alpha$  line at the line centre very accurately scales as  $\lambda_i \cdot f_{osc,i}$ . Using this, we compute the strength of all Lyman transitions relative to Ly $\alpha$ . The result is shown Fig. 7.1. As can be seen, the absorption due to each of the first ten transitions is of the order  $10^{-2}$  weaker than that of the Ly $\alpha$ . For higher-order transitions, the strength is of the order  $10^{-3}$  or smaller than that of Ly $\alpha$ .

Moreover, due to their decreasing central wavelength, the contribution of higher order lines to the total absorption due to intergalactic H I is only important at short wavelengths for a given redshift, or at higher redshifts for a fixed wavelength range, *i.e.* a given pass-band. Because of these reasons, and in order to safe computational time, we restrict ourselves to the inclusion of the first ten transitions in the computation of the absorption of a single absorber. The above shows that our results should not be grossly affected by this fact.

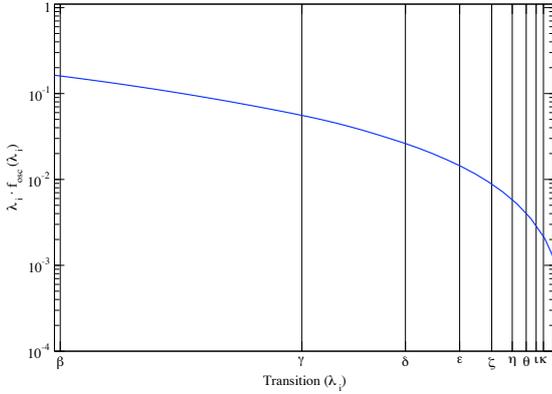
Much more important for the computation of the absorption is the approach that is used, as becomes apparent when comparing a mean absorption curve computed

<sup>1</sup> It should be clear that  $z_{abs}$  and  $\tau$  – through the dependence on  $N_{HI}$ ,  $\sigma_i$ , and  $b$  – in equation (7.2) are different in general for each absorber. However, we do not write this explicitly by *e.g.* introducing a new subscript in order to avoid confusion.

<sup>2</sup> The code to generate an absorption mask for any desired redshift will be made publicly available together with the paper.

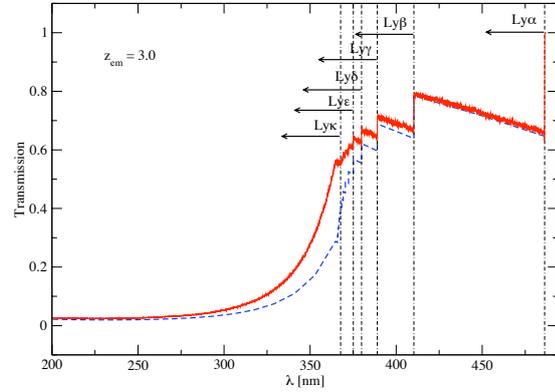


**Fig. 7.2:** Example of an absorption mask for a single line of sight out to a redshift  $z_{em} = 3.5$ , including the absorption at the Lyman edge and the first 10 Lyman resonant transitions for each absorber. For this redshift, the expected number of Ly $\alpha$  forest absorbers and LLSs is, according to equation (5.23) and the MMC input distribution functions, 2504.94, and 6.271, respectively. The actual number of systems of each type for this particular realisation, drawn from a Poisson distribution, is 2359 and 10, respectively. This unique line-of-sight hence intersects a total of 2369 absorbing systems. The arrow in the upper panel and the vertical line in the lower panel show, respectively, the Lyman break and the Ly $\alpha$  absorption feature of a Lyman Limit system at a redshift  $z_{abs} = 3.3066$ , with a column density  $N_{\text{HI}} = 19.6 \text{ cm}^{-2}$  and a Doppler parameter  $b = 35 \text{ km s}^{-1}$ . The red dashed line is a fit to the Ly $\alpha$  absorption line computed following [Tepper-García \(2006\)](#) using the above parameters.



**Fig. 7.1:** Absorption line strength at line centre, given as the product of the central wavelength  $\lambda_i$  and oscillatory strength  $f_{osc,i}$  of the corresponding transition, relative to the Ly $\alpha$  line strength. For visualisation purposes, only the transitions from Ly $\beta$  through Ly $\kappa$  are shown. Note the rapid decrease of the relative line strength with increasing order.

both using the Madau model and the MMC model. The attenuation curve for the MMC model is simply obtained by averaging the absorption at each given wavelength over a sufficiently large number of lines-of-sight. In other words, the mean absorption is the average of an ensemble of absorption masks at a fixed redshift  $z_{em}$ . This comparison is illustrated in Fig. 7.3 for  $z_{em} = 3.0$ . Here we show the arithmetical mean at each wavelength over an ensemble of  $4 \cdot 10^3$  absorption masks, which include the absorption at the Lyman edge and the first ten Lyman transitions for each absorber, computed according to the MMC model. When compared to the attenuation curve resulting from Madau’s analytical approach, it is evident that even using the same input distributions does not quite result in the same mean absorption computed by different methods, as already recognised by Bershady et al. (1999). The absorption according to Madau’s method is over-estimated with respect to the MMC model. Clearly, the difference is due to the different approaches used, since both methods use the same distribution functions of the absorber properties. It should be noted that the difference is already significant for the first ten transitions, and that the addition of further lines would certainly not dramatically increase the effect, for the reasons regarding the line strengths of higher-order transitions relative to Ly $\alpha$  explained above.



**Fig. 7.3:** Comparison of the mean absorption computed according to the MMC model averaging over  $10^4$  lines-of-sight (red, solid line) and the Madau model (blue, dashed line). The noisy appearance of the MMC curve is due to the stochastic nature of the absorption along random lines-of-sight. The wavelength at which the absorption due to successive resonant transitions becomes effective, giving rise to the characteristic stair-case profile, is indicated for the first ten transitions by the vertical dashed line. However, for simplicity, the transitions between Ly $\epsilon$  and Ly $\kappa$  have been omitted. Note the difference in the predicted absorption in the region between the Ly $\beta$  and Ly $\kappa$  lines.

### 7.2.2 The Statistics

Once we have an ensemble with a sufficiently large number of absorption masks, we may compute for any given spectrum an ensemble of values of any desired photometric observable, *e.g.* the flux in a given pass-band. The advantage of disposing of such an ensemble of values is that it can be easily treated statistically to obtain quantities of interest such as the mean and standard deviation around the mean. However, since the distribution function of such an observable is unknown a priori, we choose to treat our data using quantiles. Formally, the quantile  $q_x(A)$  of a random variable (observable)  $A$  can be defined as

$$x = 100 \cdot \int_0^{q_x} P_A(a) da, \quad (7.3)$$

where  $P_A$  is the normalised distribution function of  $A$

According to this definition, the median value of the observable  $A$  is  $q_{50}(A)$ . In analogy to Gaussian statistics, we define the mean of the distribution of  $A$  as the median, and the  $\pm 1$ -,  $\pm 2$ - and  $\pm 3$   $\sigma$  ranges around this

value as  $q_{50\pm 34.13} (A)$ ,  $q_{50\pm 47.72} (A)$ , and  $q_{50\pm 49.86} (A)$ , respectively.

In Paper II, we used this approach to compute the mean evolution and the scatter of the cosmic flux decrement  $D_A$ . We now apply it here to the analysis of the effect of the variation in the absorption along random lines-of-sight on the luminosities of galaxies and their photometric redshifts.

### 7.2.3 The Model Galaxy Spectra

We use as input spectra the model spectra obtained from the GALEV code (Bicker et al. 2004, and references therein), and we briefly explain it here for completeness. In order to keep the number of free parameters as small as possible, the model spectra are computed as one-zone, closed-box models, in which the different types of galaxies correspond to different time evolutions of their star-formation rates (SFRs), as derived from observations by Sandage (1986). For instance, an elliptical-type galaxy has an exponentially declining SFR, with an  $e$ -folding time of  $\tau_* = 10^9$  yr. The galaxies of type Sa-Sc have a SFRs which linearly scale with the evolving gas mass of the galaxy, *i.e.* SFR  $\propto G(t)/M_{tot}$ , where  $M_{tot}$  is the total mass (stars and gas). A Sd-type galaxy is described by a constant star-formation rate (CSFR). The characteristic times  $t_*$  for the transformation of gas into stars range from  $4 \cdot 10^9$  yr for Sa-type galaxies to larger than  $1.5 \cdot 10^{10}$  yr for Sd-type galaxies.

The distribution in mass of newly formed stars is assumed to be given by a Salpeter Initial Mass Function (IMF) with lower and upper mass limits  $m_l = 0.08 M_\odot$  and  $m_u \approx 70 M_\odot$ , respectively. The IMF is normalised to a fraction of visible mass (FVM) of 0.5 in order for the model galaxies to match after  $1.2 \cdot 10^{10}$  yr the mass-to-light (M/L) ratios of galaxies observed in the local Universe. Correspondingly, the mass of the galaxy is scaled in such a way that the luminosity in the  $B$  band matches the mean observed value of the corresponding galaxy type in the Virgo cluster (cf. Sandage 1986).

The stellar evolution is computed using isochrones from the Padova Group (Bertelli et al. 1994) in its version of November 1999, and later extended to include the thermal pulsing AGB phase for intermediate mass stars (cf. Schulz et al. 2002). The stellar spectra are taken from the library of models atmospheres by Lejeune et al. (1997, 1998), which covers the whole range of stellar metallicities, spectral types, and luminosity classes, and a wide wavelength range from 0.09

to 160,000 nm.

The redshift evolution of the galaxy spectra is computed using the GALEV code coupled to a cosmological tool assuming a formation redshift  $z_{form} = 9.8$ , embedded in a cosmology characterized by the set  $(H_0, \Omega_m, \Omega_\Lambda) = (0.72 h_{100}, 0.27, 0.73)$ , consistent with the most recent measurements of the cosmological parameters by Spergel et al. (2006). For our purpose, we had to increase by linear interpolation the "resolution" of the spectra to a constant resolution of  $\Delta\lambda = 0.25 \text{ \AA}$ , covering a wavelength baseline from 90.00 nm to  $121.60 \cdot (1 + z_{gal})$  nm, *i.e.* the range affected by intergalactic attenuation.

Magnitudes in any desired filter system are obtained by folding the spectra with the respective response functions for this filters (and detectors).

### 7.3 VARIATIONS IN MAGNITUDES

As an illustrative example, we calculate the effect of the attenuation in Johnson's  $U$ ,  $B$ ,  $V$ , and  $R$  as a function of galaxy redshift  $z_{gal}$  for the particular case of an Sd-type galaxy, which corresponds to a constant SFR. We choose this type of galaxy, because it has the higher UV flux at all epochs compared to the other types (E-Sa), where the absorption due to intergalactic H I is most severe. All magnitudes are given throughout this work in the AB photometric system (Oke 1974, Oke & Gunn 1983). The zero-point  $z_p(T)$  for a given normalised passband  $T(\lambda)$  in this system is given by

$$z_p(T) \equiv 2.5 \log \left[ \int_0^\infty f_{AB}(\lambda) \cdot T(\lambda) d\lambda \right], \quad (7.4)$$

where  $f_{AB}(\lambda) = 3.63 \cdot 10^{-20} c^2 \lambda^{-2} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ .

The apparent magnitude in the passband  $T(\lambda)$  of a galaxy at redshift  $z_{gal}$  with intrinsic flux  $f_{em}$  along a particular random line-of-sight is given by

$$m_\lambda^i(T; z_{gal}) = -2.5 \log \left[ \int_0^\infty f_{obs}^i(\lambda) \cdot T(\lambda) d\lambda \right] + z_p(T) + BDM(z_{gal}), \quad (7.5)$$

where  $f_{obs}^i$  is given by equation (7.1), and the subscript  $i$  denotes the  $i$ th line-of-sight with  $i = 1, \dots, N_{LOS}$ . The quantity  $BDM$  is the bolometric distance modulus, and it accounts for the dimming of light due to the distance of the observed source. It is important to realise that the observed flux as given by equation (7.1) encodes not only the  $k$ -correction, but the  $e$ -correction as well.

Recall that the  $e$ -correction accounts for the fact that galaxies at high redshifts are seen in younger evolutionary stages, and their intrinsic spectra are *not* equivalent to template spectra which have been redshifted.

For a given passband  $T$  and a fixed redshift  $z_{gal}$ , we compute the magnitudes for a total of  $4 \cdot 10^3$  line-of-sight, obtaining an ensemble of an equal number of magnitudes  $\{m_\lambda^i(T; z_{gal})\}$ . From these, we calculate the mean magnitude  $q_{50}(m_\lambda(T; z_{gal}))$  and the 1-, 2-, and  $3\sigma$  ranges around this value as explained in section 7.2.2. We do this for the redshift range  $0.01 < z_{gal} < 5.0$ .

### 7.3.1 The intergalactic $k$ -correction

We are primarily interested in the difference between magnitudes with and without intergalactic attenuation, a quantity designated by Meiksin (2006) as intergalactic  $k$ -correction and denoted by  $k_{IGM}$ . We find the term appropriate and thus adopt it here. However, it should not be confused with the "usual"  $k$ -correction which is due solely to the expansion of the Universe, and which is already included in equation (7.5) through  $f_{obs}^i$  and equation (7.1).<sup>3</sup> In general,  $k_{IGM}$  is a function of redshift, the given passband and, in our case, it is also depends on the line-of-sight. Thus,

$$k_{IGM}^i(T; z_{gal}) \equiv m_\lambda^i(T; z_{gal}) - m_\lambda^0(T; z_{gal}) \quad (7.6)$$

where the reference magnitude  $m_\lambda^0(T; z_{gal})$  corresponds to the magnitude *without* attenuation, and is given by equations (7.5) and (7.1) with  $\Phi \equiv 1$ . Since the intergalactic  $k$ -correction and the apparent magnitude only differ by an additive constant, the mean, and sigma ranges of  $k_{IGM}$  can be computed directly from those of  $m_\lambda$ .

