



PROBING EARLY GALAXY FORMATION WITH DAMPED LYMAN α -ABSORBERS

*A study of a $z=2.4$ Damped Lyman α -Absorber along
the line of sight to the $z=3$ quasar SDSS 0918+1636*

Thøger Emil Juul Thorsen

Thesis for the degree of:

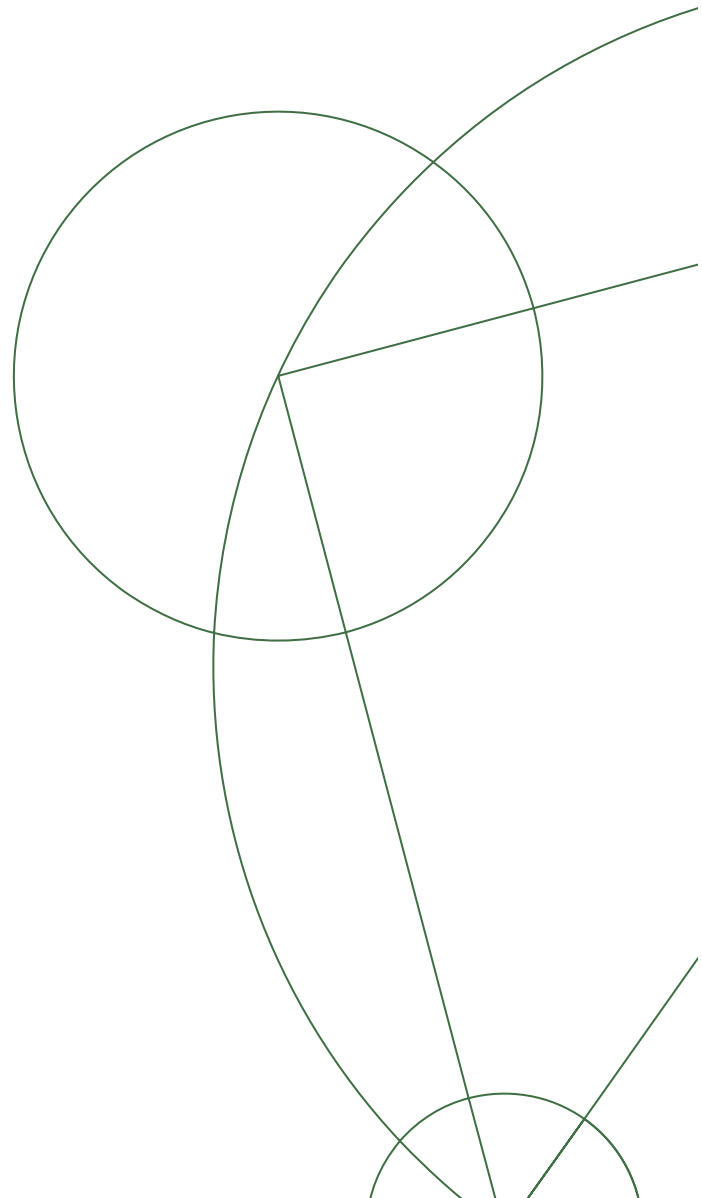
Master of Science in Physics (Astrophysics)

Submitted date:

Tuesday, March 1st, 2011

Supervisor:

Johan P. U. Fynbo, DARK Cosmology Centre.



Abstract

The overall goal of this project has been to contribute to our understanding of the nature of the first galaxies through the study of a Damped Lyman α Absorber. DLAs are clouds with a high column density of neutral hydrogen, leaving a characteristically strong and broad absorption line in the spectrum of a background light source, in this case a quasar.

The study of the DLA absorption features in the spectra of quasars is a powerful tool for probing early history of galaxy formation, as it allows us to extract detailed information about chemical composition and kinematical properties in the gas of very distant and/or faint galaxies. The information gained from the DLA can help understand the processes by which the large scale structure of the Universe has evolved from minute temperature fluctuations in the Cosmic Microwave Background to the galaxies, clusters and enormous voids of the local Universe; and hence understand the nature of the Dark Matter, the gravitational influence of which is governing this evolution. Furthermore, it can help understand how the early clouds of hydrogen and helium have since been chemically enriched with the elements that today make up our planet and a major part of ourselves.

The data studied here are produced by a group led by Johan Fynbo in an attempt to find weak lines of emission from the galaxies hosting the DLA clouds. The data investigated consist of three exposures of a quasar of redshift $z \approx 3.07$, taken with the X-Shooter spectrograph at the ESO Very Large Telescope in Paranal, Chile. In these spectra were found two DLA systems with redshifts $z \approx 2.58$ and $z \approx 2.412$, respectively, of which the latter was the subject for this work.

The system was found to be a relatively hydrogen rich DLA with a metal abundance and evidence of dust. It shows a dynamical structure consistent with the assumption that DLA systems are young galaxies with active star formation, and that the more massive systems have the higher star formation rate and correspondingly higher abundances of metals.

Resumé

Formålet med dette specialeprojekt har været at bidrage til vor forståelse af de første galakser igennem analyse af en Dæmpet Lyman- α -Absorber, forkortet DLA. Disse er skyer med et højt indhold af neutral brint, som efterlader en karakteristisk kraftig og bred absorptionslinje i spektret fra en lyskilde i baggrunden - i dette tilfælde en quasar. Studiet af DLA'ers absorption i quasarspektre er et virkningsfuldt middel til at sondere den tidligste galaksedannelseshistorie, idet den sætter os i stand til at skaffe detaljeret viden om kemisk sammensætning og kinematiske forhold i gassen i meget fjerne og/eller lyssvage galakser.

Denne information kan hjælpe til at forstå de processer, ved hvilke Universets storskalastruktur har udviklet sig fra uhyre små temperaturforskelle i den Kosmiske Mikrobølgebaggrund til de nuværende galakser, galaksehobe og enorme gasfilamenter, og dermed også at forstå det Mørke Stof, hvis gravitationelle indflydelse styrer denne udvikling. Den kan desuden hjælpe os til at forstå, hvordan de oprindelige skyer af brint og helium er blevet beriget med de grundstoffer, der idag udgør vores planet samt en stor del af os selv.

De anvendte data er produceret af en gruppe ledet af Johan Fynbo med det formål at søge efter svage emissionslinjer fra den galakse, der skaber DLA-absorptionslinjen i quasarens spektrum. De anvendte data indeholder tre eksponeringer af en quasar med en rødforskydning på $z \approx 3.07$, taget med X-Shooter-spektrografen på ESO's Very Large Telescope i Paranal, Chile. I disse spektra fandtes to DLA-systemer med rødforskydninger på hhv. $z \approx 2.583$ og $z \approx 2.412$. Af disse er sidstnævnte blevet undersøgt i dette projekt; ingen emission blev dog fundet i dette.

Det undersøgte system fandtes at være en relativt brintrig DLA med et højt indhold af metaller tegn på tilstedeværelse af støv. Det udviser en dynamisk struktur i overensstemmelse med den antagelse, at DLA-systemerne er unge galakser med aktiv stjernedannekse, og at systemerne med størst masse udviser den kraftigste stjernedannelsesaktivitet og tilsvarende hurtigste berigelse med tunge grundstoffer.

CONTENTS

1	Introduction	7
1.1	The Formation of structure	7
1.2	Chemical evolution	9
1.2.1	Big Bang Nucleosynthesis	9
1.2.2	The Dark Ages and reionization	9
1.3	Quasars	10
2	Methods	15
2.1	Absorption line systems	15
2.1.1	The Lyman α transition	15
2.1.2	Damped Lyman α Absorbers	19
2.1.3	Voigt Profiles	24
2.1.4	Robust Lower Limits	29
2.2	Sample selection	32
2.3	Error propagation	34
3	Data Reduction	39
3.1	The X-Shooter	39
3.2	Observation Conditions	40
3.2.1	Resolution	41
3.3	Extraction and Normalization	43
3.3.1	The X-Shooter pipeline	43
3.3.2	Wavelength correction	46
3.3.3	Stacking & Extraction	47
3.3.4	Continuum modeling	49
3.3.5	Normalization	52
3.3.6	Telluric absorption	54
4	Analysis and results	57
4.1	DLA line fitting	57
4.2	Robust lower limits	58

Contents

4.3	Line profile fitting	60
4.3.1	vpfit	60
4.3.2	vpguess	61
4.3.3	Velocity components	63
4.3.4	States of ionization	64
4.3.5	Turbulent motion	64
4.3.6	Thermal motion	69
4.3.7	Metallicities	72
4.4	Dynamics	73
4.5	Discussion	75
4.6	Summary and outlook	77
4.6.1	Possible future work	77
Appendices		81
A	Reverse flux-calibration of the composite quasar spectrum	81
A.1	Extragalactic extinction	81
A.2	Galactic extinction and reddening	82
A.3	Atmospheric extinction and instrument response	84
B	Source Code	85
B.1	Pipeline-script	87
B.2	Normalization of VIS arm	94
B.3	DLA fitting routine	102
B.4	Metallicity calculation	106

PREFACE

*I know how hard it is in Latian verse
To tell the dark discoveries of the Greeks
Chiefly because our pauper-speech must find
Strange terms to fit the strangeness of the thing;
yet worth of thine and the expected joy
Of thy sweet friendship do persuade me on
To bear all toil and wake the clear nights through,
Seeking with what of words and what of song
I may at last most gloriously uncloud
For thee the light beyond, wherewith to view
The core of being at the centre hid.*
Titus Lucretius Carus: *On The Nature Of Things*

In this thesis, I am going to present my work of the last year on the study of Damped Lyman α Absorbers in the spectra of Quasi-Stellar Objects, called QSOs or quasars. Quasars are the extremely luminous, dense regions surrounding Supermassive black holes in the centers of distant galaxies. When a cloud of neutral hydrogen in a galaxy intersects the line of sight to a quasar, it absorbs some of its light, leaving a characteristic strong, broad absorption line in the spectrum of the quasar, the so-called Damped Lyman α line. This line is easily detected, even in automated surveys, where it can be marked for follow-up spectroscopy, in which absorption lines left by heavier elements in the absorption system combined with the DLA line can reveal detailed information about the absorbers chemical composition, kinematic properties and inner structure, which in turn can help us understand the history of formation and evolution of galaxies and their place in the evolution of the Universe as a whole.

The data used were provided by my advisor, Johan Fynbo, whose group has acquired the spectra with the X-Shooter spectrograph on the Very Large Telescope in Paranal, Chile. The exposures were made in February 2010, a few weeks before my work on the project started. My work has mainly consisted in reducing and analyzing the data after acquisition, a process that has constituted the majority of the work on this thesis, which I also hope to have reflected in the structure of this report.