In Fig. 7.4, we show the evolution of  $k_{IGM}$  in Johnson's  $U$ ,  $B$ ,  $V$ , and  $R$  bands for a Sd-type galaxy as a function of  $z_{gal}$ . Shown are the mean  $k_{IGM}$  and the  $\pm 1$ -,  $\pm 2$ -, and  $\pm 3\sigma$  ranges around this value. We include for comparison the values of  $k_{IGM}$  computed from the Madau's and the Meiksin's models (cf. Paper II). The evolution of  $k_{IGM}$  is qualitatively the same in all passbands, and therefore we will focus on the discussion of the  $U$  band. The  $k_{IGM}$  becomes noticeable at redshifts  $z_{gal} \gtrsim 1.5$ , at which the Ly $\alpha$  line is shifted into the filter. Between redshifts 2 and 2.3, the Ly $\beta$  and higher-order transitions are shifted into the passband and correspondingly, the  $k_{IGM}$  increases. Furthermore, the scat-

ter in the absorption and consequently in  $k_{IGM}$  also increases, since the number of absorbers, especially of those with the highest column densities, dramatically increases with redshift. At  $z_{gal} = 2.35$ , the Lyman edge enters the passband and the intergalactic  $k$ -correction rises correspondingly. At this point the scatter in the absorption increases even more, since now the Lyman edge of each absorber contributes to the dimming of the flux and the presence of a few less or a few more LLSs significantly changes the absorption and consequently  $k_{IGM}$ . At redshifts slightly lower than the redshift at which the Lyman edge is shifted to the wavelength range at which the transmission of the filter reaches its maximum, the ever increasing scatter in the absorption becomes even more apparent, until the galaxy becomes so faint that is not longer observable in the passband. The redshift as which this happens is known as the drop-out redshift, and we define as follows: If  $T(\lambda)$  is the response function of a given filter, and  $\lambda_{eff}$  the transmission-averaged wavelength of the filter response function, *i.e.*

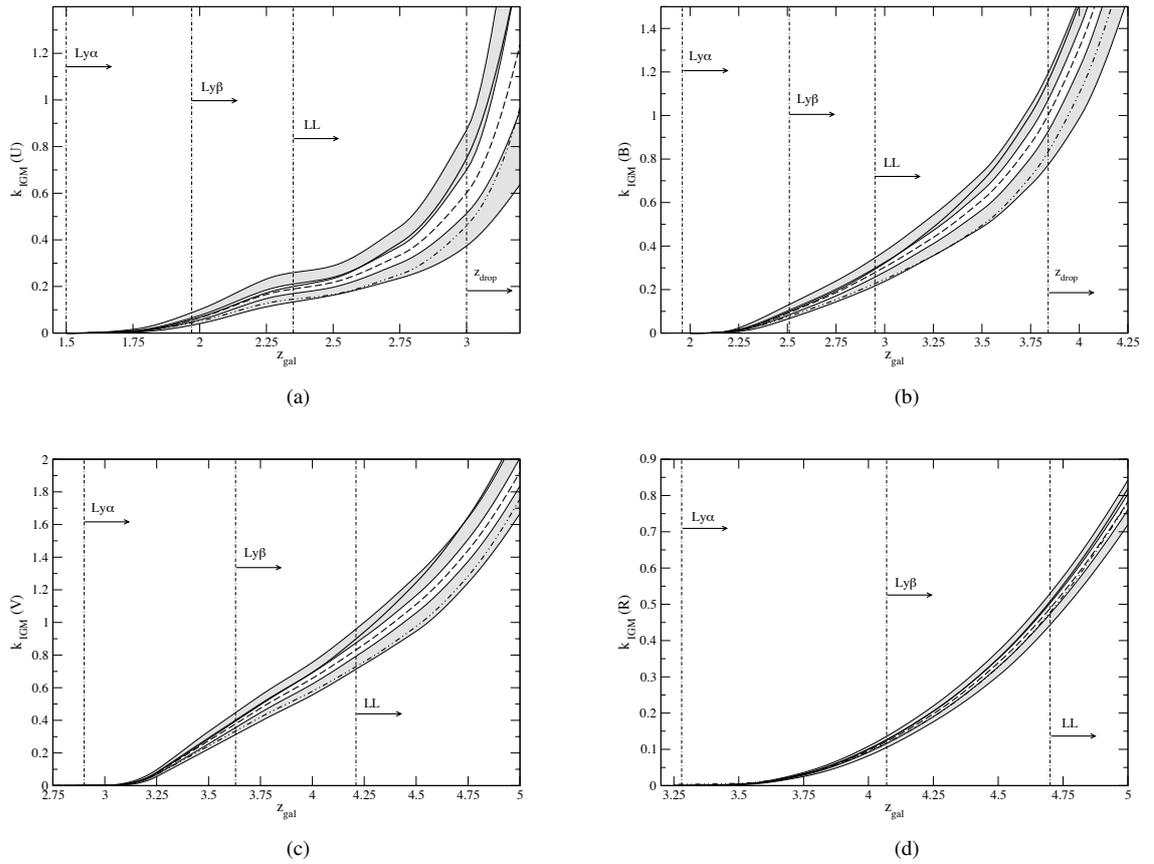
$$\lambda_{eff} \equiv \frac{\int_0^\infty \lambda T(\lambda) d\lambda}{\int_0^\infty T(\lambda) d\lambda}, \quad (7.7)$$

then the drop-out redshift  $z_{drop}$  is the redshift of the galaxy at which its Lyman break is just shifted rearward of  $\lambda_{eff}$ , *i.e.*  $z_{drop} \equiv \lambda_{eff}/\lambda_L - 1$ , where  $\lambda_L = 91.175$  nm. Accordingly, the drop-out redshift for the standard Johnson  $U$  band is  $z_{drop} = 366.6 \text{ nm}/\lambda_L - 1 = 3.0$ . Any estimate of the (mean) luminosity and scatter around it in this passband is in principle of no use for  $z > 3.0$ , since the galaxy is no longer observable, as a result of the fact that with increasing redshift, its Lyman edge moves through the filter, and the galaxy gets gradually fainter, ultimately dropping below the detection threshold. The drop-out redshift, however, should not be taken as a rigorous limit, but only as an indication that the predicted magnitude of the galaxy should be taken with caution. We list in table 7.1 the drop-out redshifts of further passbands, together with the redshifts at which the Lyman edge, the Ly $\beta$  and the Ly $\alpha$  lines are shifted into the corresponding passband.

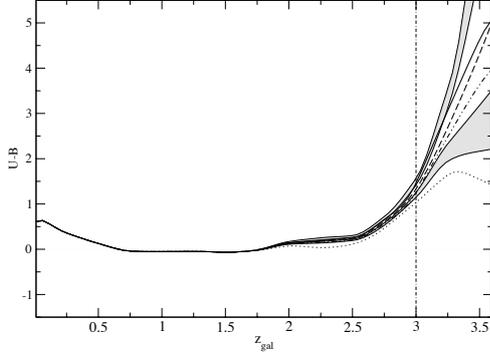
The increase in the scatter of  $k_{IGM}$  at redshifts in the vicinity of  $z_{drop}$  is best seen in the evolution of the colors, as shown in Figs. 7.5, 7.6, and 7.7, where we show the evolution of the  $U-B$  and  $B-V$  colors as function of redshift, where the  $k_{IGM}$  correction computed according to the MMC, the Madau, and the Meiksin models.

Let us highlight two important facts: First, as already

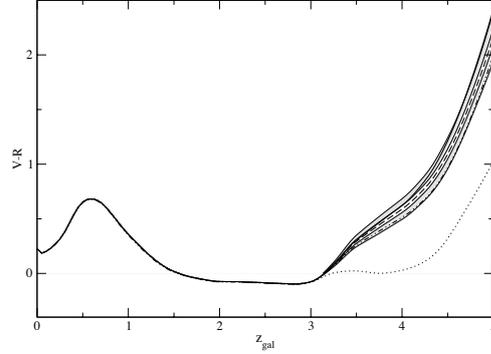
<sup>3</sup> For an excellent reference on the computation of the  $k$ -correction see Hogg et al. (2002).



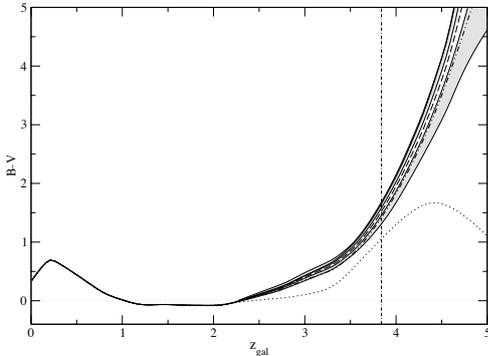
**Fig. 7.4:** Intergalactic  $k$ -correction in Johnson's (a)  $U$ , (b)  $B$ , (c)  $V$ , and (d)  $R$  bands for an Sd-type galaxy using different models for the attenuation: Madau's model (solid line), Meiksin's model (dot-dashed line), and the MMC model, where the median is given by the dashed line, the  $\pm 1\sigma$  range is the white area around the median, and the  $\pm 3\sigma$  range is given by the shaded area. The vertical dashed lines indicate the redshift at which a given absorption feature, *e.g.* the Lyman edge or a absorption line, is shifted into the corresponding passband, as indicated. The vertical dot-dashed line indicates the drop-out redshift, *i.e.* the redshift at which the Lyman edge of the galaxy is shifted redward of the wavelength at which the transmission in the corresponding filter reaches its maximum. These values are summarised in table 7.1.



**Fig. 7.5:** Evolution of  $U - B$  of an Sd-type galaxy including the effect of the absorption along random lines-of-sight using the Madau model (thick solid line), the Meiksin model (thick double-dot dashed line), and the MMC model: mean (thick dashed line),  $\pm 1 \sigma$  range (white area around the mean), and  $\pm 3 \sigma$  range (white plus gray area around the mean). The vertical dot-dashed line indicates in each case the drop-out redshift of the bluest passband.



**Fig. 7.7:** Evolution of  $V - R$  of an Sd-type galaxy including the effect of the absorption along random lines-of-sight. Lines and shaded areas as in Fig. 7.5.



**Fig. 7.6:** Evolution of  $B - V$  of an Sd-type galaxy including the effect of the absorption along random lines-of-sight. Meaning of lines and shaded areas as in Fig. 7.5.

argued, the difference in the approach to account for the intergalactic attenuation even when using the same input distributions for the absorber parameters becomes evident from the evolution of the mean  $k_{\text{IGM}}$  using the Madau model and the median as computed with the MMC model. However, the differences between these

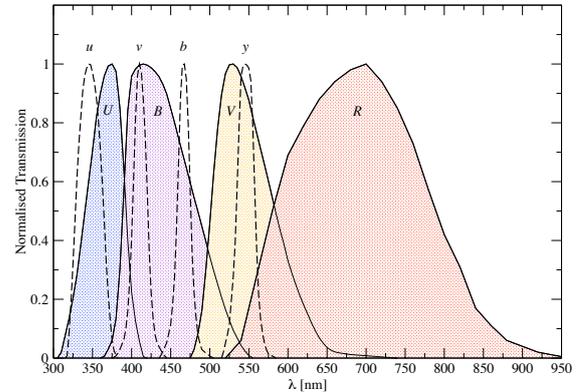
models, and between the Meiksin model amount to less than 0.2 mag at  $z_{\text{gal}} \leq z_{\text{drop}}(U)$ , and both the Madau model and the Meiksin model are all consistent with the MMC model at the  $3 \sigma$  level. The larger differences at redshifts above the drop-out redshift need not be considered, since they escape observations. The second important fact to be highlighted is the modest scatter in the evolution of  $k_{\text{IGM}}$  or  $m_{\lambda}$ , as already pointed out by Bershady et al. (1999). This is somehow counterintuitive, since one would expect that the contingent presence of a few absorbers less or more than average at column densities  $N_{\text{HI}} \geq 17.21 \text{ cm}^{-2}$  – the column density demarcation between Ly $\alpha$  forest and Lyman limit systems – cause sufficient differences in the absorption for galaxies with different sight-lines, causing them to have significantly different brightness, and not just scatter around the expected magnitude at a level of  $\pm 0.1$  mag. The reason for this counterintuitive effect is that at all redshifts, the absorption features, *i.e.* Ly $\alpha$ , Lyman break, etc., of the individual absorbers are blueward of the corresponding strong absorption features of the observed galaxy. Moreover, the mean number of Ly $\alpha$  forest systems dramatically increases with redshift, but their scatter and the corresponding scatter in their absorption is not significant (cf. Paper II, Section 5.3). In other words, even though the increasing mean number of Ly $\alpha$  forest absorbers causes the mean  $k_{\text{IGM}}$  to increase, their scatter does not cause the dispersion in  $k_{\text{IGM}}$  to increase as well. Besides, these systems dominate the absorption in the line-region, *i.e.* at redshifts

lower than  $z_{LL}$ , over the optically thicker LLSs (cf. discussion in Paper II). These can have a significant contribution to the absorption only in the Lyman break region, *i.e.* at rest-frame wavelengths smaller than 91.175 nm, through photoelectric absorption of light. This effect is reflected by the increase in the scatter of  $k_{\text{IGM}}$  at redshifts just below  $z_{\text{drop}}$ . The overall trend is certainly not affected by the fact that only the first ten Lyman transitions are included.

A further reason for the scatter in the absorption in the line-region to be modest is the fact that the process of computing broad-band magnitudes effectively averages the observed flux over a wide wavelength range, typically of order 200–300 Å. For comparison, the Ly $\alpha$  absorption line of systems with a column density  $N_{\text{HI}} = 10^{18} \text{ cm}^{-2}$  and a Doppler parameter  $b = 36 \text{ km s}^{-1}$  has a rest equivalent-width of the order of 1 Å. At  $z_{\text{gal}} = 2.5$ , the mean number of systems with column densities in the range  $10^{17.21} - 10^{20} \text{ cm}^{-2}$  is 5. Hence, assuming the absorption lines of all optically thick systems at this redshift have this same equivalent width, the combined absorption is around 1/10 of the width of a typical broad-passband. This optimistic estimate assumes implicitly that all absorbing systems are just at the right redshift in order for their absorption lines to fall into the wavelength range of the given passband. Such a clustering and especially at all lines-of-sight is highly improbable. For the case of the more numerous, but optically thinner Ly $\alpha$  forest clouds, the net effect is that, by averaging the observed flux over a given filter, the absorption lines are "smeared out", which diminishes the range in transmission values and hence its scatter as computed from its quantiles.

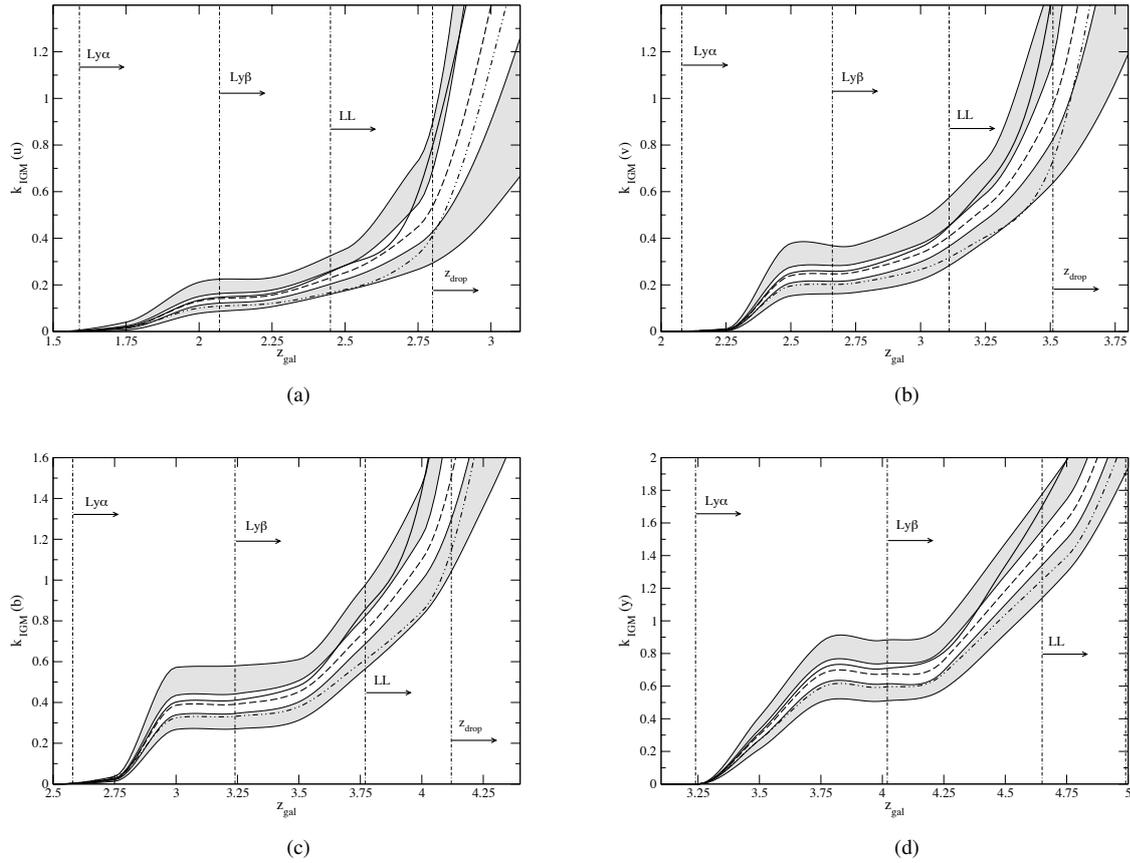
The line of reasoning presented above implies also that the scatter in the absorption due to stochasticity would be larger if averaging the observed flux, *i.e.* computing magnitudes, over narrower wavelength ranges. In other words, the variation in the absorption from line-of-sight to line-of-sight should be larger for narrower filters due to the fact that the absorption profiles correlate over small  $\Delta\lambda$ . In order to test this, we repeat the exercise described above this time using Strömrgren's  $u$ ,  $v$ ,  $b$ , and  $y$  passbands, which are narrower than Johnson's filters. A comparison of the filter response functions is shown in Fig. 7.8. The evolution of the magnitudes in the Strömrgren filters is shown in Fig. 7.9.

As suspected, the scatter in all four passbands is larger than in the Johnson filters, even though it is still of the order of  $\pm 0.2$  mag just below the drop-out redshift in the  $u$  passband. Also, and in contrast to the



**Fig. 7.8:** Comparison of the bandwidth of Standard Johnson's  $U$ ,  $B$ ,  $V$ , and  $R$ , and Strömrgren  $u$ ,  $v$ ,  $b$ , and  $y$ .

Johnson filters, in all these four passbands the characteristic features in the evolution of  $k_{\text{IGM}}$  caused by the shift of the different absorption lines into and out of the filters are more pronounced, again due to the narrower bandwidth of these filters compared to Johnson's. In particular, it is clearly seen how Ly $\alpha$  is shifted into a passband, causing the  $k_{\text{IGM}}$  to increase and to reach a local maximum at redshifts where Ly $\alpha$  is shifted to  $\lambda_{\text{eff}}$  (see equation 7.7). From this point on,  $k_{\text{IGM}}$  remains roughly constant at the  $\pm 1 \sigma$  level, with Ly $\alpha$  lines being shifted into and out of the filter, until it rises again when the Ly $\beta$  lines is shifted into the filter. At the  $\pm 1 \sigma$  level, this plateau remains visible until higher redshifts, due to the fact that Ly $\beta$ , being roughly a factor 10 weaker than Ly $\alpha$  (see Fig. 7.1), does not contribute significantly to the absorption at this level. Only when higher-order lines are shifted into the passband and Ly $\beta$  is shifted towards  $\lambda_{\text{eff}}$ , does  $k_{\text{IGM}}$  rise again. Such a plateau is evidently not present in the Johnson filters. The reason is that these filters are not narrow enough to allow the Ly $\alpha$  to be shifted to  $\lambda_{\text{eff}}$  without the Ly $\beta$  being already shifted into the filter, causing  $k_{\text{IGM}}$  to rise continuously, as can be seen in Fig. 7.4. This becomes even more apparent when comparing the evolution of  $k_{\text{IGM}}$  in the  $U$  and the  $R$  filters. The  $U$  band, being narrower than the  $R$  band, still allows for the effect to be marginally observable, while the evolution of  $k_{\text{IGM}}$  in the latter is monotonically increasing due to the enormous bandwidth of this filter. To this adds the fact that the mean number of absorbers and hence the effective absorption increases with redshift, which makes the evolution of  $k_{\text{IGM}}$  stronger.



**Fig. 7.9:** Intergalactic  $k$ -correction in Strömgren (a)  $u$ , (b)  $v$ , (c)  $b$ , and (d)  $y$  bands for an Sd-type galaxy using different models for the attenuation: Madau's model (solid line), Meiksin's model (dot-dashed line), and the MMC model, where the median is given by the dashed line, the  $\pm 1\sigma$  range is the white area around the median, and the  $\pm 3\sigma$  range is given by the shaded area. The vertical dashed lines indicate the redshift at which a given absorption feature, *e.g.* the Lyman edge or a absorption line, is shifted into the corresponding passband, as indicated. The vertical solid line indicates the drop-out redshift, *i.e.* the redshift at which the Lyman edge of the galaxy is shifted redward of the wavelength at which the transmission in the corresponding filter reaches its maximum. These values are summarised in table 7.1.

**Table 7.1:** Characteristic redshifts for two different sets of filters: Standard Johnson’s  $U$ ,  $B$ ,  $V$ , and  $R$  and Strömrgren  $u$ ,  $b$ ,  $v$ , and  $y$ . The third, fourth, and fifth column contain the redshift at which the Lyman edge, the  $\text{Ly}\beta$ , and the  $\text{Ly}\alpha$  line are redshifted in the given passband, respectively. The last column gives the drop-out redshift. The corresponding filter functions are available at the Asiago Database on Photometric Systems <sup>a</sup> (Moro & Munari 2000).

<i>System</i>	<i>Passband</i>	$z_{LL}$	$z_{\text{Ly}\beta}$	$z_{\text{Ly}\alpha}$	$z_{drop}$
Johnson	$U$	2.35	1.97	1.51	3.00
Johnson	$B$	2.95	2.50	1.96	3.84
Johnson	$V$	4.21	3.63	2.91	5.04
Johnson	$R$	4.70	4.07	3.28	6.61
Strömrgren	$u$	2.46	2.07	1.59	2.80
Strömrgren	$v$	3.11	2.66	2.09	3.51
Strömrgren	$b$	3.77	3.24	2.58	4.12
Strömrgren	$y$	4.65	4.02	3.24	4.99

<sup>a</sup> <http://ulisse.pd.astro.it/Astro/ADPS>

Since the four Strömrgren passbands have roughly the same width, the plateau-effect is stronger for the passbands at larger effective wavelength, because of the combined effect of the increase in the wavelength range between absorption lines and the increase in the strength of  $k_{\text{IGM}}$  with redshift.

In either case, however, the scatter in the evolution of  $k_{\text{IGM}}$  is small both in broad- or narrow-band filters to be observable. We conclude that the magnitudes of galaxies observed in deep surveys are not expected to scatter significantly due to differences in the absorption along different lines-of-sight. In other words, the stochasticity effects in the absorption do not contribute significantly to the actual observed scatter, which must hence be caused by differences in the intrinsic properties of galaxies.

#### 7.4 IMPLICATIONS FOR THE ESTIMATES OF PHOTOMETRIC REDSHIFTS

Measurements of spectroscopic redshifts are very time-consuming for distant, very faint galaxies. Instrumental as well as time allocation limitations make it practi-

cally impossible to systematically obtain spectroscopic redshifts for large numbers of distant galaxies. Accessible are only a few of the very brightest objects at any redshift beyond 2 – and these, for sure – are not representative of the population at those redshifts. Deep multi-band photometry in combination with dedicated photometric redshift techniques readily provide reliable photometric redshifts for large numbers of galaxies covering a wide range in intrinsic luminosities out to  $z \gtrsim 5$ .

The estimate of redshifts using broad-band photometry has been significantly improved in the last years, obtaining a match between spectroscopic and photometric redshifts of the vast majority of galaxies of unprecedented accuracy (see *e.g.* Sawicki et al. 1997, Benítez 2000, Coe et al. 2006). However, there are still a few cases where the spectroscopic redshifts have been accurately and reliably determined, and the corresponding photometric redshifts deviate dramatically from them (so-called *catastrophic* redshifts). This fact indicates that the technique used to estimate photometric redshifts still has some problems. Classical sources of error are the photometric errors, even though it is thought that these do not dominate the residuals between spectroscopic and photometric redshifts (Lanzetta et al. 1998). Further uncertainties in the estimate of the photometric redshifts are introduced by the different methods used, *i.e.* the *empirical training-set* and the *SED-fitting* methods (Yee 1998), due to the particular approach used. The *training-set* approach consists basically in using a set of observed SEDs of galaxies for which redshifts are available to derive an empirical, functional relation between redshifts  $z$  and colours  $C_i$  of the form  $z = z(C_i)$ . This relation is then used to compute the redshift of an object for which the colors are known – hence the term “*training-set*”. However, it is a strong assumption that such a functional form exists and is well-behaved, *i.e.* bijective, and this assumption is not necessarily given in general. Furthermore, the extrapolation of this functional form, assuming that it exists, to values of the colours out of the range of the training set, may lead to wrong estimates of the object’s redshift. The *SED-fitting* technique consists in comparing the observed spectral energy distributions (SEDs) of an object with a manifold of theoretical SEDs and corresponding redshifts, computed using evolutionary synthesis models such as GISSEL (Bruzual A. & Charlot 1993), PEGASE (Fioc & Rocca-Volmerange 1997), or GALEV (Bicker et al. 2004, and references therein). Compared to methods based on local template spectra, these evolutionary synthesis libraries have the advantage to

include the evolutionary effect, *i.e.* account for the fact that galaxies at higher redshifts are observed in younger evolutionary stages. GALEV models, in particular, have the additional advantage that they also account for the importance of subsolar stellar metallicities in galaxies at higher redshifts. This fact is of particular importance in the context of photometric redshifts for galaxies from deep multi-band imaging.

The comparison between measured and theoretical SEDs can be achieved, for example, through a  $\chi^2$ -technique, in which case the object is assigned the redshift corresponding to the best matching model, *i.e.* that with the minimum  $\chi^2$ . Recently, another approach to find a best-matching model based on Bayesian statistics has been shown to provide very accurate photometric redshifts when compared to spectroscopic ones (see *e.g.* Benítez 2000).

Independently of the statistical treatment used to find a best model, several factors may introduce large uncertainties in the redshift estimate for an object. For instance, local-galaxy template SEDs may not be compared to high-redshift objects, since their evolutionary stages are very different. In order to account for this, the template SEDs used should include the evolutionary (or *e*-)correction, *i.e.* the fact that galaxies have different colors at different evolutionary stages (Bicker *et al.* 2004). They also should account for the *k*-correction, *i.e.* the dimming and redshifting of the emitted intrinsic flux due to the expansion of the Universe. Another important factor that affects the observed SEDs is the restframe UV attenuation due to absorption of light in the intergalactic H<sub>I</sub>. A widely used, almost standard, approach to correct for this is to add the mean absorption using the Madau model (Madau 1995), as done *e.g.* by Fontana *et al.* (2000). There are two concerns about this: First, this approach does not take into account that the absorption along random line-of-sight is stochastic in nature. We showed above that the scatter in the intergalactic absorption can have a non-negligible impact on the colors of different types of galaxies, and this may introduce a further uncertainty in the estimate of photometric redshifts, as first noted by Yee (1998). Furthermore, it was shown above that the mean attenuation at a given redshift computed using the Madau model does not match the mean attenuation predicted by the MMC model and, as first shown by Bershadsky *et al.* (1999), it turns out that the correct way of accounting for the mean intergalactic attenuation is the Monte Carlo approach. It is true that the scatter in *individual* luminosities and colors due to the stochastic nature of the

absorption is not significant. However, it needs to be explored how this scatter affects the estimate of photometric redshifts, given the complex topology of the manifold of solutions. The goal of this section is thus to quantify the uncertainties introduced in the estimate of photometric redshifts when the model SEDs only account for the mean intergalactic attenuation.