In the first chapter of this document, I shall briefly summarize the current

Contents

standard model of cosmology, in order to set the stage and provide some context for this work. In chapter two, I shall present in greater depth some of the theoretical methods used in the reduction and analysis of the spectra. In chapters 3 and 4 I shall give an account of the process and results of the reduction and analysis, respectively. There is a considerable topical overlap between chapters 1 and 2 which provide the theoretical background, and between chapters 3 and 4 which document the practical process and present the results.

ACKNOWLEDGEMENTS

A bunch of people deserve a lot more than just a “thank you” here, but a “thank you” you still shall have.

First of all **Christina**; for putting up with my sometimes insane stress level during the last two years.

To **Johan Fynbo** for having been a patient, helpful and supportive advisor, even when I’ve been confusion incarnated.

To **Martin Sparre** for invaluable help reviewing and proof-reading.

To my **parents, sisters and brothers** for having been supportive and inspiring, and my parents specifically for having taught me the courage to wrestle the big questions in Life.

Ása for being such a great help in the initial stages of my project. I hope I can provide similar help to some newcomer around now.

To **Catarina, Karen, Martina, Bo, Peter, Mitch, Henrik, Christiano** and **Diana** and all the other DARKers.

Majken, Ursula, Sanne, Kenneth and the other Astronomy people who happen to not be DARKers but still make up a good crew.

My **friends**, because you are awesome, but too many to list, so I’ll only list those I believe would possibly read the list.

And last but not least, **Bruce Lee**.

CHAPTER 1

INTRODUCTION

The sort of study of Damped Lyman α systems carried out in this thesis project can provide information about the history of both structure formation and chemical enrichment in the CMB Universe. To set the stage for this, I shall now give a brief account of early structure formation and chemical composition.

1.1 THE FORMATION OF STRUCTURE

The currently most widely accepted theory of the origin of the Universe is the **inflationary big bang** theory. It states that at an age of $T = 10^{-36}$ s after the Big Bang, the universe started expanding exponentially until around $T \approx 10^{-34}$.

During this extremely short time span, the Universe would have expanded by roughly a factor of $\sim 10^{43}$; corresponding to a sphere with the radius of 1% that of the electron expanding to the entire observable Universe today in a time span of 10^{-34} seconds. This would mean that even if there were any inhomogeneities in the Universe before inflation, these would be stretched so much that they would be completely smoothed out; after inflation, the Universe would be completely flat.

There would however be some very small inhomogeneities, due to quantum mechanical fluctuations taking place *during* inflation. These would leave little energy density ripples on the fabric of spacetime, which would then instantly stretch with the rest of the Universe until inflation ends, leaving density fluctuations on practically all scales imaginable. At the end of inflation, the field believed to be responsible for inflation would decay into an ordinary energy-momentum-tensor, which would fill up the hitherto empty and extremely cold universe in what is called the epoch of reheating.

How the density fluctuations grow with time is a competition between gravity, that tends to make the perturbations grow, and high temperature which will

Chapter 1. Introduction

tend to smooth them out. In the early, hot, dense Universe, matter and radiation were in thermal equilibrium, making up a photon-baryon plasma with very high temperature which would not allow any gravitational collapse; and hence the density perturbations of the inflation era were left to oscillate as extremely faint sonic waves. Only when the Universe grew large and cool enough that baryons and photons would decouple, and baryons and electrons combine into neutral hydrogen and helium atoms would baryons be able to undergo gravitational collapse. Hence, if the Universe only consisted of baryonic matter, these density fluctuations would only start growing at the era of recombination, and the Cosmic Microwave Background that we see today would be completely smooth.

However, this is not what the CMB looks like. The CMB has a remarkably homogeneous temperature, but it does have small density fluctuations of $\sim \frac{\delta T}{T}$, which is already too large to have been created by the gravitational collapse of baryons. The strength of the fluctuations in the CMB is one important piece of evidence for the existence of non-baryonic **Dark Matter**; this does not interact with photons and will hence have decoupled from the rest of the Universe at a significantly earlier stage, giving it time to undergo gravitational collapse and settle into the higher-density structures. Studies of the CMB fluctuations have shown, in agreement with measurements of spiral galaxies and galaxy clusters, that there is roughly ten times as much Dark Matter as there is baryonic matter in the Universe, meaning that the gravitation of Dark Matter is dominating. After recombination, the baryonic matter simply falls into the potential wells already established by Dark Matter; and the tiny differences in gravitational potential red- and blueshifts the photons emitted there slightly, creating the temperature fluctuations in the CMB.

These tiny density fluctuations are the seeds that have since grown into the structures of the Universe today, with stars, planets, galaxies and intergalactic gas. How this evolution happened is the underlying question in a large part of modern observational cosmology, including this thesis project.

HOT AND COLD DARK MATTER

There are many candidates for what could be the evasive Dark Matter; and they generally fall in two categories, *hot* or *cold* Dark Matter. Hot Dark Matter moves at relativistic speeds, while Cold Dark Matter moves at nonrelativistic speeds and carry correspondingly smaller kinetic energies. This has an important impact on the formation of structure in the Universe: Hot Dark Matter will move freely at larger distances than cold dark matter, tending to smooth over inhomogeneities at small scales. Hence, if the Universe is dominated by HDM, large scale structures will collapse gravitationally before smaller scale structures, the so-called *top-down* scenario. In contrast, CDM domination will see smaller structures col-

lapse gravitationally first, and then merge into larger structures, the *bottom-up* scenario.

Observing the early Universe to see if gas collapses on supercluster scales and then develop smaller substructures, or gas collapses at smaller scales like stars and star clusters to then merge into larger structures is therefore a powerful tool to constrain the ratio of hot to cold Dark Matter in the Universe.

1.2 CHEMICAL EVOLUTION

1.2.1 BIG BANG NUCLEOSYNTHESIS

At around 1 second after Big Bang, the Universe had cooled enough that neutrinos could freeze out of equilibrium, fixing the neutron-to-proton ratio at 0.2. Free neutrons are unstable and will decay with a half-time of about 15 minutes, but as the Universe cools down further, the neutrons bind to an equal number of protons to form Helium 4-nuclei through the process of **Big Bang Nucleosynthesis**. The temperature does however not get below the binding energy until a time of around 200 s [31], a time that is non-negligible compared to the half-life of free neutrons. At this time, the neutron-to-proton ratio has decreased to around 0.15, at which level it is frozen due to the stability of nucleus neutrons. After the end of nucleosynthesis at $T \sim 300$ s, there is roughly 2 neutrons for every 14 protons, tied up in one helium atom and 10 hydrogen atoms, leaving a primordial chemical composition of the Universe of $\sim 25\%$ Helium and $\sim 75\%$ Hydrogen by mass, or 10% Helium and 90% hydrogen by numbers. In addition to this, there are traces of Lithium and Beryllium. Any heavier elements and raised Helium abundance can only be the result of stellar activity, making metal lines in the ISM and IGM a useful tracer of current or past star formation.

1.2.2 THE DARK AGES AND REIONIZATION

At a redshift of $z \sim 1100$ [4], corresponding to an age of the universe of around 300.000 years, the Universe finally cools to a level below the ionization energy of hydrogen. Nuclei and electrons form neutral atoms in what is somewhat misleadingly called the **era of recombination**. Since the atoms are neutral, they stop interacting with radiation, and for the first time, the Universe becomes transparent, and the photons present everywhere can now travel freely through the Universe. The background radiation is soon redshifted into low-energetic infrared and microwave wavelengths, and there are no stars or other processes providing light, apart from the 21 cm Hydrogen spin-flip line; the Universe has entered the so-called **dark ages**. From this point and until a redshift $z \sim 20$, the Universe

is completely dark, and dominated by neutral hydrogen and Dark Matter, slowly collapsing into the gravitational wells of what shall later become the first stars and galaxies. The oldest galaxy observed so far has a redshift of $z \sim 10$ [5], meaning that star formation must have been going on for a while before that.

Whether the first stars have formed before the first galaxies or vice versa depends on the hot to cold Dark Matter ratio and is still unknown. The hard UV radiation from the first stars marks the beginning of the **era of reionization**, the end of which is defined as the time when only a fraction of $\sim 10^{-3}$ of the hydrogen of the Universe is still neutral. This is approximately the limit at which the Universe becomes transparent to radiation at the Lyman α wavelength of 1216 Å (see next chapter).

1.3 ACTIVE GALACTIC NUCLEI AND QUASARS

Shortly after the first galaxies - maybe even at the same time - the their slightly more curious cousins the **quasars** appear. The term “quasar” is an abbreviation of a “Quasi-Stellar Radio Source”, a class of objects first discovered in the 1950’s. The term quasi-stellar was given them because of their very small angular size; most are, like stars, to be considered point sources, yet they show an emission spectrum very different from that of stars, and they were found to be strong radio sources as well - in fact, the first quasars were discovered with radio telescopes. However, it was soon discovered that not all of these stellar-like high redshift objects were emitting at radio wavelengths, and the name “quasi-stellar objects”, or **QSOs**, was given to the radio-quiet ones. Today, it has been found that there is no clear distinction between these two kinds of objects, and the term “quasar” is often used more loosely about any kind of QSO, a practice that is also adopted in this thesis.

Quasars were soon found to be one subspecies of the class of objects called **Active Galactic Nuclei**, a class that also includes e.g. Seyfert Galaxies and Radio galaxies. These objects are, from an observational point of view, very different: some are radio loud, some are radio quiet, some have strong UV and visible spectrum emission, some show none of that. However, what they all have in common is a very high redshift, implying cosmological distances, and a small angular sizes.

The spectra of quasars show some characteristic features, including strong and very broad emission lines. Typical width of these lines would indicate that they are emitted from systems with a velocity dispersion of several percent of the speed of light. Some of the sources also show changes in both continuum and line emission on time scales of days or weeks, permitting the emitting region to be no more than a few light-weeks across. Adopting an example width of the region of $R \sim 0.01$ pc and a width of the emission lines of $v \sim 10^4$ km s $^{-1}$ [18], and

inserting into the usual formula for escape velocity, the central region is found to have a mass of:

$$v_e = \sqrt{\frac{2GM}{R}} \Rightarrow$$

$$M = \frac{v^2 R}{2G} \approx 1.16 \times 10^8 M_\odot$$

A mass measured in millions to billions of solar masses enclosed in a volume with a radius a few times that of the Solar System is almost certain to be a supermassive black hole, which is also the only known engine that can produce the energies detected, and later modeling has shown that also the characteristics of other kinds of AGN are consistent with them being the highly energetic inner surroundings of supermassive black holes in galactic centers, their apparent differences being due to the face they show us rather than differences in physical properties.

Quasars are by far the most luminous of these objects. While the other classes of AGN are visible galaxies with strongly radio-emitting cores, quasars appear to be “cores with no galaxies”, appearing very similar to Galactic stars at first glance; but they show a stronger UV and IR emission than stars in addition to their very strongly redshifted emission lines.

Deep imaging has since shown that also the quasars are indeed the extremely luminous cores of galaxies, but the quasar typically outshines the galaxy by a factor of $\sim 10^3$.

All AGN emission is today believed to be different manifestations of potential energy released by infall of matter into the central black hole. When a massive central black hole has formed in a young galaxy, matter is falling into the potential well of the black hole and forming an accretion disk around the center. Friction within the disc material will cause it to lose its potential energy and release it as thermal radiation. The accreting matter will typically radiate $\sim 10\%$ of its rest mass during infall due to the enormous gradient in potential energy around the central black hole [18], making for extremely high temperatures in the disk, and the continuum emission contribution will show correspondingly high energies.

At these temperatures, most matter will be in the plasma phase of freely moving nuclei and electrons, and the strong rotation and friction will cause a strong magnetic field to build up, sometimes strong enough to power narrow jets of charged particle outflow at highly relativistic speeds, feeding the radio lobes if present. These jets emit strong synchrotron radiation in UV and γ bands. When one such jet happens to point in our direction, we observe a luminosity many times stronger yet, a so called *blazar*. It is not known precisely what fuels a

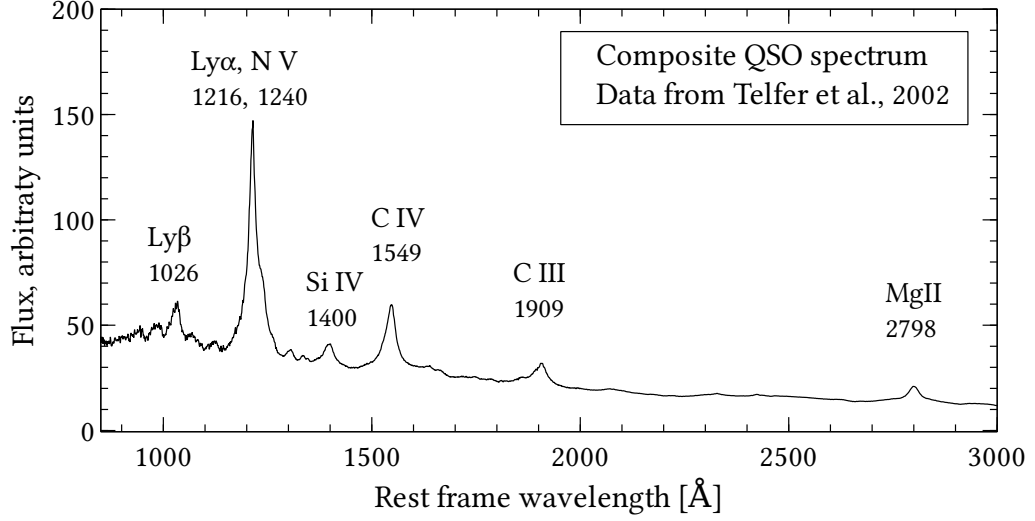


Figure 1.1: Part of a composite QSO spectrum, made by Telfer et al. [33]. The strong, broad emission lines are clear, as is the relatively strong UV continuum.

quasar, as it seem that a large abundance of neutral gas is necessary but not sufficient.

QUASARS AS COSMOLOGICAL PROBES

The strong luminosity of QSOs means that they can be observed out to very high distances; the farthest observed QSO so far has a redshift of $z \gtrsim 7$ ¹, meaning that the light we observe is no more than 12.7 Gyr old, at which time the Universe was only $\sim 12\%$ its present scale and $\sim 5.5\%$ its present age. This means this light was emitted before the end of the era of reionization at $z \sim 6$ (see section 2.1).

QSOs are seen within the full range of observable redshifts [18], but their comoving number density per unit redshift is variable with a peak around $z \sim 2 - 3$, above which it starts dropping again. This era seems to be a peak age of QSO activity. The abundance of quasars at different redshifts can, provided a correct theory of the fueling mechanisms and their relation to the mechanisms of galaxy formation, they can provide information about the evolution of galaxies in the Universe. One hypothesis is that quasars have a relatively short life time but can be “re-ignited” as a result of galaxy mergers, in which case the abundance of quasars could also carry information about galaxy collision rates and hence the history of structure formation. The relative strengths of the strong emission

¹Bram Venemans, talk at DARK Cosmology Center

lines bear information about the chemical evolution of the Universe and in young galaxies in particular.

But maybe more important, quasars can serve as cosmic “lighthouses”, providing backlight for absorption spectroscopy of primeval galaxies and the intergalactic medium. The bright light of quasars make it possible to probe the IGM as far back as it is possible to observe quasars, well into the pre-ionization era.

As is explained further below, the strong light from quasars provide statistical information about the structure of the IGM from reionization and to the local Universe, as well as observing young galaxies that are not luminosity selected and hence are likely to be more representative of the typical galaxy at high redshifts than galaxies observed directly.

CHAPTER 2

METHODS

2.1 ABSORPTION LINE SYSTEMS

In much the same way that a cloud obscuring the sun or the moon can give us information about the thickness and structure of the cloud, background light from e.g. a quasar can give us information about any matter partly obscuring this light. The atoms of the obscuring matter leave absorption lines in the spectrum of the light source, revealing information about the absorber like e.g. chemical composition, number of atoms encountered, and more. Correctly interpreted, this information can reveal even more information about e.g. structure and dynamics of the absorber.

When the light source is e.g. a quasar, beaming at us over cosmological distances, there is an added benefit from the fact that the light is redshifted as the fabric of spacetime stretches while the photons traverse it. When meeting an absorbing system, the QSO spectrum will already be redshifted, and the absorber will leave its absorption line imprints on it bluewards of the quasar emission line of the corresponding transition. Given a correctly modeled relation between redshift and distance, absorption lines will then provide a mapping of the distances, compositions and concentrations of matter encountered on the way. Thanks to the very large distances at which quasars can be observed, this allows us to obtain statistical information about the structure and composition of the Universe out to very high distances and corresponding early times; in recent years, observation of extremely distant quasars has even allowed a dating of the completion of reionization of the Universe to a redshift of $z \approx 6$ [3] (see below).

2.1.1 THE LYMAN- α TRANSITION LINE

By number, hydrogen makes up $\sim 90\%$ of the atoms in the Universe, and any photon is far more likely to encounter a hydrogen atom than any other element.

The spectrum of hydrogen is therefore an important tool in mapping the mass distribution in the Universe.

The strongest line in the in the Hydrogen spectrum is the 1216 Å Lyman α transition between the ground state and the first excited state of the hydrogen atom. Lyman α photons are mainly produced in regions where photons of energies above the **Lyman limit** at 912 Å can ionize hydrogen which then recombines and decays back into the ground state. This happens mainly in hot, star forming regions and inside AGNs and, to a lesser degree, in the cooling of collapsing gas [21]. Young, star-forming galaxies will typically be very bright in the Ly α line, which makes them appear significantly brighter in a narrow band around their (redshifted) Ly α line than in a broad continuum, a fact that can be utilized for bulk mapping of galaxy redshifts in wide-field images like in e.g. [36].

When light of Lyman α wavelength passes through a cloud of neutral hydrogen, of which at any time most will be in the ground state, a number of the photons will be absorbed to excite the atoms into the upper state, which will almost immediately decay back into the ground state by reversing the process and emitting a Ly α photon. The life time of the upper state is long enough that the photon will be re-emitted in a random direction, meaning that along the line of sight to the light source, the cloud has effectively absorbed light at this wavelength. The efficiency of the absorber can be expressed as the absorption **cross section area** σ . Lyman α absorption in neutral hydrogen has a high interaction cross section, meaning that small abundances of neutral hydrogen can leave noticeable or even saturated line features in a spectrum. This means that the distribution of neutral hydrogen in the Universe can be efficiently mapped by aid of Lyman α absorption in QSO spectra. Different distribution patterns will leave different imprints on the QSO continuum. Some of these features, as well as their physical implications, are described below.

THE GUNN-PETERSEN TROUGH

Since Lyman α is so efficiently absorbed in neutral hydrogen, the Universe would have to be almost completely ionized before becoming transparent to this wavelength. Light emitted before this time at wavelengths of $\lambda \leq 1216\text{Å}$ will be continually redshifted into and through the Ly α wavelength and absorbed, sparing the one half of the broad Ly α emission line that has been Doppler shifted to redwards from the rest-frame wavelength of the transition. A deep absorption feature will work its way bluewards from the center of the Ly α emission line. This effect will end when the Universe has been almost completely reionized, giving the absorption feature a lower wavelength limit of the Ly α wavelength at the redshift at the end of the epoch of reionization. This limit can well be on the blue side of what was the 912 Å Lyman Limit at emission; in which case there

will be strong absorption everywhere bluewards of the $\text{Ly}\alpha$ emission line of the quasar.

This absorption feature was first predicted by Gunn and Petersen in [14] and is called the Gunn-Petersen effect or the **Gunn-Petersen Trough**, but the first detection of the feature was not until 2001, when it was announced by Robert H. Becker et al. [3] that they had found a schoolbook example of a Gunn-Petersen trough in the spectrum of the $z = 6.28$ QSO SDSS 1030+0524 and similar features in three quasars of redshifts between 5.80 and 6.28, determining the end of the epoch of reionization to be around $z \sim 6$.

THE LYMAN α FOREST

In the Universe after reionization, only a small fraction of the Hydrogen is still neutral. This hydrogen is concentrated in the filaments and inter-void walls described in section 1.1. Following the general evolution of the large scale structure, the largest concentration of hydrogen will be found in the nodes and filaments, where parts of it will be shielded against ionizing radiation.

When reionization is complete, the large majority of the Universe has gone transparent to the Lyman α transition, but these clouds of neutral hydrogen still absorb in these wavelengths, thus revealing their thickness, position and other properties.

Most of these clouds are small, dilute and cold, though often with column densities large enough for saturated absorption. When the light beam of a QSO emitted after reionization travels through intergalactic space, continually being redshifted, it encounters a large number of these clouds, each leaving a thin but often deep absorption line at the $\text{Ly}\alpha$ wavelength of the absorber redshift. The thinner and smaller of these clouds, with column densities $N \lesssim 1.6 \times 10^{17} \text{ cm}^{-2}$ [20], are also the more numerous. They are large enough to be optically thick in the strongest hydrogen lines including $\text{Ly}\alpha$ and $\text{Ly}\beta$, but are optically thin in metal lines. A large number of these lines can be found bluewards of the $\text{Ly}\alpha$ emission feature in many high-redshift spectra, resulting in a comb-teeth pattern commonly known as the **Lyman α forest**.

As the Universe expands further, the IGM is further ionized, and the thin HI systems become fewer and smaller, resulting in a thinning out of the $\text{Ly}\alpha$ forest at lower redshifts.

An example of a $\text{Ly}\alpha$ forest is shown in figure 2.1, where a low- and a high-redshift quasar spectrum are shown together for comparison.