#### 7.4.1 Effects on the photometric redshifts

In order to assess the effect of variations in the intergalactic absorption from line-of-sight to line-of-sight, we generate at a given redshift  $z_{gal}$  a sample of mock SEDs which shall be compared to a grid of model SEDs. The mock SEDs are computed in Johnson's *U*, *B*, *V*, and *R* bands for a given model galaxy input spectrum which has been affected by intergalactic attenuation using our ensemble of absorption masks, and which includes the *e*- and *k*-corrections as computed with the GALEV code. Since the absorption encoded in each mask is unique, it is clear that the SEDs computed in this way will, in general, differ from each other. This procedure mimics the observations of galaxies of a given type and at a given redshift, *i.e.* with an identical intrinsic spectrum, along different directions in the sky. The model SED at each given redshift is computed for the same input spectrum, corrected for intergalactic attenuation with the Madau model. This is done for redshifts  $0.01 < z_{gal} < 5.0$ .

The comparison of mock and model SEDs is performed with a modified version of the AnalySED tool which makes use of the  $\chi^2$  minimisation technique (for a detailed description of the AnalySED tool see Anders *et al.* 2004). As usual, the  $\chi^2$  value is given by

$$\chi^2 \equiv \sum_i^{N_{mag}} \frac{[(m_{mock}^i - m_{off}) - m_{mod}^i]^2}{\sigma_i^2}. \quad (7.8)$$

where  $N_{mag}$  is the number of points, *i.e.* magnitudes in the SED. The quantity  $m_{off}$  is the offset between the the mock and model SEDs averaged over all magnitudes, and is needed to compute the photometric mass and to gauge the mock SEDs. The uncertainty  $\sigma_i$  of each term in the sum is given by the quadratic sum of the individual uncertainties in the corresponding magnitude of the model SED and the mock SED, *i.e.*

$$\sigma_i^2 = \sigma_i^2(mock) + \sigma_i^2(model). \quad (7.9)$$

We estimate that magnitudes computed from our model galaxy spectra, and hence both our model and

mock SEDs, are accurate to within  $\pm 0.1$  mag at the  $1\sigma$  level.

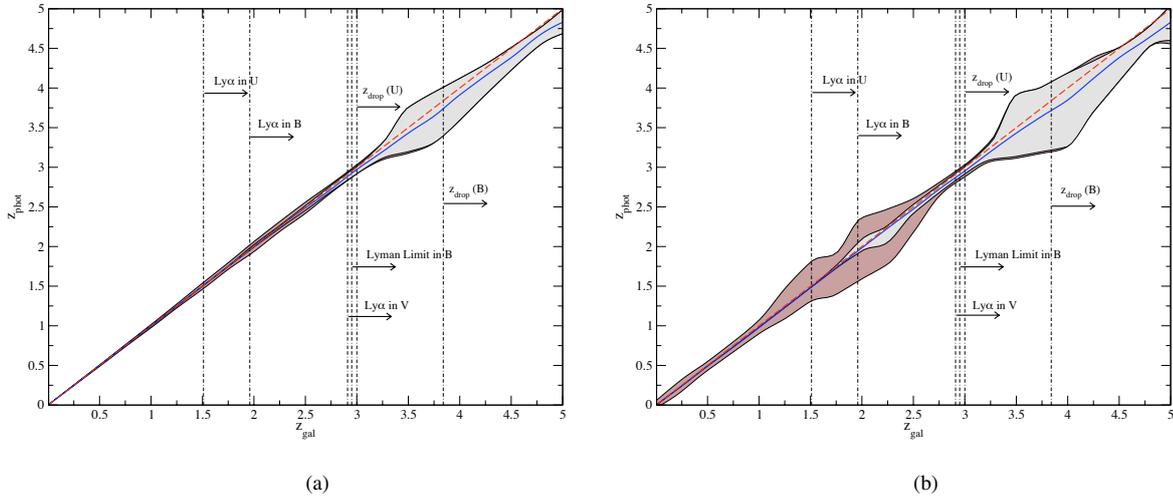
Each model SED is assigned a probability according to  $\exp(-\chi^2)$ . The best-matching model SED is obviously that with the highest probability. A  $\pm 1\sigma$  uncertainty interval is defined, in analogy to Gaussian statistics, by summing the probabilities assigned to each model SED arranged in descending order up to a total value of 0.68. The parameter, in our case the redshift, corresponding to the template SEDs with the largest and smallest probability within this range define the  $\pm 1\sigma$  uncertainty in the estimate of the parameter. We do not introduce further constraints into this analysis, as *e.g.* flagging mock SEDs of galaxies at drop-out redshifts for a given filter.

The goal of the procedure described above is to quantify to which extent the redshift of a model galaxy whose spectrum includes stochastic intergalactic attenuation is recovered when compared to a grid of model SEDs that are affected only by mean intergalactic attenuation. In a first approach, we limit our analysis to an Sd-type input spectrum, both for the mock and model SEDs. Following this approach, we obtain for each mock SED in the sample at a fixed  $z_{gal}$  a corresponding ensemble of best estimates  $BE(z_{phot})$  for  $z_{phot}$ , and their corresponding uncertainties  $\pm \sigma_m(z_{phot})$  as given by the procedure described above. In addition, we compute for the ensemble of  $BE(z_{phot})$  the median value  $\mu(z_{phot})$  and define the dispersion of  $BE$ ,  $\pm \sigma_{BE}(z_{phot})$ , by  $\max(BE) - \mu$  and  $\mu - \min(BE)$ , respectively, at each redshift. We follow two approaches, which differ only by the fact that the average off-set  $m_{off}$  between the mock and model SEDs is included or not. The result of this comparison is shown in Fig. 7.10. Here, we show at each  $z_{gal}$  the median value  $\mu(z_{phot})$  for the distribution of the best estimates  $BE(z_{phot})$  (blue-solid line), together with its corresponding dispersion,  $\pm 1\sigma_m(z_{phot})$  (brown-shaded area) and  $1 \pm \sigma_{BE}(z_{phot})$  (gray-shaded area), for: (a) no off-set, and (b) with off-set. Let us first focus on case (a). Note first that the galaxy redshifts are almost exactly recovered at redshifts  $z_{gal} \lesssim 3$  with  $1\sigma \leq 0.01$ . Above this redshift, the galaxy drops out of the  $U$  band and also the scatter in the attenuation becomes large, up to  $1\sigma \approx 0.25$ . Both these results are expected from the small scatter in the colors due to stochastic intergalactic attenuation, as shown in Fig. 7.5, and the increase in the scatter at redshifts larger than  $z_{drop}$ , as shown in Fig. 7.4. Note further that the scatter as computed from the quantile formalism and from the AnalySED tool are almost identical at all red-

shifts. Interestingly, the value of  $\mu(z_{phot})$  is slightly lower at redshifts  $z_{gal} \gtrsim 3$ , which means that photometric redshifts estimated using a set of SEDs affected only by a mean intergalactic absorption are slightly *underestimated*. This may be due to the fact that the scatter in the evolution of the magnitudes and colors is asymmetric with respect to the median, as can be seen in Figs. 7.4 and 7.5.

The bottom line of this analysis is that the stochastic intergalactic attenuation has little or no effect on the estimate of photometric redshifts under ideal circumstances, specifically, when the type and especially the mass of the galaxy is known. This is, however, not given in general, and for this reason it is necessary to include the magnitude off-set in the computation of the  $\chi^2$  value (cf. equation 7.8). Surprisingly, it turns out that in this case the effect of the stochastic attenuation is amplified, as illustrated in Fig. 7.10, panel (b). As can be seen, both the scatter as given by  $\sigma_m(z_{phot})$  is larger at all redshifts compared to its value for case (a). In particular,  $\sigma_m(z_{phot})$  increases slowly at  $z_{gal} \gtrsim 1$ , and becomes large in the range  $1.5 < z_{gal} < 2.75$ . In particular,  $\sigma_m(z_{phot})$  increases significantly at redshifts for which the Ly $\alpha$  line is shifted into the  $U$  and the  $B$  bands, respectively, making the estimate of the photometric redshift uncertain to within  $\Delta z_{phot} \leq 0.25$ , in agreement with estimates by Meiksin (2006) using similar passbands. At  $z_{gal} \approx 3$ , a series of coincidences decreases the scatter in  $z_{phot}$ : first, this redshift corresponds to the drop-out redshift for  $U$  band as defined above; also at this redshift, the Lyman edge is shifted into the  $B$  band, while Ly $\alpha$  moves into the  $V$  band. At  $z_{gal} > 3$ , the scatter again increases. Note, however, that the behaviour of  $\mu(z_{phot})$  in this case is almost identical to that in case (a), and that this is close to the identity line at all redshifts. Furthermore,  $\sigma_{BE}(z_{phot})$  does not differ significantly from its counterpart in case (a), which is consistent with the result that the best estimates of  $z_{gal}$  are not significantly affected by the stochastic effect of intergalactic attenuation.

The reason for the large scatter  $\sigma_m(z_{phot})$  introduced by the magnitude off-set is rather puzzling. However, this effect is not an artifact of the code, since we made sure that a perfect match between  $z_{gal}$  and  $z_{phot}$  is obtained in both cases, *i.e.* with and without introducing the magnitude off-set, for identical mock and model SEDs. Since, as explained above, the magnitude off-set is related to the overall scaling of the flux, and hence to the mass of the galaxy, we expect the magnitude of the scatter to be larger when estimating pho-



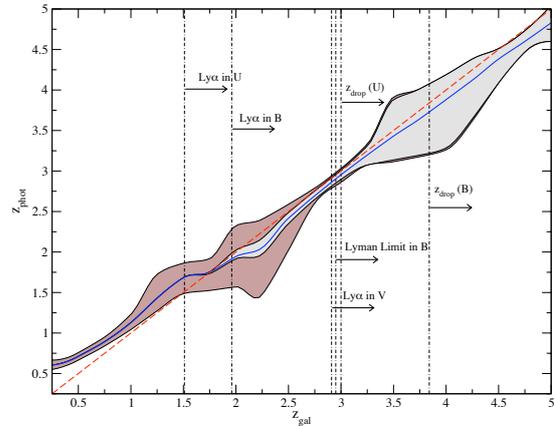
**Fig. 7.10:** Photometric estimates obtained from a sample of mock SEDs, which were computed in Johnson’s  $U$ ,  $B$ ,  $V$ , and  $R$  bands for an sd-type model galaxy spectrum affected by intergalactic attenuation using the MMC model, with (b) and without (a) including the magnitude offset  $m_{off}$ . The best estimate  $BE(z_{phot})$  for the photometric redshift in each case is calculated with a  $\chi^2$  minimisation technique by comparing the sample of mock SEDs to a grid of model SEDs, computed using the same input spectrum affected by intergalactic absorption following [Madau \(1995\)](#). At each redshift, the mock sample consists of  $4 \cdot 10^3$  SEDs and of an equal number of  $BE(z_{phot})$ ’s. The red-dashed line indicates a perfect match between  $z_{gal}$  and  $z_{phot}$ . The blue-solid line gives for each  $z_{gal}$  the median value  $\mu(z_{phot})$  of the ensemble of  $BE(z_{phot})$ ’s. The gray-shaded and brown-shaded areas correspond to the  $\pm 1 \sigma_q(z_{phot})$  and  $\pm 1 \sigma_m(z_{phot})$  ranges, respectively (see text for details). For reference, the vertical dashed lines denote the redshifts at which various absorption features ( $Ly\alpha$ , Lyman limit) are shifted into the indicated passband. The drop-out redshift indicates the redshift at which the Lyman limit is moved across the effective wavelength of the given passband.

tomeric redshifts for different galaxies with a priori unknown types. We illustrate this by comparing mock SEDs of an Sd-type input spectrum to a grid of model SEDs computed using a Sa-type spectrum. This approach should give a hint on the uncertainties in the estimates of photometric redshifts due to the confusion of different galaxy types, which are in general not known, especially at high redshift. The result is shown in Fig. 7.11. In general terms, it can be seen that the photometric redshifts dramatically deviate from the actual redshift at  $z_{gal} \leq 1.5$ , with photometric redshift in general being overestimated, as seen on the median and the scatter as given by  $\sigma_m(z_{phot})$ . Note, however, that  $\sigma_{BE}(z_{phot}) = 0$  in this range, which is consistent with the fact that the intergalactic attenuation does not have any effect at  $z_{gal} < 1.5$ . At  $z_{gal} > 2$ , the behaviour of the estimated redshifts is qualitatively the same as in Fig. 7.10, even though the scatter as given by  $\sigma_m(z_{phot})$  is larger, especially at  $2 < z_{gal} < 2.75$ . In this range and also at higher redshifts, photometric redshifts are in general underestimated. This scatter at this redshifts is again introduced by the magnitude off-set due to the fact that typical Sa and Sd galaxies have very different masses. Note that at  $z_{gal} > 3$ , the scatter as given by  $\sigma_m(z_{phot})$  is comparable to the scatter as given by  $\sigma_{BE}(z_{phot})$ . The reason for this is that the colors of galaxies at these redshifts are very similar, irrespective of their type, due to the strong cumulative intergalactic absorption, and also because at those redshifts all galaxies had stronger star formation and younger stellar populations

## 7.5 SUMMARY & CONCLUSIONS

We have presented a model that accounts for the intergalactic absorption due to neutral hydrogen in the form of Ly $\alpha$  forest and Lyman limit systems, based on the approach by Bershady et al. (1999), complemented with the input distributions for the properties of Ly $\alpha$  absorbers used by Madau (1995). This combination was suggested by our comparison of different input distribution functions of the absorbers properties with respect to their predictions for the redshift evolution of the cosmic flux decrement reported in a previous work.