The number density per redshift bin and shape of the Lyman α forest lines provide valuable statistical constraints on the formation of structure. As expected from the picture of an expanding and reionizing early Universe, the density of the lines in a QSO spectrum per unit redshift is decreasing with increasing redshift.

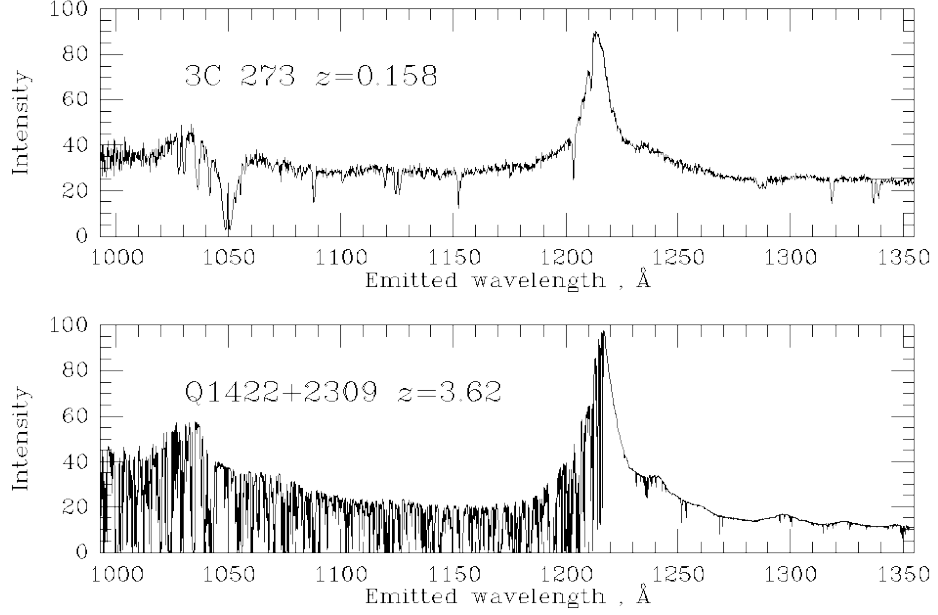


Figure 2.1: Quasar spectra from a low redshift (upper panel) and a high redshift (lower panel) quasar for comparison. The Lyman α forest is clearly visible in the lower panel. Image adapted from [19].

The density of lines can normally be fitted as a power law

$$N(z) = N_0(1 + z)^\gamma; \quad (2.1)$$

with N_0 being a function of line strength, as weaker lines are more numerous than strong lines [29]. The exponent $\gamma \approx 2.5$ is explained further by Petitjean in [29]. The Ly α Forest lines show less clustering than lines of higher column density, consistent with what would be expected from a scenario where these dilute clouds are associated with IGM filaments and walls, and only to a small degree with galaxies [18]. The number of Ly α Forest lines is roughly constant at redshifts of roughly $z \lesssim 1.5$, indicating that the dilute clouds at this point have either been ionized and diffused into the hot IGM or collapsed into denser clouds [18]. Even the most extended, dilute Ly α Forest systems show a small but measurable abundance of metals. Since it is unlikely that present day galaxies have had the time to enrich the dilute IGM with these metals, this suggests a first generation of stars forming in the early and much more homogeneous Universe, which have subsequently merged into galaxies [18], lending further support to a Cold Dark Matter scenario in which small scale density fluctuations have collapsed first, and the larger scales have followed as the Universe cooled and expanded.

LYMAN LIMIT SYSTEMS

Systems at column densities larger than the Ly α Forest systems, at column densities $1.6 \times 10^{17} \lesssim N \leq 2 \times 10^{20} \text{ cm}^{-2}$ are optically thick in the Lyman α line, and in addition also optically thick in wavelengths below the 912 Å ionization energy of hydrogen, the so called Lyman Limit, after which they have their name, the **Lyman Limit systems**. These systems are mixed ionization states, and are believed to be found mostly within virial radii of galaxies [20], associating them physically to the galaxies. They are known to dominate the absorption of ionizing photons in the Universe [20], but they are not very well understood and shall not be treated in further detail here.

Systems with column densities greater yet, of $N \geq 2 \times 10^{20} \text{ cm}^{-2}$, are the so-called **Damped Lyman α Absorber** systems and are considerably better understood; and since they form the topic of this thesis, they shall be treated separately in the next section.

2.1.2 DAMPED LYMAN α ABSORBERS

Systems denser yet than the Lyman Limit systems are rarer, but are found in some spectra of quasars and GRBs. They show heavily saturated Lyman α absorption lines with broad dampening wings in the sides, the origin of which is described further in section 2.1.3, and which have given the systems their names: the Damped Lyman α Absorbers, or DLA. The $N \geq 2 \times 10^{20} \text{ cm}^{-2}$ limit is in fact slightly artificial, as these broad dampening wings are seen in systems with column densities as low as $N \sim 10^{18} \text{ cm}^{-2}$ [29], the so-called sub-DLA systems. The $N \sim 10^{20} \text{ cm}^{-2}$ limit stems part from the studies of local spiral galaxies which show that this is the typical order of magnitude of HI column density in these; part from this being the level at which also metal lines become unambiguously identifiable [13]. Hence, the limit is set not just from the occurrence of the dampening wings, but also from what seems to be “galaxy-like” conditions.

Figure 2.2 from [29] shows the distribution of HI clouds at $z \sim 2.8$ as a function of column density. With column densities $10^{12} \lesssim N \lesssim 10^{14} \text{ cm}^{-2}$, the Lyman α forest systems are by far the more numerous. However, the major part of HI in the Universe is found in the DLA systems to the lower right in the figure. The distribution of the stars is well fitted with a power law $\frac{dn}{dz} \propto N^{-1.5}$ [29] [13]. The number of clouds with a column density within the interval $[N, N + dN]$ is proportional to $N^{-1.5}$. The amount of gas in clouds within this interval is then proportional to the gas content per cloud times the number of clouds $N \times N^{-1.5} = N^{-0.5}$. The total gas content in a given column density

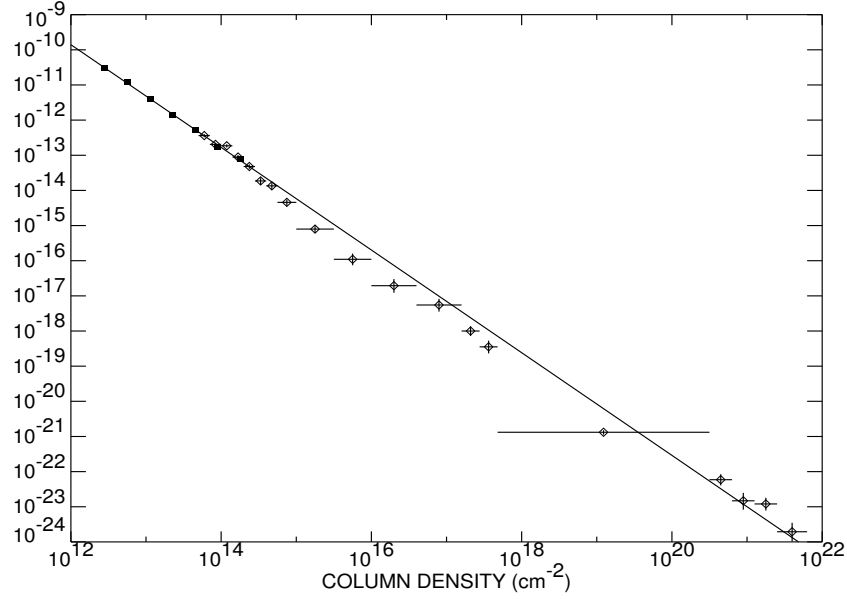


Figure 2.2: The column density distribution of HI clouds in the $z \sim 2.8$ Universe from [29]. The thin ($N \lesssim 10^{14} \text{ cm}^{-2}$) clouds are by far the more numerous, but the majority of HI is found in the large ($N \gtrsim 10^{20}$) systems.

interval $[N_1, N_2]$ is then found by

$$n([N_1, N_2]) = \int_{N_1}^{N_2} N^{-0.5} dN \quad (2.2)$$

Since $1/N^{0.5} > 1/N$, and the latter does not converge when integrated, neither will the former, so the upper end of the column densities will dominate the integral, and the majority of the mass will indeed be found in the upper end of the column density scale, regardless of the lower number of systems.

Figure 2.3 can sum up the presentation of the Ly α transition quite nicely. It shows the part of the UVB arm of the spectrum of Q 0918+1636, the QSO that is the subject of this work. Two DLA systems as well as the Lyman α Forest are clearly visible. Around 3550 Å is seen the sharp cutoff of a Lyman limit system with a redshift of $z \approx 2.85$, whose Lyman α line is visible as a semi-strong absorption feature around 4750 Å. The system treated in this work is the one responsible for the DLA line centered around 4149 Å, just bluewards of the Lyman β emission line of the QSO.

DLA AS (EARLY) GALAXIES

Neutral hydrogen is the fuel and building material for star formation and hence for galaxy formation, and since the major part of neutral hydrogen is concen-

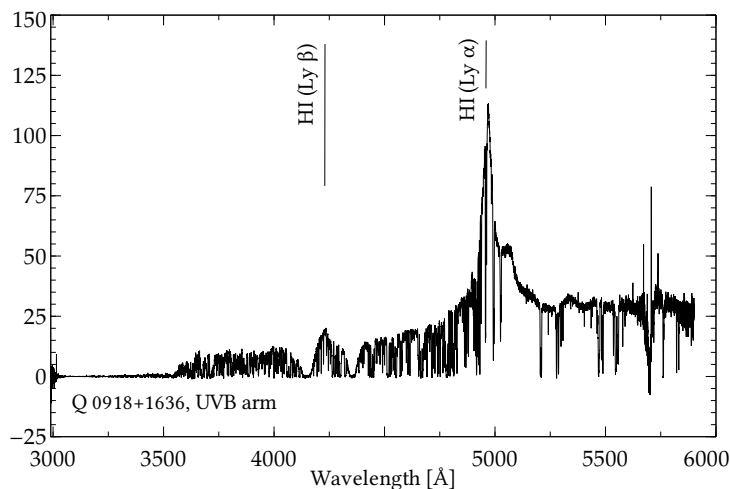


Figure 2.3: The spectrum of *Q 0918+1636*, with the $\text{Ly}\alpha$ and Lyman β emission lines marked. Two DLA features as well as the Lyman α forest and a Lyman limit system are visible in the spectrum

trated in the DLA systems, these are considered the reservoirs of HI from which early Universe galaxy formation has drawn its material [39]. This means that an analysis of the chemical properties of DLA systems at different redshifts can provide valuable information about the history of galaxy evolution. DLA systems are our only source of knowledge about non-luminous material in galaxies outside the Milky Way. The fact that the QSO-DLA systems are not physically bound to the background source means that the range of redshifts that can be probed by this technique is not downwards limited by the low number of QSOs in the local Universe like e.g. the abundance of GRB at a given redshift limits the number of DLA systems probed by GRB there; the entire range of redshifts from the first quasars to the local Universe can be probed with this technique. This gives a unique opportunity to get detailed and easily comparable information about the chemical evolution of the galaxy population over a wide range of redshifts, covering a time line from a very young Universe and almost until present times.