We showed that the scatter in the magnitudes of a galaxy in a given filter due to the random distribution of intergalactic neutral hydrogen along the line-of-sight is rather small for redshifts lower than the corresponding drop-out redshift  $z_{drop}$  of a filter, and can be signifi-



**Fig. 7.11:** Photometric redshifts for a sample of mock SEDs computed in Johnson’s  $U$ ,  $B$ ,  $V$ , and  $R$  for an Sd-type galaxy, compared to a grid of model SEDs computed in the same filter system for an Sa-type galaxy. Lines and colored areas as in Fig. 7.10.

cantly larger close to and above this redshift. However, galaxy luminosities and magnitudes can no longer be accurately determined for  $z > z_{drop}$ , and hence their scatter need no longer be consider in that range.

We apply our model in order to asses the impact of the stochastic attenuation on the estimates of photometric redshifts, and find that photometric redshifts are in general underestimated, even though the effect is small. However, we find that the uncertainty due to the stochastic intergalactic attenuation is amplified due to the confusion of the different types of galaxies with different masses, present in deep surveys.

We conclude that the stochastic nature of the H I distribution in the IGM has no observable impact on the colors and luminosities of galaxies, but may have a non-negligible effect on the estimates of photometric redshifts. Furthermore, the mean intergalactic attenuation is not negligible at all, as previously and convincingly shown by Madau (1995) and Bershady et al. (1999), and it should be properly taken into account as presented here. However, the evolution of the mean magnitudes and colors of normal galaxies as computed by Bicker et al. (2004) and which include the correction for the mean attenuation, are not affected by the stochastic absorption in the IGM.

An extensive and detailed comparison of photometric redshifts obtained with our methods with observed spectroscopic redshifts is the next step in this investigation.

---

## Summary & Outlook

---

The formation and evolution of galaxies is one of the most active fields in Astrophysics to date. Important questions to be addressed concern the mechanisms behind the formation of different types of galaxies, and their individual properties such as stellar populations, morphologies, chemical abundances, all of which determine their spectro-photometric properties.

For galaxies in the nearby Universe, spectra can be obtained, and their global morphologies and structures within and associated to the galaxy (such as star clusters, star-forming regions, etc.) can be individually resolved, with which a deep insight into the on-going physical processes can be gained. Also, a wealth of information about their past evolution, *i.e.* star formation histories, interactions, etc., can be obtained by the analysis of their Spectral Energy Distributions (SEDs).

Even though spectra can also be taken for some of the brightest galaxies at higher redshifts, this few objects are truly not representative of the mean galactic population at those redshifts. Hence, in general, the study of the properties of these galaxies is limited to the analysis of their observed SEDs.

However, the observed SEDs substantially differ from their intrinsic SEDs because of various effects, such as the evolutionary effect, the cosmological effect, and the intergalactic attenuation. Any model describing the redshift evolution of the galaxies should account for all these effects. It turns out that the only tool that allows for the consistent description of the evolution of galaxy properties – in particular of their colors and luminosities – including evolutionary and cosmological

effects as well as the intergalactic attenuation, are the Evolutionary Synthesis models, such as GALEV.

This particular model, as many others, accounts for the intergalactic attenuation in a more or less standard way, *i.e.* by including only the *mean* absorption in the intergalactic medium due to neutral hydrogen. However, the distribution of intergalactic neutral hydrogen (HI) along the line-of-sight is random in nature, and deviations from a mean absorption are expected. Correspondingly, a scatter in the colors and luminosities observed along different line-of-sight should be observed as a consequence of this stochasticity. Hence, it is necessary to include in the GALEV code a correction for the variations in the intergalactic attenuation, in order to correctly interpret observations of galaxies, especially at high-redshift, where the attenuation is stronger.

The accurate description of the intergalactic attenuation, in particular of its variation along random lines-of-sight, requires a detailed knowledge of the distribution and physical state of the intergalactic HI along the line-of-sight. In other words, the evolution with redshift of the number density, of the column densities, and of the temperatures – as given by the Doppler parameter – of the intervening (Ly $\alpha$ ) absorbing systems embedded in the intergalactic medium (IGM) is needed.

These properties can be determined, for example, from the analysis of the absorption lines imprinted in the spectra of background sources which are associated to the Ly $\alpha$  absorbers. The usual method consists in looking for the best set of values for the column density and Doppler parameter that best match the observed

profile. In the case of the absorption lines found on QSO spectra and attributed to Ly $\alpha$  absorbers, the absorption line profiles are best modeled by Voigt profiles, which are mathematically given in terms of the Voigt-Hjerting function.

In this work, I presented a simple analytic approximation to the Voigt-Hjerting function with which Voigt profiles can be modeled with high accuracy for an arbitrary range in wavelength (or frequency), and for the range of column densities and Doppler parameters characteristic to intergalactic neutral hydrogen. This approach offers a great compromise between speed, accuracy, and flexibility in its implementation. Furthermore, it is not limited to intergalactic neutral hydrogen, but is valid to model absorption Voigt profiles found in any type of spectrum (stellar, solar, etc.), which arise in a medium whose damping parameter and column density fall in the range characteristic to intergalactic H $\text{I}$ . This work is certainly an invaluable contribution to the improvement of the analysis of QSO absorption lines, and consequently to our knowledge of the evolution of the properties of the Ly $\alpha$  absorbers.

The properties of these absorbers as obtained from observations are usually parametrized by a set of distribution functions which describe the evolution with redshift of their average number, column densities, and Doppler parameters. However, all these parameters are truly random variables and thus their values are in general different for different absorbers and along different lines-of-sight. Due to the complex dependence of the absorption along a particular line-of-sight on all this parameters, the only possible method to describe it correctly is the Monte Carlo approach. The procedure in order to construct a model that accounts for the intergalactic absorption along random line-of-sight using this approach was developed and presented with great detail in this work.

It is clear that, independently of the approach followed, the intergalactic absorption due to intergalactic H $\text{I}$  cannot be properly modelled without a consistent set of input distributions for the physical parameters of the absorbers constrained from observations. In order to discriminate between two widely used sets of such distributions found in the literature, we used our model to compute the redshift evolution of the cosmic flux decrement  $D_A$ , and found that the Monte Carlo method in combination with a particular set of distribution functions very accurately matches the measurements of the cosmic flux decrement in the redshift range  $0.2 < z < 5.41$ . The measurements of  $D_A$  in the range

$2.71 < z < 5.41$  were actually performed by ourselves, with which we expanded the redshift range of measurements available so far in the literature.

Moreover, we derive from first principles that the distribution of values of the cosmic flux decrement at a given redshift is well described by a lognormal distribution at low redshifts and even better by a Gaussian distribution at high redshifts, a result which was confirmed by numerical simulations of the evolution of the  $D_A$  performed with our model. This result implies that at redshifts where  $D_A$  is distributed lognormally, the effective optical depth of the intergalactic H $\text{I}$  obeys a normal distribution. The result that  $D_A$  is distributed lognormally leads to the question of whether this fact is linked in some way to the assumption that the primordial density fluctuations that gave origin to the structure formation in the Universe, and that ultimately gave rise to the Ly $\alpha$  absorbers, obey a lognormal distribution as well. In the case of an affirmative answer, this result would provide a great insight into the mechanism responsible for the formation of structure in the Universe.

Once we identified a suitable set of distributions for the physical properties, we use our model for the intergalactic absorption due to neutral hydrogen in the form of Ly $\alpha$  forest and Lyman limit systems to compute the variations in the luminosities and colors of galaxies in the redshift range  $0 < z < 5$ , and found that the effect is too small to be observable. We also explored for the first time the impact of the stochastic attenuation on the estimates of photometric redshifts, and find that photometric redshifts are in general slightly underestimated, even though the effect is small. However, we find that the uncertainty due to the stochastic intergalactic attenuation is amplified due to the confusion of the different types of galaxies with different masses, as is the case of the galaxies observed in deep surveys.

It turns out in the light of the above results that the scatter in the colors and luminosities of galaxies observed in Deep Fields cannot be accounted for by the stochastic absorption of their light due to the presence of randomly distributed intergalactic neutral hydrogen along their lines-of-sight. Nevertheless, the mean attenuation does have a significant impact on the colors and luminosities, and the only correct method to account for this effect is the Monte Carlo approach in combination with a suitable set of input distribution for the properties of the Ly $\alpha$  absorbers, as described in this work.

Moreover, the mild effect of the stochastic attenuation on the photometric redshifts found here may be observable with current facilities. However, our approach

to estimate photometric redshifts has to be tested first against reliable spectroscopic redshifts. Once the reliability of our method is established, we may use our approach to analyse large sample of multi-wavelength observations of galaxies in order to test whether the stochastic attenuation does affect the photometric redshifts as predicted by our model. This is, however, beyond the scope of this work and will be the starting point of future work.



---

## Bibliography

---

- Abramowitz, M. & Stegun, I. A. 1965, Handbook of mathematical functions with formulas, graphs, and mathematical tables (Dover Books on Advanced Mathematics, New York: Dover, —c1965, Corrected edition, edited by Abramowitz, Milton; Stegun, Irene A.)
- Anders, P., Bissantz, N., Fritze-v. Alvensleben, U., & de Grijs, R. 2004, MNRAS, 347, 196
- Arimoto, N. & Jablonka, P. 1991, A&A, 249, 374
- Arimoto, N. & Yoshii, Y. 1986, A&A, 164, 260
- Atwood, B., Baldwin, J. A., & Carswell, R. F. 1985, ApJ, 292, 58
- Bahcall, J. N. & Peebles, P. J. E. 1969, ApJ, 156, L7+
- Bahcall, J. N. & Salpeter, E. E. 1965, ApJ, 142, 1677
- Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, ApJ, 327, 570
- Barkana, R. & Loeb, A. 2007, Reports of Progress in Physics, 70, 627
- Baron, E., Carswell, R. F., Hogan, C. J., & Weymann, R. J. 1989, ApJ, 337, 609
- Bechtold, J., Weymann, R. J., Lin, Z., & Malkan, M. A. 1987, ApJ, 315, 180
- Benítez, N. 2000, ApJ, 536, 571
- Bernardi, M., Sheth, R. K., SubbaRao, M., et al. 2003, AJ, 125, 32
- Bershady, M. A., Charlton, J. C., & Geoffroy, J. M. 1999, ApJ, 518, 103
- Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275
- Bi, H. G., Boerner, G., & Chu, Y. 1992, A&A, 266, 1
- Bianchi, S., Cristiani, S., & Kim, T.-S. 2001, A&A, 376, 1
- Bicker, J., Fritze-v. Alvensleben, U., Möller, C. S., & Fricke, K. J. 2004, A&A, 413, 37

- Bruzual A., G. & Charlot, S. 1993, *ApJ*, 405, 538
- Burbidge, E. M., Lynds, C. R., & Stockton, A. N. 1968, *ApJ*, 152, 1077
- Carroll, S. M. 2001, *Living Reviews in Relativity*, 4, 1
- Carswell, R. F., Morton, D. C., Smith, M. G., et al. 1984, *ApJ*, 278, 486
- Carswell, R. F. & Rees, M. J. 1987, *MNRAS*, 224, 13P
- Carswell, R. F., Webb, J. K., Baldwin, J. A., & Atwood, B. 1987, *ApJ*, 319, 709
- Charlton, J., Churchill, C., & Murdin, P. 2000, *Encyclopedia of Astronomy and Astrophysics*
- Coc, A., Vangioni-Flam, E., Descouvemont, P., Adahchour, A., & Angulo, C. 2004, *ApJ*, 600, 544
- Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, *AJ*, 132, 926
- Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., & Savaglio, S. 1995, *MNRAS*, 273, 1016
- Cristiani, S. & D'Odorico, V. 2000, *AJ*, 120, 1648
- Cristiani, S., Giallongo, E., Buson, L. M., Gouiffes, C., & La Franca, F. 1993, *A&A*, 268, 86
- Davé, R., Hernquist, L., Weinberg, D. H., & Katz, N. 1997, *ApJ*, 477, 21
- Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, *ApJ*, 253, 423
- Dawson, H. G. 1898, *London Mathematical Society Proceedings*, 29, 519
- de Bernardis, P., Ade, P. A. R., Bock, J. J., et al. 2000, *Nature*, 404, 955
- Delgado, R. M. G., Cerviño, M., Martins, L. P., Leitherer, C., & Hauschildt, P. H. 2005, *MNRAS*, 357, 945
- Demiański, M., Doroshkevich, A. G., & Turchaninov, V. 2000, *MNRAS*, 318, 1189
- Desjacques, V., Nusser, A., & Sheth, R. K. 2007, *MNRAS*, 374, 206
- Dobrzycki, A. & Bechtold, J. 1991, *ApJ*, 377, L69
- Duncan, R. C., Ostriker, J. P., & Bajtlik, S. 1989, *ApJ*, 345, 39
- Fall, S. M. & Pei, Y. C. 1989, *ApJ*, 337, 7
- Fan, X., Narayanan, V. K., Strauss, M. A., et al. 2002, *AJ*, 123, 1247
- Finn, G. D. & Mugglestone, D. 1965, *MNRAS*, 129, 221
- Fioc, M. & Rocca-Volmerange, B. 1997, *A&A*, 326, 950
- Fontana, A. & Ballester, P. 1995, *The ESO Messenger*, 80, 37
- Fontana, A., D'Odorico, S., Poli, F., et al. 2000, *AJ*, 120, 2206
- Forster, O. 1983, *Analysis*, 4th edn., Vol. 1 (Vieweg Verlag, Braunschweig)
- Fritze-v. Alvensleben, U. 1999, in *Astronomical Society of the Pacific Conference Series*, Vol. 192, *Spectrophotometric Dating of Stars and Galaxies*, ed. I. Hubeny, S. Heap, & R. Cornett, 273–+
- Fritze-v. Alvensleben, U. & Bicker, J. 2006, *A&A*, 454, 67