A number of fits of such low-redshift DLA systems has been done by Meiring et al. in [24] and is shown in figure 2.4, taken from same. These spectra have been obtained by the Cosmic Origins Spectrograph mounted on the Hubble Space Telescope, as DLA systems at these low redshifts are still well into the near-UV bands in which the Earth's atmosphere is not transparent.

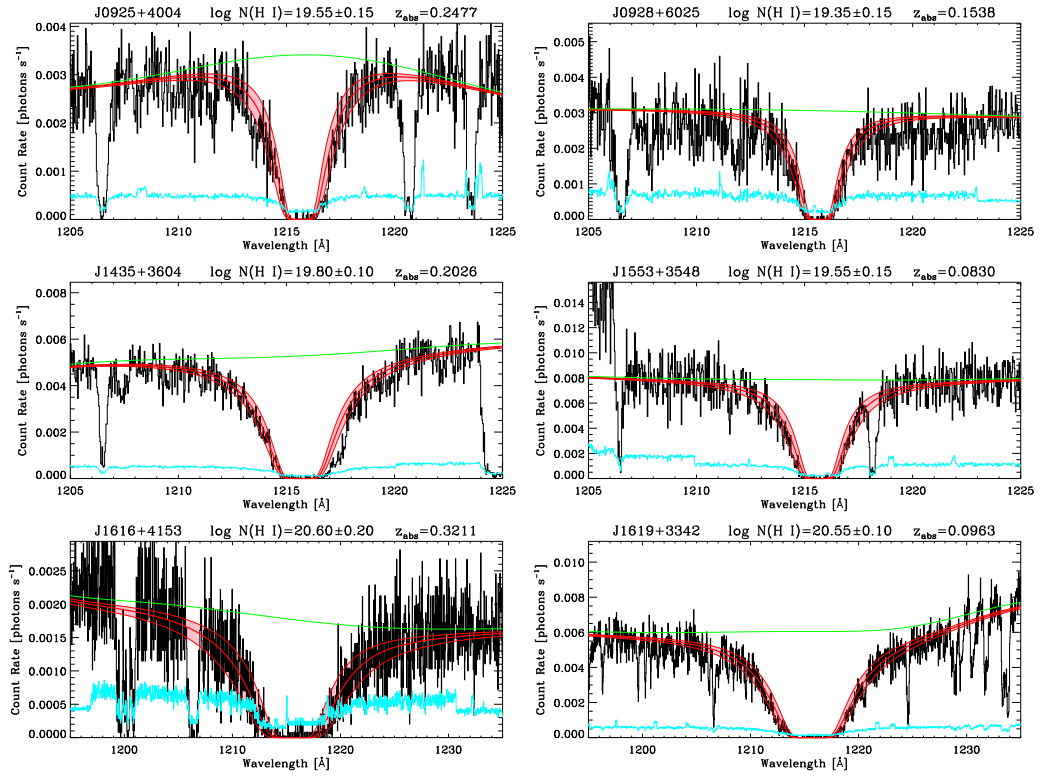


Figure 2.4: DLA line fits for a number of low-redshifts QSO-DLA, done by Meiring *et al.* in [24], from which this figure is also taken.

DYNAMIC VELOCITY AND MASS

Besides information about the chemical composition of distant galaxies, Damped Lyman α systems can also, by the velocity-space morphology of their line compounds, provide information about the dynamical structure of the system. Specifically, it is possible to find the **Dynamic velocity**, which is defined as the Doppler velocity corresponding to the spectral width of 95% of the integrated optical depth of the line (see section 4.4). The dynamic velocity plays a measure of the velocity dispersion within the system. Due to the virial theorem, and assuming there is no mechanism besides gravity correlating the velocities of the subsystems, the dynamic velocity can also be interpreted as a measure of the total mass of the host galaxy.

In [22], Ledoux et al. present data showing a correlation between dynamic velocity and metallicity for a sample of 70 DLA or strong sub-DLA systems. Their results are shown in figure 2.5

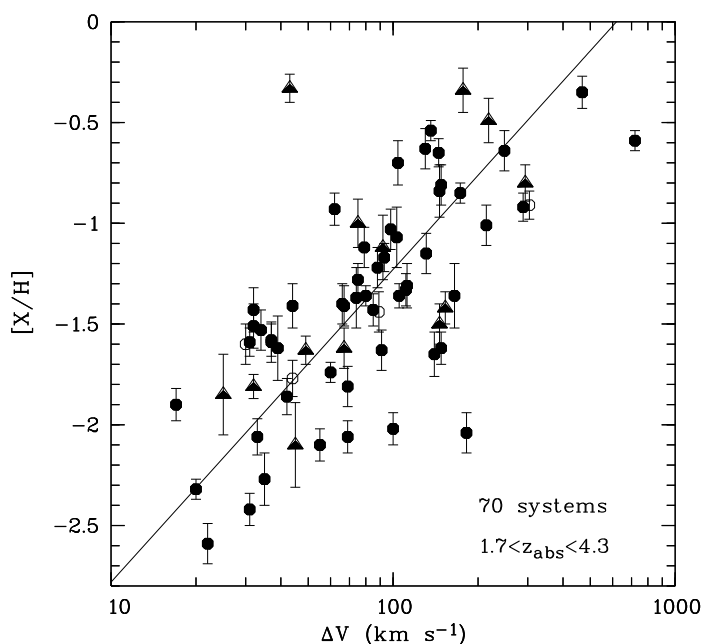


Figure 2.5: The metallicity-Dynamic velocity relation as found by Ledoux et al. in [22].

If in fact the mass-metallicity relation holds true, this also implies a mass-luminosity relation due to the luminosity-metallicity relation presented in 2.2. Assuming a CDM model, this could potentially also explain why high-redshift DLA hosts have been so difficult to observe in emission, as Cold Dark Matter models predict that galaxies at earlier stages will be less massive and hence less

luminous than galaxies in the local Universe. This would also mean that statistical information about dynamic velocity from a number of high-redshift galaxies could help constrain the fractions of hot and cold Dark Matter in the Universe.

2.1.3 VOIGT PROFILES

This section deals with some of the physics that allows to draw physical conclusions about the absorber from the shape of absorption lines. While the depth of absorption lines provides information about the number of this exact ion encountered in the cloud and hence about the chemical composition of the cloud; the width and more general shape of the line feature tells of the size of the entire system as well providing hints about its internal structure.

An atom in a dilute HI gas has a certain probability of absorbing a passing photon proportional to the interaction **cross section area**, denoted σ . This probability is, for a single atom at rest and with energy levels with an infinite life time, a delta function $\delta(\lambda)$ centered around the transition wavelength λ_0 , but this distribution is broadened by two mechanisms.

The first, **natural broadening**, stems from the fact that, due to the finite life time of the electron energy state and Heisenberg's uncertainty relation $dE \cdot dt > \hbar/2$, there is a finite probability that an atom will absorb a photon that is somewhat detuned from the nominal transition wavelength. The probability for this to happen at a given wavelength is described by a **Lorentz distribution** in λ . A generic expression for the Lorentz function can be found on [38] as:

$$L(x, \gamma) = \frac{\gamma}{\pi(x^2 + \gamma^2)} \quad (2.3)$$

The distribution can be seen in figure 2.6, along with the Gauss and Voigt profiles described below.

The above description is assuming the atom is fixed in the rest frame of the absorbing cloud, but in general, any system has some internal motion. In the reference frame of an atom moving within the cloud, a photon of wavelength λ_0 is generally red- or blueshifted away from the transition wavelength, while slightly detuned photons might be red- or blueshifted closer to (or further away from) λ_0 , and hence have an increased or decreased probability of interacting with the atom relative to its Lorentz probability of interacting with the same atom, had it been fixed. This kinetically line broadening is described by the **Gauss distribution**, in generic terms [38]:

$$G(x, \sigma) = \frac{e^{-x^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}} \quad (2.4)$$

The Gauss distribution is also shown in figure 2.6. This distribution is broader and less pointed than the Lorentzian near the line center at λ_0 , but it drops off to zero faster than does the Lorentz distribution, which stays non-negligible further out.

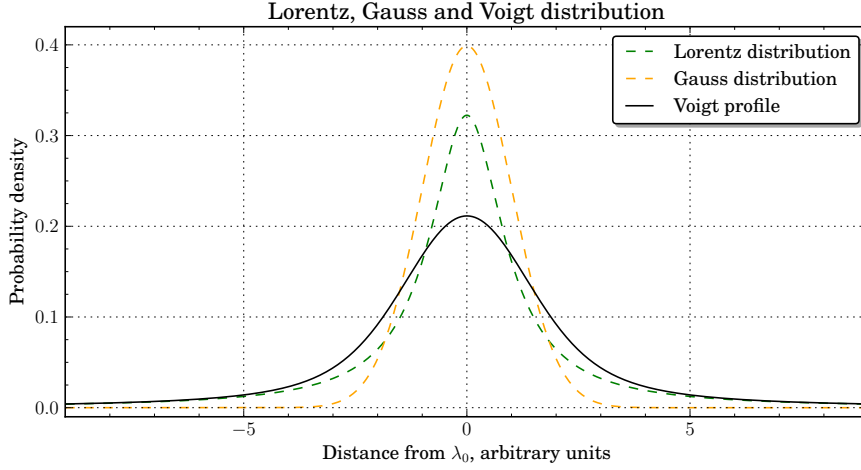


Figure 2.6: Probability density distributions for the Gauss, Lorentz and Voigt profiles with values $\sigma = 1$ and $\gamma = 1$. The Voigt distribution is dominated by the shape of the Gaussian near the line center and by the Lorentzian in the wings.

If the gas density is sufficiently high, atoms will have a high collision frequency, and kinetic energy will be distributed according to the equipartition theorem, giving each element a velocity distribution according to its atomic mass. In very dilute clouds, the inter-atom distance is generally so large that collisional interactions are very rare, and the only line broadening components will be due to turbulent motion in which all ions will have similar velocity distributions. For the systems treated here, the gas was originally assumed to be so dilute that the motion broadening is taken to be purely turbulent, but also the case of thermal broadening has been explored.

When physically combining these two broadening effects by sending photons of wavelengths near the transition wavelength through a gas in turbulent motion, a line profile emerges that is a Gaussian distribution of Lorentz profiles or, equivalently, a Lorentzian distribution of Gauss profiles. Mathematically, this resulting probability density distribution is described as the convolution of the two functions and is called the **Voigt distribution**:

$$V(x, \sigma, \gamma) = \int_{-\infty}^{+\infty} G(x', \sigma) L(x - x', \gamma) dx' \quad (2.5)$$

Since convolution is commutative, the order of the distributions is arbitrary.