- Giallongo, E., Cristiani, S., & Trevese, D. 1992, *ApJ*, 398, L9
- Giallongo, E., D’Odorico, S., Fontana, A., et al. 1994, *ApJ*, 425, L1
- Giallongo, E., Gratton, R., & Trevese, D. 1990, *MNRAS*, 244, 450
- Greenstein, J. L. & Schmidt, M. 1967, *ApJ*, 148, L13+
- Gunn, J. E. & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Guth, A. H. 1981, *Phys. Rev. D*, 23, 347
- Harris, III, D. L. 1948, *ApJ*, 108, 112
- Hjerting, F. 1938, *ApJ*, 88, 508
- Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, *ArXiv Astrophysics e-prints*
- Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A., & Rauch, M. 1995, *AJ*, 110, 1526
- Hui, L. & Rutledge, R. E. 1999, *ApJ*, 517, 541
- Humlíček, J. 1982, *J. Quant. Spec. Radiat. Transf.*, 27, 437
- Izotov, Y. I., Thuan, T. X., & Stasińska, G. 2007, *ApJ*, 662, 15
- Janknecht, E., Baade, R., & Reimers, D. 2002, *A&A*, 391, L11
- Jenkins, E. B. & Ostriker, J. P. 1991, *ApJ*, 376, 33
- Karzas, W. J. & Latter, R. 1961, *ApJS*, 6, 167
- Kielkopf, J. F. 1973, *Journal of the Optical Society of America (1917-1983)*, 63, 987
- Kim, T.-S., Carswell, R. F., Cristiani, S., D’Odorico, S., & Giallongo, E. 2002a, *MNRAS*, 335, 555
- Kim, T.-S., Cristiani, S., & D’Odorico, S. 2001, *A&A*, 373, 757
- Kim, T.-S., Cristiani, S., & D’Odorico, S. 2002b, *A&A*, 383, 747
- Kim, T.-S., Hu, E. M., Cowie, L. L., & Songaila, A. 1997, *AJ*, 114, 1
- Kirkman, D. & Tytler, D. 1997, *ApJ*, 484, 672
- Kirkman, D., Tytler, D., Lubin, D., & Charlton, J. 2007, *MNRAS*, 131
- Kirkman, D., Tytler, D., Suzuki, N., et al. 2005, *MNRAS*, 360, 1373
- Kirkman, D., Tytler, D., Suzuki, N., O’Meara, J. M., & Lubin, D. 2003, *ApJS*, 149, 1
- Lamb, D. Q. & Reichart, D. E. 2000, *ApJ*, 536, 1
- Lanzetta, K. M. 1991, *ApJ*, 375, 1
- Lanzetta, K. M., Fernández-Soto, A., & Yahil, A. 1998, in *The Hubble Deep Field*, ed. M. Livio, S. M. Fall, & P. Madau, 143–+
- Laor, A., Fiore, F., Elvis, M., Wilkes, B. J., & McDowell, J. C. 1997, *ApJ*, 477, 93
- Le Borgne, D., Rocca-Volmerange, B., Prugniel, P., et al. 2004, *A&A*, 425, 881

- Lehner, N., Savage, B. D., Richter, P., et al. 2006, ArXiv Astrophysics e-prints
- Lejeune, T., Cuisinier, F., & Buser, R. 1997, A&AS, 125, 229
- Lejeune, T., Cuisinier, F., & Buser, R. 1998, A&AS, 130, 65
- Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M. 1996, ApJ, 472, 509
- Lynds, R. 1971, ApJ, 164, L73+
- Madau, P. 1991, ApJ, 376, L33
- Madau, P. 1992, ApJ, 389, L1
- Madau, P. 1995, ApJ, 441, 18
- Madau, P. 2000, ArXiv Astrophysics e-prints
- Maraston, C. 2005, MNRAS, 362, 799
- McDonald, P., Seljak, U., Cen, R., Bode, P., & Ostriker, J. P. 2005, MNRAS, 360, 1471
- Meiksin, A. 2006, MNRAS, 365, 807
- Meiksin, A. & Madau, P. 1993, ApJ, 412, 34
- Meiksin, A. & White, M. 2004, MNRAS, 350, 1107
- Melott, A. L. 1980, ApJ, 241, 889
- Mihalas, D. 1970, Stellar atmospheres (Series of Books in Astronomy and Astrophysics, San Francisco: Freeman, —c1970)
- Miller, W. L. & Gordon, A. R. 1931, J. Phys. Chem., 35, 2785
- Miralda-Escude, J. & Ostriker, J. P. 1990, ApJ, 350, 1
- Mollá, M. 2007, ArXiv Astrophysics e-prints
- Möller, C. S., Fritze-v. Alvensleben, U., & Fricke, K. J. 1999, in Astronomische Gesellschaft Meeting Abstracts, ed. R. E. Schielicke, 14–+
- Møller, P. & Jakobsen, P. 1990, A&A, 228, 299
- Monaghan, J. J. 1971, MNRAS, 152, 509
- Moro, D. & Munari, U. 2000, A&AS, 147, 361
- Morton, D. C. 2003, ApJS, 149, 205
- Muecket, J. P., Petitjean, P., Kates, R. E., & Riediger, R. 1996, A&A, 308, 17
- Murdoch, H. S., Hunstead, R. W., Pettini, M., & Blades, J. C. 1986, ApJ, 309, 19
- Murphy, M. T. 2002, Ph.D. Thesis
- Neugebauer, G., Oke, J. B., Becklin, E. E., & Matthews, K. 1979, ApJ, 230, 79
- O'Brien, P. T., Wilson, R., & Gondhalekar, P. M. 1988, MNRAS, 233, 801

- Oke, J. B. 1974, *ApJS*, 27, 21
- Oke, J. B. & Gunn, J. E. 1983, *ApJ*, 266, 713
- Oke, J. B. & Korycansky, D. G. 1982, *ApJ*, 255, 11
- Ostriker, J. P., Bajtlik, S., & Duncan, R. C. 1988, *ApJ*, 327, L35
- Ostriker, J. P. & Ikeuchi, S. 1983, *ApJ*, 268, L63
- Ostriker, J. P. & Steinhardt, P. J. 1995, *ArXiv Astrophysics e-prints*
- Palenius, H. P., Kohl, J. L., & Parkinson, W. H. 1976, *Phys. Rev. A*, 13, 1805
- Peimbert, M., Luridiana, V., & Peimbert, A. 2007, *ArXiv Astrophysics e-prints*
- Penton, S. V., Stocke, J. T., & Shull, J. M. 2004, *ApJS*, 152, 29
- Percival, W. J., Nichol, R. C., Eisenstein, D. J., et al. 2007, *ApJ*, 657, 51
- Perlmutter, S., Turner, M. S., & White, M. 1999, *Physical Review Letters*, 83, 670
- Peterson, B. A. 1978, in *IAU Symp. 79: Large Scale Structures in the Universe*, ed. M. S. Longair & J. Einasto, 389–392
- Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F., & Lanzetta, K. 1993, *MNRAS*, 262, 499
- Pettini, M., Hunstead, R. W., Smith, L. J., & Mar, D. P. 1990, *MNRAS*, 246, 545
- Pierre, M., Shaver, P. A., & Iovino, A. 1988, *A&A*, 197, L3
- Press, W. H. & Rybicki, G. B. 1993, *ApJ*, 418, 585
- Press, W. H., Rybicki, G. B., & Schneider, D. P. 1993, *ApJ*, 414, 64
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in Fortran 77*, 2nd edn. (Cambridge University Press)
- Rao, S. M. 2005, in *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, ed. P. Williams, C.-G. Shu, & B. Menard, 125–147
- Rauch, M. 1998, *ARA&A*, 36, 267
- Rauch, M., Carswell, R. F., Chaffee, F. H., et al. 1992, *ApJ*, 390, 387
- Rauch, M. & Haehnelt, M. G. 1995, *MNRAS*, 275, L76
- Rauch, M., Miralda-Escude, J., Sargent, W. L. W., et al. 1997, *ApJ*, 489, 7
- Richter, P., Fang, T., & Bryan, G. L. 2006, *A&A*, 451, 767
- Riediger, R., Petitjean, P., & Mucket, J. P. 1998, *A&A*, 329, 30
- Sandage, A. 1961, *ApJ*, 134, 916
- Sandage, A. 1986, *A&A*, 161, 89
- Sargent, W. L. W., Steidel, C. C., & Boksenberg, A. 1989, *ApJS*, 69, 703
- Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, *ApJS*, 42, 41

- Sawicki, M. J., Lin, H., & Yee, H. K. C. 1997, *AJ*, 113, 1
- Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W. 2000, *MNRAS*, 318, 817
- Schneider, D. P., Schmidt, M., & Gunn, J. E. 1989, *AJ*, 98, 1951
- Schneider, D. P., Schmidt, M., & Gunn, J. E. 1991, *AJ*, 101, 2004
- Schulz, J., Fritze-v. Alvensleben, U., Möller, C. S., & Fricke, K. J. 2002, *A&A*, 392, 1
- Shectman, S. A., Landy, S. D., Oemler, A., et al. 1996, *ApJ*, 470, 172
- Spergel, D. N., Bean, R., & Doré, O. e. a. 2006, *ArXiv Astrophysics e-prints*
- Spitzer, L. 1978, *Physical processes in the interstellar medium* (New York Wiley-Interscience, 1978. 333 p.)
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629
- Steidel, C. C. & Sargent, W. L. W. 1987a, *ApJ*, 318, L11
- Steidel, C. C. & Sargent, W. L. W. 1987b, *ApJ*, 313, 171
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, *ApJ*, 565, 773
- Tepper-García, T. 2006, *MNRAS*, 369, 2025
- Tepper-García, T. & Fritze, U. 2007, *ArXiv e-prints*, 705
- Tytler, D. 1987, *ApJ*, 321, 49
- Tytler, D., Kirkman, D., O'Meara, J. M., et al. 2004a, *ApJ*, 617, 1
- Tytler, D., O'Meara, J. M., Suzuki, N., et al. 2004b, *AJ*, 128, 1058
- Umemura, M. & Ikeuchi, S. 1985, *ApJ*, 299, 583
- Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, *AJ*, 122, 549
- Vázquez, G. A., Carigi, L., & González, J. J. 2003, *A&A*, 400, 31
- Walshaw, C. D. 1955, *Proceedings of the Physical Society A*, 68, 530
- Weinberg, D. H., Davé, R., Katz, N., & Kollmeier, J. A. 2003, in *American Institute of Physics Conference Series*, Vol. 666, *The Emergence of Cosmic Structure*, ed. S. H. Holt & C. S. Reynolds, 157–169
- Weymann, R. J., Carswell, R. F., & Smith, M. G. 1981, *ARA&A*, 19, 41
- Weymann, R. J., Jannuzi, B. T., Lu, L., et al. 1998, *ApJ*, 506, 1
- Whiting, E. E. 1968, *J. Quant. Spec. Radiat. Transf.*, 8, 1379
- Williams, P., Shu, C.-G., & Ménard, B., eds. 2005, *Probing Galaxies through Quasar Absorption Lines*
- Williger, G. M., Heap, S. R., Weymann, R. J., et al. 2006, *ApJ*, 636, 631
- Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, *ARA&A*, 43, 861
- Yee, H. K. C. 1998, *ArXiv Astrophysics e-prints*
- Young, P., Sargent, W. L. W., & Bokkenberg, A. 1982, *ApJS*, 48, 455

Zaghloul, M. R. 2007, MNRAS, 16

Zhang, Y., Anninos, P., Norman, M. L., & Meiksin, A. 1997, ApJ, 485, 496

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 475, 469

Zuo, L. 1993, A&A, 278, 343

Zuo, L. & Lu, L. 1993, ApJ, 418, 601

## APPENDIX **A**

---

### The Lyman series

---

**Table A.1:** Values of the central wavelength, oscillatory strength and damping constant for the first 24 Lyman transitions.

Transition	Central Wavelength [ nm]	$f_{\text{osc}}$	$\Gamma[\text{s}^{-1}]$
H $1\alpha$	121.56701	0.416400	$6.265 \cdot 10^8$
H $1\beta$	102.57223	0.079120	$1.897 \cdot 10^8$
H $1\gamma$	97.25368	0.029000	$8.126 \cdot 10^7$
H $1\delta$	94.97431	0.013940	$7.640 \cdot 10^7$
H $1\epsilon$	93.78035	0.007799	$4.423 \cdot 10^7$
H $1\zeta$	93.07483	0.004814	$1.236 \cdot 10^7$
H $1\eta$	92.62257	0.003183	$8.249 \cdot 10^6$
H $1\theta$	92.31504	0.002216	$5.782 \cdot 10^6$
H $1\iota$	92.09631	0.001605	$4.208 \cdot 10^6$
H $1\kappa$	91.93514	0.00120	$3.158 \cdot 10^6$
H $1\zeta$	93.07483	0.004814	$1.236 \cdot 10^7$
H $1\eta$	92.62257	0.003183	$8.249 \cdot 10^6$
H $1\theta$	92.31504	0.002216	$5.782 \cdot 10^6$
H $1\iota$	92.09631	0.001605	$4.208 \cdot 10^6$
H $1\kappa$	91.93514	0.00120	$3.158 \cdot 10^6$
H $1\lambda$	91.81294	0.000921	$2.430 \cdot 10^6$
H $1\mu$	91.71806	0.000722	$1.910 \cdot 10^6$
H $1\nu$	91.6429	0.000577	$1.529 \cdot 10^6$
H $1\xi$	91.5824	0.000469	$1.243 \cdot 10^6$
H $1\omicron$	91.5329	0.000386	$1.024 \cdot 10^6$
H $1\pi$	91.4919	0.000321	$8.533 \cdot 10^5$
H $1\rho$	91.4576	0.000270	$7.186 \cdot 10^5$
H $1\sigma$	91.4286	0.000230	$6.109 \cdot 10^5$
H $1\tau$	91.4039	0.000197	$5.237 \cdot 10^5$
H $1\nu$	91.3826	0.000170	$4.523 \cdot 10^5$
H $1\phi$	91.3641	0.000148	$3.933 \cdot 10^5$
H $1\chi$	91.3480	0.000129	$3.443 \cdot 10^5$
H $1\psi$	91.3339	0.000114	$3.030 \cdot 10^5$
H $1\omega$	91.3215	0.000101	$2.679 \cdot 10^5$