Chapter 2. Methods

It is seen from the figure that the Voigt profile is dominated by the broad, round Gaussian shape around the line center, while the quick drop off of the Gaussian allows the Lorentzian to dominate far away from the line center, as was expected by comparing the two former profiles.

OPTICAL DEPTH AND ABSORPTION LINE SHAPE

The **optical depth** of a cloud at a given transition wavelength is a measure of the fraction of passing light that is absorbed. This quantity depends on the the average number of atoms any photon can expect to meet on its path through the system, measured as the **column density** N measured in cm^{-2} , and the probability of interacting with each of them, the cross section σ , which in turn depends on the oscillator strength of the given transition, the half-life of the upper level and other quantities; see [29] for more details.

The optical depth is a delta function centered at the nominal transition wavelength, which due to natural and Doppler broadening is transformed into a Voigt profile. But while the above formulation of the Voigt profile is well suited for gaining intuition for the nature and behavior of Voigt functions, it is not very suitable for scientific application. A more useful expression for the Voigt function in terms of physical quantities is given by Petitjean [29]:

$$H(a, u) = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(u - y)^2 + a^2} dy \quad ; \quad (2.6)$$

$$a = \frac{\lambda \gamma}{4\pi b} \quad ; \quad (2.7)$$

$$u = -\frac{c}{b} \left(\left[1 + \frac{v}{c} \right] - \frac{\lambda}{\lambda_0} \right) \quad (2.8)$$

Here, λ_0 is the nominal wavelength of the transition, and γ is the de-excitation rate of the higher state, which acts as a dampening parameter; and b is the **Doppler parameter**, a measure of the width of the thermal or turbulent velocities, in km/s [35].

Using the above formulation, the broadened optical depth becomes:

$$\tau(\lambda) = \tau_0 \cdot H(a, u) = 1.498 \times 10^{-2} \frac{N f \lambda}{b} \times H(a, u) \quad (2.9)$$

with τ_0 being the optical depth at the line center, for which an explicit expression is given in [29]. In terms of intensities before and after passing a material, the optical depth is defined as:

$$f = I/I_0 = e^{-\tau} \quad (2.10)$$

where I is the measured intensity and I_0 is the emitted intensity. From this, the line profile, in units of flux relative to the interpolated non-absorbed continuum at the profile wavelengths, is then simply given as $f(\lambda) = e^{-\tau(\lambda)}$, or:

$$f(\lambda) = e^{-\tau(\lambda)} = e^{1.498 \times 10^2 \frac{N f \lambda}{b} H(a, u)} \quad (2.11)$$

The right side of the last equality is from [29]. The integrals in equations 2.5 and 2.6, have no analytical solutions and must be solved numerically, a task that has been the subject of much effort. One approximation used in this work can be found in [34], and others are used indirectly through the software packages VPfit [7] and VPGuess [23]. On figure 2.7 is shown a number of Voigt profiles for fixed broadening parameters at different optical depths. All quantities on the figure are given in arbitrary units.

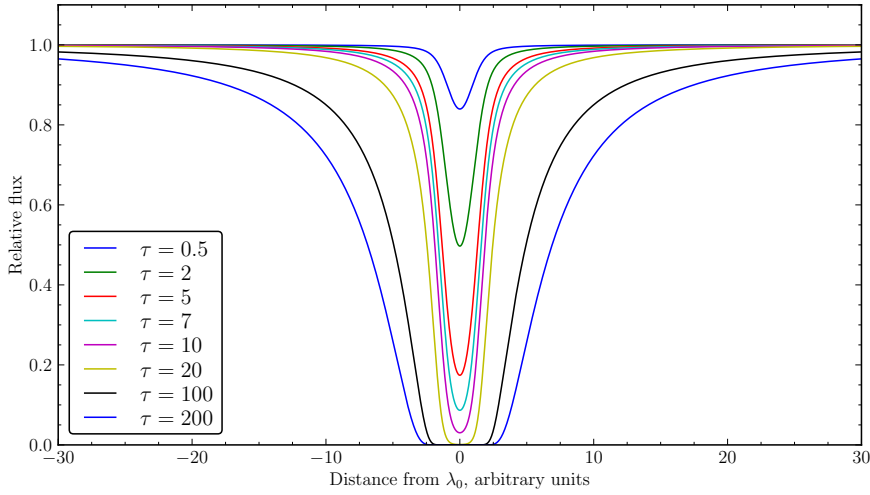


Figure 2.7: Examples of Voigt profiles for different optical depths. Note how, at a certain point, w becomes very unresponsive to changes in τ . Furthermore, it is seen that the damping wings become prominent only for high optical depths.

Physically, the shape of the Voigt profile being dominated by the Gaussian profile near the line center means that at wavelengths close to λ_0 , any photon is more likely to be absorbed by an atom in turbulent motion, since the Lorentzian profile at this point has dropped off faster than the Gaussian. At larger distances from the line center, the Lorentzian profile takes over as the dominating feature, signifying that the probability of a photon at these wavelengths to be absorbed is governed by the natural broadening, while the velocity of the absorbing atom has little influence. The transition between these two regions is more than twice the FWHM away from the line center, meaning that it takes a large column density

Chapter 2. Methods

to see any prominent damping wings, which is also clearly visible on figure 2.7. A more detailed description of this is found in [29].

COMPOUND SYSTEMS

As described in section 1.2, an absorption line system often consists of several subsystems with individual peculiar velocities and corresponding redshift components. This results in a line profile of several overlapping absorption features. Different combinations of subsystems can yield the same resulting line profile, and as is more thoroughly treated in [35], a low resolution can strongly affect the apparent shape of the line, which in turn affects the inferred values of b and N .

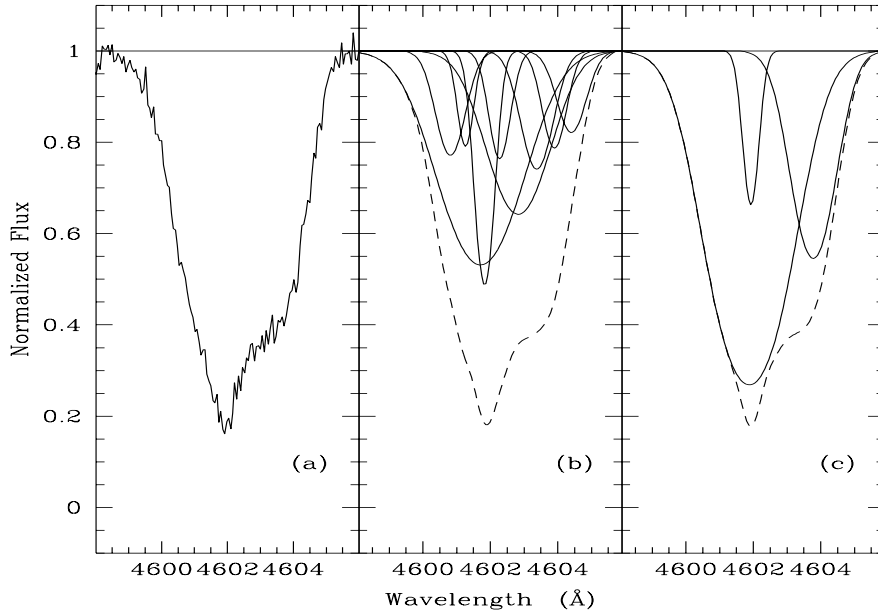


Figure 2.8: Fitting is not unique: a set of data with a fitted line profile, along with two different sets of individual subprofiles that both yield the observed profile, taken from [29].

THE CURVE OF GROWTH

An absorption line has an **equivalent width** w defined as:

$$w \equiv \int \frac{I - I_0}{I_0} d\lambda = \int (1 - e^{-\tau}) d\lambda \quad (2.12)$$

The equivalent width is the width of a rectangle of height 1 and the same area as the total flux absorbed in the line (the “area over the graph”), see section 4.2 for

examples. For a redshifted system, $w_{obs} = w_{abs}(1 + z_{abs})$. The equivalent width is easily measurable, and since it is simply an integral over a count of photons absorbed in the line relative to the interpolated nearby continuum, it is robust to distortions of the line feature due to resolution. Furthermore, w can be calculated from eqs. 2.12 and 2.6 and therefore, in theory, it should be possible to calculate N from w . Unfortunately, the relation between w , b and N requires that b be known to a very high precision.

In figure 2.9 is shown the **Curve of Growth**, taken from [29], which describes how w varies with N . The middle graph is the CoG itself, the surrounding graphs show some of the peculiarities of the three different regimes. In the optically thin regime, w does not depend on b , but only on N , and is simple to calculate by the following expression which is also used in sec. 4.2:

$$N = 1.13 \times 10^{20} \frac{w_r[\text{\AA}]}{\lambda^2[\text{\AA}]f} \quad (2.13)$$

where w_r means the equivalent width in the rest frame of absorption [29]. As can be seen on fig. 2.9, the equivalent width is unchanged even though b varies. This also means that N can be uniquely determined once w is known.

In the flat, logarithmic regime on the graph, for intermediate values of N , the situation is quite different: for the same value of N , different values of the Doppler parameter can yield very different values of w and, consequently, the equivalent width is no longer enough to determine the column density. As is shown in the middle figure, a modest change in b can cause the CoG to take widely different paths, on which an almost identical value of w correspond to strongly differing values of N . The top graph shows how, for a fixed value of N , moderate alterations of b can cause large differences in w . In this case, many transitions of the same ion are in principle necessary to accurately determine N , but mostly, as was the case for this project, there is not sufficiently many detectable transitions for each ion present in the data to make this possible.

A generally accepted solution is to assume that all ions of similar ionization degree belong to the same phase with the same physical properties, and then fitting them with tied b parameters, either as one turbulent b parameter, or for thermal contributions with per-element b parameters according to atomic masses. This assumption makes it possible to do some fairly accurate determinations of N and b and, as will be evident in 4.3, the results seem to agree well with this assumption.