---

## The Gauß approximation to the Voigt-Hjerting function

---

The Laplace transform  $L_f[s]$  of a function  $f(t)$  is defined as

$$L_f[s] \equiv \int_0^{\infty} f(t) e^{-st} dt, \quad s > 0. \quad (\text{B.1})$$

For example, if für  $f(t) = \cos(bt)$ ,  $b \in \mathbb{R}$ , then

$$L_{\cos}[s] = \int_0^{\infty} \cos(bt) e^{-st} dt = \frac{s}{b^2 + s^2}. \quad (\text{B.2})$$

It follows from this, that the Voigt-Hjerting function

$$H(a, x) = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(x-y)^2 + a^2} dy \quad x, a \in \mathbb{R} \quad (\text{B.3})$$

can be rewritten as ([Mihalas 1970](#))

$$\begin{aligned} H(a, x) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-y^2} \frac{a}{(x-y)^2 + a^2} dy \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-y^2} \int_0^{\infty} e^{-av} \cos((x-y)v) dv dy \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-y^2} \int_0^{\infty} e^{-av} \cos(xv) \cos(yv) dv dy \\ &= \frac{1}{\pi} \int_0^{\infty} e^{-av} \cos(xv) \int_{-\infty}^{+\infty} e^{-y^2} \cos(yv) dy dv. \end{aligned} \quad (\text{B.4})$$

Furthermore, it is true that

$$\int_{-\infty}^{+\infty} e^{-y^2} \cos(yv) dy = \sqrt{\pi} e^{-(v/2)^2}. \quad (\text{B.5})$$

From equations (B.4) and (B.5) we get

$$H(a, x) = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-av} \cos(xv) e^{-(v/2)^2} dv. \quad (\text{B.6})$$

If  $a < 1$ , then the term  $e^{-av}$  in equation (B.6) may be written in terms of the power series (Mihalas 1970)

$$e^{-av} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (av)^n \quad (\text{B.7})$$

It follows then

$$H(a, x) = \frac{1}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} a^n \int_0^{\infty} v^n e^{-(v/2)^2} \cos(xv) dv, \quad (\text{B.8})$$

and hence

$$H(a, x) = \sum_{n=0}^{\infty} H_n(x) a^n, \quad (\text{B.9})$$

where the functions  $H_n(x)$  are given by

$$H_n(x) \equiv \frac{(-1)^n}{\sqrt{\pi} n!} \int_0^{\infty} v^n e^{-(v/2)^2} \cos(xv) dv. \quad (\text{B.10})$$

We now want to analyse the behaviour of the functions  $H_n(x)$  for all  $n$  and  $v$ . In particular, we want to know if this functions are bounded. To this end, we take a look at their absolute value  $|H_n(x)|$ . First note that

$$\left| \int_I f(x) dx \right| \leq \int_I |f(x)| dx \quad (\text{B.11})$$

for any function  $f(x)$  and arbitrary integration interval  $I$ . It follows from this that

$$\begin{aligned} |H_n(x)| &= \left| \frac{(-1)^n}{\sqrt{\pi} n!} \int_0^{\infty} v^n e^{-(v/2)^2} \cos(xv) dv \right| \\ &\leq \frac{1}{\sqrt{\pi} n!} \left| \int_0^{\infty} v^n e^{-(v/2)^2} \cos(xv) dv \right| \\ &\leq \frac{1}{\sqrt{\pi} n!} \int_0^{\infty} |v^n e^{-(v/2)^2} \cos(xv)| dv. \end{aligned} \quad (\text{B.12})$$

Since  $e^{-(v/2)^2} > 0 \forall v \in \mathbb{R}$  and the integral is computed over  $(0, \infty)$ , it follows that  $|v^n e^{-(v/2)^2} \cos(xv)| = v^n e^{-(v/2)^2} |\cos(xv)|$ ,  $\forall n \in \mathbb{N}_0$ , and  $x, v \in \mathbb{R}$ . Furthermore,  $|\cos(xv)| \leq 1 \forall x, v \in \mathbb{R}$ , and thus

$$\begin{aligned} |H_n(x)| &\leq \frac{1}{\sqrt{\pi} n!} \int_0^{\infty} v^n e^{-(v/2)^2} dv \\ &= \frac{2^{n+1}}{\sqrt{\pi} n!} \int_0^{\infty} v^n e^{-v^2} dv, \end{aligned} \quad (\text{B.13})$$

where  $v$  was replaced by  $v/2$  in the last step. Bearing in mind that

$$\int_0^{\infty} v^n e^{-v^2} dv = \begin{cases} \frac{\sqrt{\pi} 1 \cdot 3 \cdots (2k-1)}{2} & n = 2k \\ \frac{1}{2} k! & n = 2k + 1 \end{cases},$$

it follows, using equation (B.13) that

$$\begin{aligned} |H_n(x)| &\leq \frac{1}{\sqrt{\pi} n!} \int_0^\infty v^n e^{-(v/2)^2} dv \\ &= \frac{2^{n+1}}{\sqrt{\pi} n!} \cdot \begin{cases} \frac{\sqrt{\pi}}{2} \frac{1 \cdot 3 \cdots (2k-1)}{2^k} & n = 2k \\ \frac{1}{2} k! & n = 2k + 1 \end{cases}, \end{aligned}$$

and hence

$$|H_n(x)| \leq \begin{cases} \frac{1}{k!} & n = 2k \\ \frac{1}{\sqrt{\pi}} \frac{2^{2k+1}}{(2k+1)(2k) \cdots (k+1)} & n = 2k + 1 \end{cases} \quad (\text{B.14})$$

Since  $k \in \{0, 1, 2, \dots\}$ , it follows immediately that

$$|H_n(x)| \leq \frac{2}{\sqrt{\pi}} \approx 1.123 \quad \forall \quad n \in \mathbb{N}_0, \quad x \in \mathbb{R}. \quad (\text{B.15})$$

The functions  $H_n(x)$  are thus bounded to values of the order of unity. It follows then that for values of  $a$  small compared to unity, ie  $a \ll 1$ , the VHF in the form of equation (B.9) can be approximated to zeroth order in  $a$  by the first term of the series, *i.e.*

$$H(a, x) \approx H_0(x), \quad a \ll 1, \quad x \in \mathbb{R}, \quad (\text{B.16})$$

Note that this result is exact in the limit  $a \rightarrow 0$ . Using equation (B.5) it is straightforward that

$$H_0(x) = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-(v/2)^2} \cos(xv) dv = e^{-x^2}, \quad (\text{B.17})$$

and from this, we get that

$$H(a, x) \approx e^{-x^2}, \quad a \ll 1, \quad x \in \mathbb{R}. \quad (\text{B.18})$$

---

## Dependence of the density on the scale factor

---

The Friedmann equations (4.5) and (4.6), now with a generalised density, can be rewritten as

$$\dot{a}^2 = \frac{1}{3}(8\pi G\rho + \Lambda)a^2 - kc^2 \quad (\text{C.1})$$

$$3\ddot{a} = \Lambda a - 4\pi G\left(\rho + 3\frac{p}{c^2}\right)a \quad (\text{C.2})$$

Differentiating equation (C.1) with respect to  $t$  gives

$$2\dot{a}\ddot{a} = \frac{2}{3}(8\pi G\rho + \Lambda)a\dot{a} + \frac{1}{3}8\pi G\dot{\rho}a^2 \quad (\text{C.3})$$

Multiplying equation (C.2) by  $-2/3\dot{a}$  gives

$$-2\dot{a}\ddot{a} = -\frac{2}{3}\Lambda a\dot{a} + \frac{8}{3}\pi G\left(\rho + 3\frac{p}{c^2}\right)a\dot{a} \quad (\text{C.4})$$

Adding up these equations leads to

$$0 = \frac{2}{3}8\pi G\rho a\dot{a} + \frac{1}{3}8\pi G\dot{\rho}a^2 + \frac{8}{3}\pi G\left(\rho + 3\frac{p}{c^2}\right)a\dot{a} \quad (\text{C.5})$$

Solving for  $\dot{\rho}$  leads finally to

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{1}{c^2}p\right) \quad (\text{C.6})$$

Assuming that the evolution of the density  $\rho_i$  for each component obeys an equation of the form of equation (C.6), and that the equation of state for each species can be expressed as  $p_i = w_i c^2 \rho_i$  (cf. equation 4.9), it follows that

$$\begin{aligned} \frac{\dot{\rho}_i}{\rho_i} &= -3(1+w_i)\frac{\dot{a}_i}{a_i} \\ \iff \frac{d\rho_i}{\rho_i} &= -3(1+w_i)\frac{da_i}{a_i} \\ \iff \ln\rho_i &= -3(1+w_i)\ln a_i + C \end{aligned}$$

and from this it is straightforward to see that

$$\rho_i \propto a^{-3(1+w_i)} \quad (\text{C.7})$$

For ordinary matter,  $p = 0$  implying  $w_m = 0$ , and hence, the matter density evolves proportional to  $a^{-3}$ , as expected. Relativistic matter, in particular radiation, obeys an equation of state with  $w_r = 1/3$ , and thus the radiation density evolves as  $a^{-4}$ . This is usually interpreted as being the result, first, of the density dilution due to expansion, which is proportional to  $a^{-3}$ , and of an additional decrease in the radiation energy, due to the cosmological redshift which is proportional to  $a^{-1}$ . These factors explain the  $a^{-4}$  dependence of the radiation–density evolution. Furthermore, the fact that the  $\Lambda$ –term in the Friedmann equation (4.5) is independent of  $a$  implies that  $w_\Lambda = -1$ , and that the ”curvature” density, which evolves as  $a^{-2}$ , obeys an equation of state with  $w_k = -1/3$ .

---

## The proper path length

---

To derive equation (4.22), first consider that the deceleration parameter (eq. 4.7) can be rewritten as

$$q = -H^2 (\dot{H} + H^2) = -\frac{\ddot{a}a}{\dot{a}^2} = -\frac{\ddot{a}}{aH^2}. \quad (\text{D.1})$$

On the other hand, the second Friedmann equation (4.6) including all forms of density reads

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G \sum_i (1 + 3w_i)\rho_i, \quad (\text{D.2})$$

where  $\sum_i \rho_i \equiv \rho$ , and  $p$  has been replaced using equation (4.9). Inserting the above equation into equation (D.1), and using equation (4.11) yields

$$q = \sum_i \frac{1 + 3w_i}{2} \Omega_i. \quad (\text{D.3})$$

Now, inserting this last expression in equation (4.21), and assuming  $\Omega_\Lambda \equiv 0$  it is straightforward that

$$\left| \frac{dl}{dz} \right| = \frac{c}{H_0} (1+z)^{-2} (1 + 2q_0 z)^{-\frac{1}{2}}. \quad (\text{D.4})$$

where the value at present time for all cosmological parameters has been assumed.

An alternative derivation of this last equation for  $k = +1$  is the following: The first ingredient is the Friedmann equation for vanishing pressure and cosmological constant

$$\left( \frac{\dot{a}}{a} \right)^2 + 2\frac{\ddot{a}}{a^2} = -\frac{k c^2}{a^2}, \quad (\text{D.5})$$

where  $a(t)$  is the scaling factor of the Universe and  $k/a^2$  is the Riemannian curvature. This equation can be solved in parametric form for a space with positive curvature ( $k = +1$ ) by the functions (Sandage 1961)

$$a(\theta; r) = r(1 - \cos \theta), \quad (\text{D.6})$$

$$t(\theta; r) = \frac{r}{c} (\theta - \sin \theta). \quad (\text{D.7})$$

These are the parametric equations of a cycloid, where  $r$  is the radius of the generating circle, and  $\theta$  is the angle through which the circle has rolled. A more physical interpretation of this quantity can be obtained by computing the radial path length of a photon, emitted at time  $t$  and observed at time  $t_0$ :

$$u = \int_t^{t_0} \frac{c \, dt}{a(t)}. \quad (\text{D.8})$$

Plugging equations (D.6) and (D.7) into the above equation and integrating gives  $u = \theta(t_0) - \theta(t)$ , which means that the quantity the angle  $\theta$  is actually the comoving proper path length, *i.e.*  $d\theta \equiv du$ .

Now, from the definition of the cosmological redshift,  $1 + z(t) \equiv a(t_0)/a(t)$  and the above equations it is straightforward to show that (Sandage 1961)

$$\cos \theta = \frac{z + \cos \theta_0}{1 + z}, \quad (\text{D.9})$$

where  $\theta_0 \equiv \theta(t_0)$ .

For the Hubble parameter (eq. 4.5) and the deceleration parameter (eq. 4.7) follows quite analogously that

$$H = \frac{c}{r} \frac{\sin \theta}{(1 - \cos \theta)^2}, \quad (\text{D.10})$$

and

$$q = \frac{1 - \cos \theta}{\sin^2 \theta}. \quad (\text{D.11})$$

The rate of change of the physical distance  $l$  as a function of redshift is given by

$$\frac{dl}{dz} \equiv a(t) \frac{du}{dz} = a(t) \frac{d\theta}{dz}. \quad (\text{D.12})$$

Solving equation (D.9) for  $\theta$  and taking the derivative with respect to  $z$  gives

$$\frac{d\theta}{dz} = - \frac{1 - \cos \theta_0}{(1 + z) [2z(1 - \cos \theta_0) + \sin^2 \theta_0]^{1/2}}. \quad (\text{D.13})$$

Now, using equations (D.6), (D.10), and (D.11) the last equation can be brought in the form

$$\left| \frac{d\theta}{dz} \right| = \frac{c}{H_0 a_0} (1 + z)^{-1} (1 + 2q_0 z)^{-\frac{1}{2}}, \quad (\text{D.14})$$

where  $q_0 \equiv q(\theta_0)$ , and we have taken the absolute value to get rid of the minus sign. Finally, using equation (D.12)

$$\left| \frac{dl}{dz} \right| = \frac{c}{H_0} (1 + z)^{-2} (1 + 2q_0 z)^{-\frac{1}{2}}. \quad (\text{D.15})$$

Recall that this expression is valid *only* for the special case of a *vanishing*  $\Lambda$  and *zero* pressure, and hence, is very unlikely to be valid for our Universe, as the most recent measurements of the fundamental cosmological parameters point to the fact that  $\Lambda \neq 0$ .

## APPENDIX E

---

### Evolutionary parameters of the Ly $\alpha$ absorbers

---

**Table E.1:** Compilation of the evolutionary parameters of the Ly $\alpha$  absorbers found in the literature. For the meaning of the different parameters see Chapter 4.

$z$	$\log N_{\text{HI}}$ [cm <sup>2</sup> ]	$N_0$	$\gamma$	B	$\beta$	$LOS/lines$	Reference
1.5 – 3.78	$W > 0.32$	–	$2.17 \pm 36$	–	–	11/277	Murdoch et al. (1986)
0.2 – 3.5	13.0 – 22.0	–	–	$1.45 \cdot 10^8$	$1.51 \pm 0.02$	34/83	Tytler (1987)
0.67 – 3.58	$> 17.2$	0.76	$0.68 \pm 0.54$	$10^{8.45 \pm 0.30}$	$1.51 \pm 0.02$	90/54( <i>LLS</i> )	Sargent et al. (1989)
0.67 – 3.58	$> 21$	$0.09 \pm 0.04$	$0.68 \pm 0.54$	$10^{8.45 \pm 0.30}$	$1.51 \pm 0.02$	90/6( <i>DLA</i> )	Sargent et al. (1989)
0.36 – 2.5	17.2 – 21.8	$1.2^{+1.2}_{-0.6}$	$0.3 \pm 0.9$	$10^{3.52 \pm 0.05}$	$1.25 \pm 0.03$	52/35( <i>LLS</i> )	Lanzetta (1991)
2.5 – 3.7	17.2 – 21.8	$(8.1 \pm 0.5) \cdot 10^{-4}$	$5.7 \pm 1.9$	$10^{3.52 \pm 0.05}$	$1.25 \pm 0.03$	52/35( <i>LLS</i> )	Lanzetta (1991)
2.7 – 3.4	$> 13.75$	–	$2.1 \pm 0.5$	–	$1.74 \pm 0.06$	1/301	Rauch et al. (1992)
2.5 – 4.3	13 – 22	–	$2.46 \pm 0.37$	–	$1.43 \pm 0.04$	–	Press et al. (1993)
0 – 5	12.3 – 17.20	$34.56^a$	2.46	$7.1 \cdot 10^5$	1.5	–	Madau (1995)
0 – 5	17.20 – 20	$0.94^a$	0.68	$2.08 \cdot 10^8$	1.5	–	Madau (1995)
2.53 – 3.19	12.3 – 22	–	–	$4.9 \cdot 10^7$	$-1.46^{+0.09}_{-0.05}$	4/1056	Hu et al. (1995)
2.95 – 3.6	$> 13.3$	–	2.2/2.5	–	1.8/1.4	1/178	Cristiani et al. (1995)
0 – 5	$> 14.0$	$16.8(+0.687)^b$	$0.8(+2.5)^b$	–	–	–	Muecket et al. (1996)
2.43 – 3.04	$> 12.21$	–	–	$6.2 \cdot 10^8$	1.5	1/	Kirkman & Tytler (1997)
2.0 – 3.5	12.8 – 16.0	3.12	$2.78 \pm 0.71$	–	–	5/536	Kim et al. (1997)
0 – 3.6	12.8 – 16.0	6.89	$2.15 \pm 0.21$	–	–	5/536	Kim et al. (1997)
0 – 4	12.8 – 16.0	–	$2.41 \pm 0.18$	–	–	5/536	Kim et al. (1997)
2.1 – 3.5	13.1 – 14.0	39.44	$1.29 \pm 0.45$	$4.9 \cdot 10^7$	1.46	5/536	Kim et al. (1997)
0 – 1.5	$> 14.0^b$	–	$0.6^b$	–	–	–	Riediger et al. (1998)
1.5 – 5.0	$> 14.0$	–	$2.9^b$	–	–	–	Riediger et al. (1998)
0 – 1.5	13.6 – 17.8	–	0.16	–	1.3	63/987	Weymann et al. (1998)
1.75 – 5.5	12.0 – 14.0	$181.5^a$	1.29	$1.73 \cdot 10^5$	1.46	–	Bershady et al. (1999)
1.75 – 5.5	14.0 – 17.20	$1.29^a$	3.1	$1.31 \cdot 10^6$	1.46	–	Bershady et al. (1999)
1.75 – 5.5	17.20 – 20.0	$0.92^a$	0.68	$2.08 \cdot 10^8$	1.5	–	Bershady et al. (1999)
1.5 – 4.0	12.5 – 14.0	–	–	–	1.4 – 1.5	6/	Kim et al. (2001)
1.5 – 4.0	12.5 – 14.0	–	–	–	1.4 – 1.5	6/	Kim et al. (2001)
1.5 – 4.0	13.64 – 16.0	–	$2.19 \pm 0.27$	–	–	6/	Kim et al. (2001)
1.5 – 4.0	13.1 – 14.0	–	$1.10 \pm 0.21$	–	–	6/	Kim et al. (2001)
1.5 – 4.0	13.64 – 17.0	6.1	$2.47 \pm 0.18$	$4.9 \cdot 10^7$	1.46	8/1166	Kim et al. (2002a)
0.9 – 1.7	13.0 – 16	$10.87 \pm 11.44$	$2.23 \pm 1.21$	$10^{9.62 \pm 0.58}$	$1.61 \pm 0.04$	1/235	Janknecht et al. (2002)
0.002 – 0.069	12.3 – 14.5	–	–	$10^{10.3 \pm 1.0}$	$1.65 \pm 0.07$	30/187	Penton et al. (2004)
0.002 – 0.069	14.5 – 17.5	–	–	$10^{5.2 \pm 4.9}$	$1.33 \pm 0.30$	30/187	Penton et al. (2004)
0.002 – 0.069	13.1 – 14.0	$70.79^{+10.49}_{-7.69}$	$0.76 \pm 0.12$	–	–	30/187	Penton et al. (2004)
0.002 – 0.069	$> 14.0$	$25.11^{+5.08}_{-3.73}$	–	–	–	30/187	Penton et al. (2004)
0.020 – 0.234	13.1 – 14.0	$179.24 \pm 13.38$	–	–	$1.96 \pm 0.15$	1/104	Williger et al. (2006)
0.002 – 0.423	$> 14.0$	$29.05 \pm 5.39$	–	–	$1.96 \pm 0.15$	1/104	Williger et al. (2006)
$< 0.4$	13.2 – 16.5	$96 \pm 7$	–	$1.99 \cdot 10^{13}$	$1.84 \pm 0.06$	7/341	Lehner et al. (2006)

<sup>a</sup> This number was computed by the authors from the normalisation constant  $A$  in the expression  $n(z) = A N^{-\beta} (1+z)^\gamma$  given in the reference by requiring that the column density distribution be normalised in the proper column density interval.

<sup>b</sup> This parameters were derived from simulations rather than from observations (see corresponding reference for details).

---

## Conferences & Ph.D. related activities

---

- Participation (**poster**) at the IAU Colloquium No. 199 *Probing Galaxies Through Quasar Absorption Lines*, 2005, Shanghai, China  
Poster title: The Stochastic Effect of Ly $\alpha$  Absorbers on Broadband Luminosities of Galaxies (see § Publications)
- Participation (**poster**) at the IAU Symposium No. 235 *Galaxy Evolution Across the Hubble Time*, 2006, Prague, Czech Republic  
Poster title: The Stochastic Nature of Attenuation and Its Impact on Photometric Redshifts (see § Publications)
- Participation (**contributed talk**) at the IAU Symposium No. 235 *Galaxy Evolution Across the Hubble Time*, 2006, Prague, Czech Republic  
Talk title: Chemically Consistent Evolutionary Synthesis modelling of Galaxies (see § Publications)
- Several talks at the Seminar for Extragalactic Astronomy, Institut für Astrophysik, Göttingen
- Participation at the advanced summer school *NOVICOSMO: The Dark and the Luminous Side of the Formation of Structure*, 2005, Novigrad, Croatia, organized by SISSA (person responsible: Paolo Salucci)
- Participation at the XIII IAG-LENAC Advanced School of Astrophysics *Structure Formation in the Universe*, 2006, Foz do Iguaçu, Brasil (person responsible: Beatriz Barbuy)
- Participation at the advanced summer school SSP 1177 *Evolution of Galaxies and their Large-Scale Environment*, 2006, Bad Honnef, Germany (person responsible: Peter Schneider)
- Teaching assistant for *Introduction to Astro- and Geophysics* at the Georg-August Universität, Göttingen, Germany
- Freshman Physics-lab assistant (Winter term 2005, Summer term 2006) at the Georg-August-Universität, Göttingen, Germany
- Research assistant at the Georg-August Universität, Winter term 2005
- Research associate at the Georg-August Universität, 2005 (until completion of Ph.D.)

---

## Summary of publications (refereed and non-refereed)

---

- **The Stochastic Effect of Ly $\alpha$  Absorbers on Broadband Luminosities of Galaxies.** Tepper García, T., Fritze-v.A., U., 2005, Proceedings of the IAU Colloquium No. 199 *Probing Galaxies Through Quasar Absorption Lines*, pp. 478-450, Williams, P.R., Shu, C., Ménard, B. eds., Cambridge University Press
- **Voigt Profile Fitting to Quasar Absorption Lines: A Simple Approximation to the Voigt-Hjerting Function.** Tepper García, T., 2006, MNRAS, 369, 2025
- **The Stochastic Nature of Attenuation and Its Impact on Photometric Redshifts.** Tepper García, T., Fritze-v.A., U., 2006, Proceedings of the IAU Symposium No. 235 *Galaxy Evolution Across the Hubble Time*, in press, Cambridge University Press, astro-ph/0609358
- **Chemically Consistent Evolutionary Synthesis Modelling of Galaxies.** Fritze-v.A., U., Tepper García, T., 2006, Proceedings of the IAU Symposium No. 235 *Galaxy Evolution Across the Hubble Time*, in press, Cambridge University Press, astro-ph/0610548
- **Stochastic Absorption of the Light of Background Sources due to Intergalactic Neutral Hydrogen I. Testing different line-number evolution models via the cosmic flux decrement.** Tepper-García, T., Fritze, U., MNRAS, submitted, astro-ph/0705.1242

---

# Acknowledgements

---

I would like to express my infinite gratitude to my parents Carmen and Werner, and to my sister Elke, who made the completion of this thesis possible through their continuous support in all respects.

I am especially grateful to my thesis advisor apl. Prof Dr. Uta Fritze for offering me the opportunity of working in this project, for her support all along this Ph.D., and for her continuous efforts to improve this work.

Furthermore, I am very grateful to my advisor, and also to Rike Fischer, Bruno SAN, and Miguel Verdugo for all the reading and correcting, and whose suggestion helped to significantly improve this work.

All the people at the Institute für Astrophysics Göttingen, especially the actual and former members of the Galaxy Evolution Group (Jens Bicker, Peter Anders, Markus Hartmann, Marie Aylin Tyra, Thomas Lilly) for their support and very helpful feedback in the realisation of this thesis. I want to thank Markus and Thomas in particular, for very helpful comments regarding the contents of this thesis, as well as for his support during uncountable nights of work at the Institute. In this respect, I also want to thank Pedro and Sergio!

Very special thanks go also to Ralf Kotulla to whom I am deeply in debt for his kind help with all types of troubleshooting during the last phase of this work.

Finally, to my friends in Göttingen, with whom I have spent some of the best moments of my life during the realisation of this Ph.D. Even though the list is long, I want to mention in particular my roommate Bruno, Gonzaga, Miguel and Cristiano, thanks a lot for all the fun we had together! I want also to thank all the people I got to know during my years in Gttingen, who made it an awesome time.

Also very special to mention is Ines, for her friendship and support all along this time.

And last but not least, I want to thank Rike, my climbing partner, for pushing me more than I was sometimes willing to, thus contributing to the improvement both of my climbing skills and this work.

This work was possible thanks to the financial support of the *Mexican Council for Science and Technology* (CONACyT), of the *Göttingen Graduate School of Physics* (GGSP) through a Lichtenberg-Stipendium, and of the *Universität Göttingen*.

Thorsten Tepper García  
20th June 2007

The numerical simulations used for this thesis were performed on an Linux Workstation at the Institut für Astrophysik Göttingen. This work was typeset using  $\LaTeX$ .



---

# Lebenslauf

---

Nachname: Tepper García  
Vorname: Thorsten  
Geburtstag: 17. August 1977  
Geburtsort: Mexiko-Stadt, Mexiko  
Staatsangehörigkeit: Deutsch

## Ausbildung

- **Promotion**
  - Juli 2007 an der Universität Göttingen (Deutschland)
  - Titel der Doktorarbeit: "The Stochastic Intergalactic Attenuation and its Impact on High-Redshift Sources"
  - Betreuer: apl. Prof. Dr. U. Fritze – v. Alvensleben (Göttingen, Deutschland)
- **Master of Science (M.Sc.)**
  - Oktober 2002 - September 2004 Universität Göttingen (Deutschland)
  - Studiengang: Physik (M.Sc.)
  - Titel der Masterarbeit: "The influence of intergalactic neutral hydrogen on the spectra of high-redshift galaxies"
  - Betreuer: Prof. Dr. Uta Fritze
  - Note: Ausgezeichnet
- **Bachelor of Science (B.Sc.)**
  - August 1996 - April 2001 Universidad de la Amricas-Puebla (Mexiko)
  - Notendurchschnitt: 9.9 (von 10.0)
  - Titel der Bachelorarbeit: "Supersymmetry and the constants of motion of the two-dimensional isotropic harmonic oscillator"
  - Betreuer: Dr. Gerardo Francisco Torres del Castillo
  - Note: Summa Cum Laude
- **Schulische Ausbildung**
  - September 1983 - Juni 1996, Guadalajara (Mexiko)
  - Schulabschluss (dem Abitur gleichwertig) am 19. Juni 1996
  - Notendurchschnitt = 9.9 (von 10.0)