2.1.4 ROBUST LOWER LIMITS

When the light of the QSO is dispersed in the spectrograph, there is a lower limit to the spectral width of line profiles that can be seen, determined by the

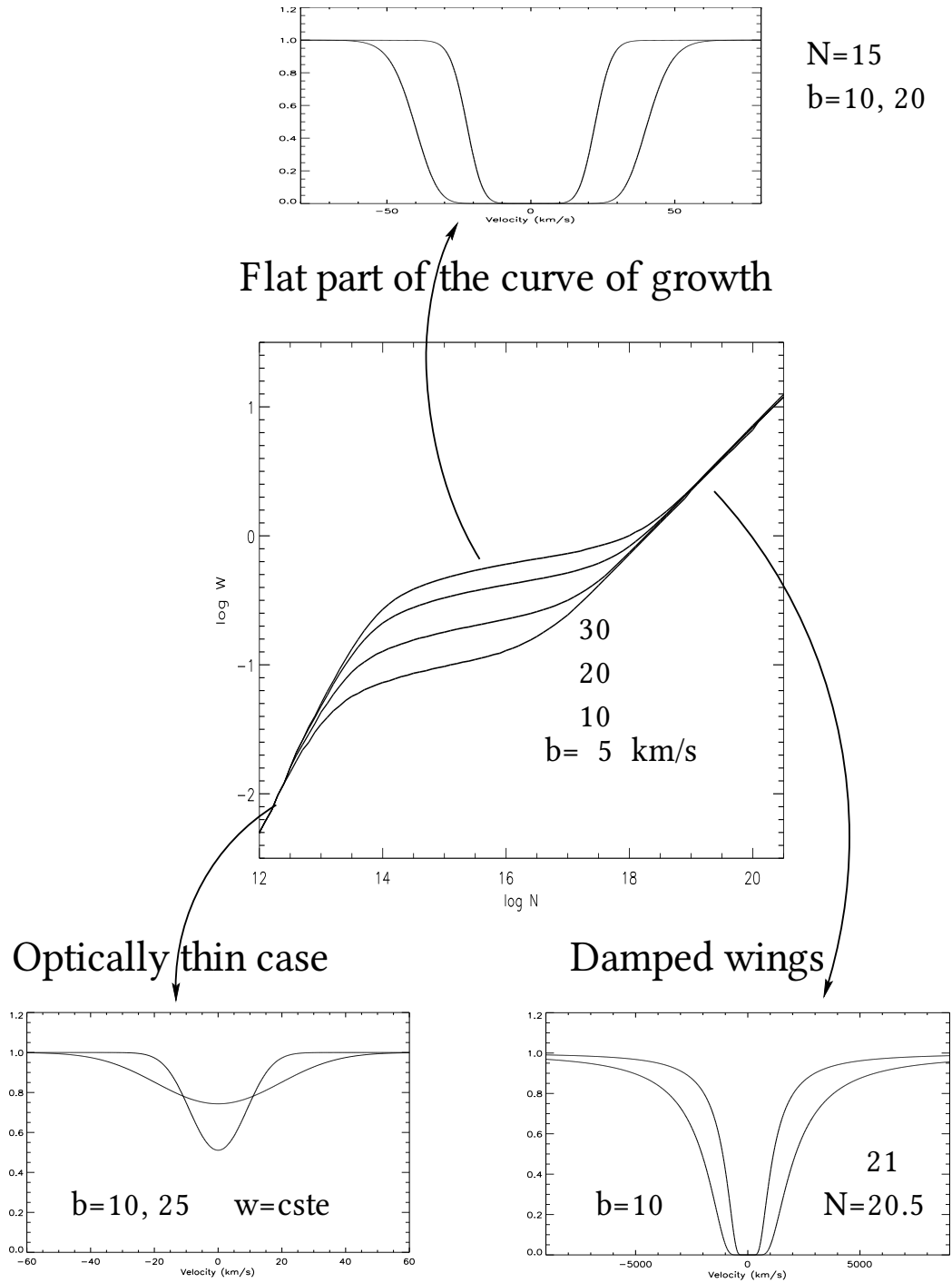


Figure 2.9: The Curve of Growth for N as a function of w . In the optically thin limit, N is uniquely determined by w , and the same is the case in the saturated limit. In the intermediate region, N depends strongly on b , and determination of N by means of w becomes highly uncertain. The figure is adapted from [29].

spectral resolution of the instrument. If there would happen to be strong absorption lines present with a spectral width $\ll R$, these lines would be smoothed by the resolution to appear weaker and broader than they really are; making a narrow, saturated line appear as broader, optically thin feature. Since, as we have seen, the relation between equivalent width and Doppler parameter is different in the optically thin, intermediate and saturated regimes, such a **hidden saturation** could lead to an incorrect estimate of w and hence the metallicity of the system. Of course, the CoG method is supposed to minimize this by fitting b for multiple lines of the same phase, but it would be preferable to have a different, less model dependent way of determining the metallicity of the system.

One way to do this is to take advantage of the facts that w is directly measurable and independent of resolution, and the column density is independent of b in the weak line limit. By picking a reasonable line and measuring w directly, N can be calculated directly from eqn. 2.13

As can be seen from figure 2.9, the assumption of the line being optically thin corresponds to assuming that the roughly linear relation in the weak line regime is generally valid, yielding a value of N which is either close to the true value or underestimated. Hence the name of the metallicity calculated by this method, the **robust lower limit**. This model-independent value can, apart from a strong lower constraint, also serve as a sanity check of the models reached through line profile fitting.

The line should be chosen weak enough that it can be safely assumed that the weak-line approximation is still valid, but on the other hand not too weak, making the measurement of w too statistically uncertain. Results are shown in section 4.2

The software package `vpguess`, described in section 4.3, has a deconvolution feature, in which resolution smoothing of the modeled spectrum can be turned on and off, showing the “real” profile without instrument resolution blurring. This can be helpful to see if it looks like there is any lines that are saturated after deconvolution, or any suspicious looking near-misses; but it cannot rule out the possibility of hidden saturation, since information is lost in convolution that cannot be recovered by deconvolution. It is assumed in `vpguess` that one line component at spectral resolution corresponds to one physical component, which is not necessarily the case.

However, for what it is worth, turning off instrument convolution of the modeled profile of the spectrum of Q 0918+1636 did not reveal any hidden saturation for any of the lines fitted and used for metallicity determination.

2.2 SAMPLE SELECTION

The spectra for the present work are taken as a part of a survey carried out by J. Fynbo and collaborators, presented in [10] and [11], the primary aim of which is to detect emission from galaxy counterparts of high metallicity Damped Lyman α Absorbers at redshifts $\sim 2 - 3$. The absorbers are selected for a large equivalent width in the SiII 1526, which is known to be a reliable proxy for metallicity [30].

It has long been expected that high-redshift DLAs are young, star-forming galaxies [39]. Surveys have shown that the DLA clouds contain the bulk of neutral hydrogen in the early Universe, and metallicity measurements have shown a metal enrichment of many DLA systems indicating that star formation has already enriched the clouds. Galaxies hosting DLAs are often very difficult to detect, because they are very faint and, according to currently favored DM theories, relatively small structures, typically the size of modern day dwarf galaxies.

Especially for DLAs in QSO spectra, direct imaging is very difficult, because the QSO will most likely be magnitudes more luminous than the DLA hosting galaxy, and the impact parameters most often comparable to the size of the QSO point spread function. However, galaxies with strong star formation will have a strong emission in the Lyman α line, which could be visible as a small peak at the bottom of the broad trough of the DLA absorption feature. These emission lines have been found, but in a surprisingly small number of the observed DLAs.

P. Møller et al. announce in [26] one such Lyman α emission line within the broad DLA absorption line, and in the same work, they propose that there exist a metallicity-luminosity relation at these redshifts similar to the one that is known to exist in the local Universe [37]. They show data that, albeit too sparse to be conclusive, are consistent with and weakly supporting the hypothesis that more luminous galaxies have higher metallicities. The authors suggest that observing a number of high-metallicity DLAs at redshifts $\sim 2 - 3$ could likely give an answer to this.

J. Fynbo and collaborators make a similar prediction in [12]. The authors set out in an effort to give an explanation to why DLA in Gamma Ray Burst spectra have higher mean metallicity than those of QSO-DLA, as shown in figure i2.10. Assuming that the GRB galaxies and QSO-DLA hosts are drawn from the same group of objects, they attempt to develop a model consistent with known properties of these objects, in which a random sample of GRBs and a random sample of QSO-DLA would show properties similar with those observed and shown in figure 2.10.

If, with a given luminosity function $\Phi(L)$, we draw a random sample of QSO-DLA hosts and GRB-hosts in a given patch of the sky, the contribution to the entire sample of luminosity bin $[L, L + \Delta L]$ can first be determined. For GRBs, this is fairly simple. The GRB-hosts are selected by their star formation rate per

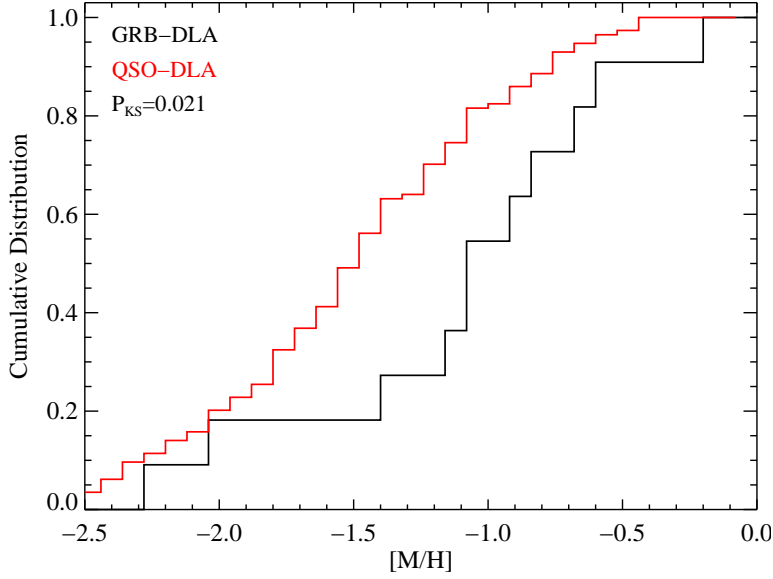


Figure 2.10: Cumulative distribution of metallicities for GRB-DLA and QSO-DLA. The metallicities of GRB-DLA are on average significantly higher than those of QSO-DLA. Figure from [12].

luminosity bin, which is proportional to its rate of GRB events. Hence, contribution of the bin $[L, L + \Delta L]$ of the sample is $f_{GRB}(L) \propto SFR(L)\Phi(L)$. The star formation cannot be measured direct, but it is assumed that it is proportional to the luminosity of the galaxy, since the majority of emission detectable from these galaxies is assumed to be due to young, bright stars; and hence, we have that $f_{GRB} \propto L\Phi(L)$.

Unlike GRBs that happen in the central regions of their host galaxies, QSOs are not physically connected to the DLA system, but can be taken to be point events randomly scattered across the sky patch. The probability of detection is determined by the cross section area of the DLA cloud. Hence, DLA hosts are **cross section selected**; their probability of intersecting the line of sight to the QSO is $f_{QSO}(L) \propto \sigma_{HI}(L)\Phi(L)$. To establish how that influences the average luminosity of QSO-DLA hosts, the authors of [12] adopt the *Holmberg relation* of the local universe, in which the radius of a galaxy is related to its luminosity as $R \propto L^{0.4}$. The cross section area, and hence the probability to intersect the QSO line of sight, is then related to the luminosity function as $f_{QSO} \propto \Phi(L) \times L^{0.8}$. The average luminosity of a QSO-DLA host is, therefore, less than the average luminosity of the GRB hosts.

To see what that means to the metallicity distribution, a local Universe luminosity/metallicity relation of the form $Z \propto (L/L_*)^\beta$ is adopted, meaning that a

higher luminosity indicates a higher metallicity, since $L \leq L_*$.

Furthermore, while GRBs practically always happen in the central region of their host galaxies [9]; QSO lines of sight can trace any path through the galaxy hosting the DLA systems. Since metallicity is determined by SFR, the central regions of a galaxy is expected to have higher metallicity than the outer parts of the same galaxy. QSO-DLA, which generally have a higher impact parameter than GRBs, must be expected to show lower metallicities than GRB-DLA, even if the sample of DLA host galaxies were assumed to all have the exact same metallicity.

The results of simulating GRB and QSO data through this model is shown in figure 2.11, also taken from [12]. In the bottom panel is shown a predicted cumulative luminosity distribution; as predicted, a random sample of GRB hosts are expected to have a greater luminosity than a random sample of QSO-DLA hosts. As always, high magnitude means low luminosity.

The middle panel shows the relation between impact parameter and luminosity for three different, simulated samples of QSO-DLA. Fainter DLA hosts have lower impact parameters due to the Holmberg relation between radius and luminosity. Low metallicity hosts have lower luminosity due to the Metallicity/-Luminosity relation and from that lower impact parameters due to the Holmberg relation. Impact parameters are assumed to be zero for GRBs.

Finally, the upper panel shows the simulated relation between metallicity and luminosity. The GRB hosts lie on a line, since their impact parameter is zero, making the measured DLA metallicity an accurate tracer of the host maximum metallicity, which is assumed to be uniquely defined by the luminosity. The QSO-DLAs are scattered on and below this line, since any non-zero impact parameter will mean a detected metallicity lower than the maximum metallicity of the host. Note that this model clearly predicts that higher metallicity DLAs will have higher luminosities, in agreement with the predictions of Møller et al. in [26].

Prediction of a generally higher luminosity for high metallicity objects has been the motivation for the sample selection for the survey in [10] and [11], which has so far shown consistency with this prediction, although more data are needed to be conclusive. The DLA treated in this work is, however, not among those showing host galaxy emission lines.

2.3 ERROR PROPAGATION

All errors in this work are propagated under the assumption that they are randomly distributed and uncorrelated, so the error of a given operation is propagated as:

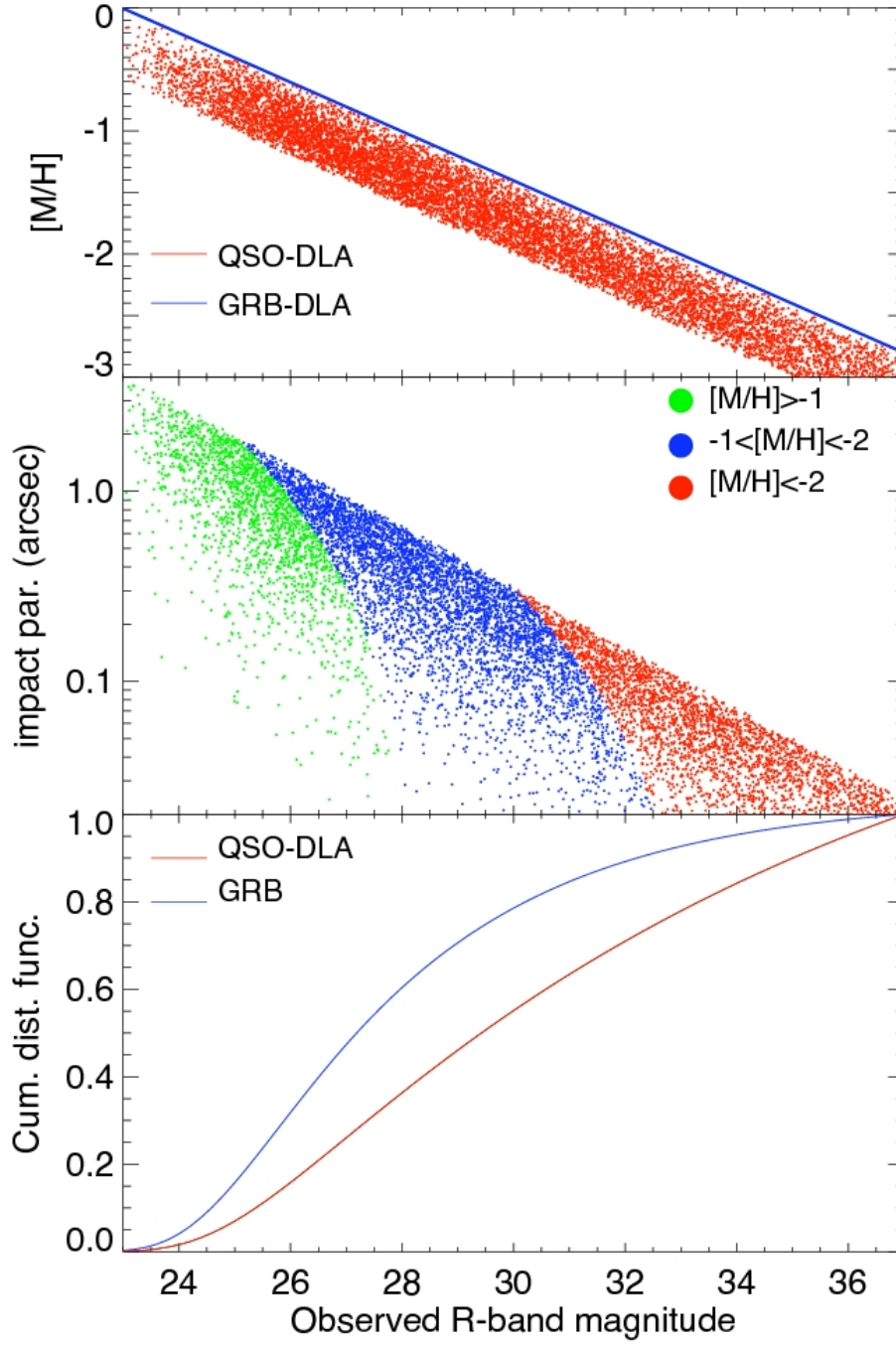


Figure 2.11: The main results of [12]. Top panel: simulated metallicity-luminosity measurements from a sample of QSO- and GRB-DLA. Middle panel: Simulated luminosity-impact parameter relation for three random samples of QSO-DLA with different metallicities. Bottom panel: predicted cumulative luminosity distribution of QSO- and GRB-DLA.

$$\delta_x = \sqrt{\left(\delta_u \frac{\partial x}{\partial u}\right)^2 + \left(\delta_v \frac{\partial x}{\partial v}\right)^2 + \dots}, \quad (2.14)$$

where

$$x = f(u, v, \dots)$$

is a function of the variables u, v, \dots [32].

Error propagation in the X-Shooter pipeline is based on the same assumptions and described shortly in [16]. Error propagation in `vpfit` is documented in the source code, the analysis of which is outside the scope of this project. Apart from these two steps, the only step for which error propagation is non-trivial is the calculation of metallicities from the column densities found by `vpfit`, which shall be described here.

For a start, let $\log(N_i)$ denote the logarithmic column density of an ion in the i 'th velocity component of the line, and $x_i = \delta \log(N_i)$ be the uncertainty of $\log(N_i)$ as given by `vpfit`. Here, $\log(u)$ denotes the logarithm base 10 of u , and $\ln(u)$ shall denote the natural logarithm of u .

The metallicity for the given ion is then found as $Z = \log(\sum_i 10^{\log(N_i)}) - \log(N_{HI})$, and the error propagation step for step is done as follows.

First, the uncertainty for N_i :

$$\delta N_i = x_i \times \ln(10) \times 10^{\log(N_i)} \quad (2.15)$$

$$= y_i \quad (2.16)$$

where y_i is a convenience renaming of the result of the preceding calculation. The next step is a simple summation of the column densities with a Pythagorean summation of the uncertainties:

$$N_{ion} = \sum_i N_i \quad (2.17)$$

and

$$\delta N_{ion} = \sqrt{\sum_i y_i^2} = z \quad (2.18)$$

Next step is taking the $\log(N_{ion})$, and the corresponding propagated uncertainty is:

$$\delta \log(N_{ion}) = \frac{1}{N_{ion} \times \ln(10)} \times z = \sigma_{ion} \quad (2.19)$$

Finally, the ion column density is expressed in terms of the HI column density by a simple subtraction, and hence the last step in the error propagation is yet a Pythagorean summation:

$$Z_{ion} = \log(N_{ion}) - \log(N_{HI}) \quad (2.20)$$

and

$$\delta Z_{ion} = \sqrt{\sigma_{ion}^2 + \sigma_{HI}^2} \quad (2.21)$$

with σ_{HI} denoting the uncertainty on the column density of HI found during the DLA fitting as described in section 4.1.

This procedure for calculating metallicities and propagating uncertainties is implemented in the script `Metallicities.py`, included in appendix B.4; along with a parallel propagation done by the Python module `uncertainties`. Reassuringly, the two implementations yielded identical results.

CHAPTER 3

DATA REDUCTION

3.1 THE X-SHOOTER

The X-Shooter is a single target spectrograph, mounted on the Unit Telescope 2 *Kuyen* of the Very Large Telescope Array (VLT) in Paranal, Chile. The spectrograph is designed to cover a wide spectral range, from the UVB to the K band, at an intermediate spectral resolution.

The light is split into a near-ultraviolet/blue (UVB), a visible/red (VIS) and a near-infrared (NIR) arm, each equipped with optimized coatings, detectors etc. The arms cover the wavelength ranges 300-550 nm in the UVB arm, 550-1000 nm in the VIS arm and 1000-2500 nm in the NIR arm, at resolutions ranging from 4000 to 14000, depending on slit width and wavelength [1].

For the observations treated in this work, the nominal resolutions are 4000, 6700 and 4300 for slit widths of $1.3''$, $1.2''$ and $1.2''$ for the UVB, VIS and NIR arms, respectively; although, as we shall see, the true resolution is higher.

In fig. 3.1 is shown a schematic overview of the X-Shooter, found in the software manual for the X-Shooter data reduction pipeline [1]. After entering the spectrograph, the light is sent through two dichroics, which redirect first the UVB and then the VIS light to their respective arms, before the remains enter the cooled-down NIR arm. Both the UVB and VIS arm have correction for atmospheric dispersion, and all three arms have piezoelectric mirrors to correct mechanical deformation of the instrument and setup [8].

Each of the three arms has its own shutter and masks and hence, exposure times can be controlled individually for each of them.

Each of the arms have a *slit* and an *IFU* configuration, of which the former was used. Each of the arms have a set of different reads-out and binning modes, an overview of which is given in table 3.1.

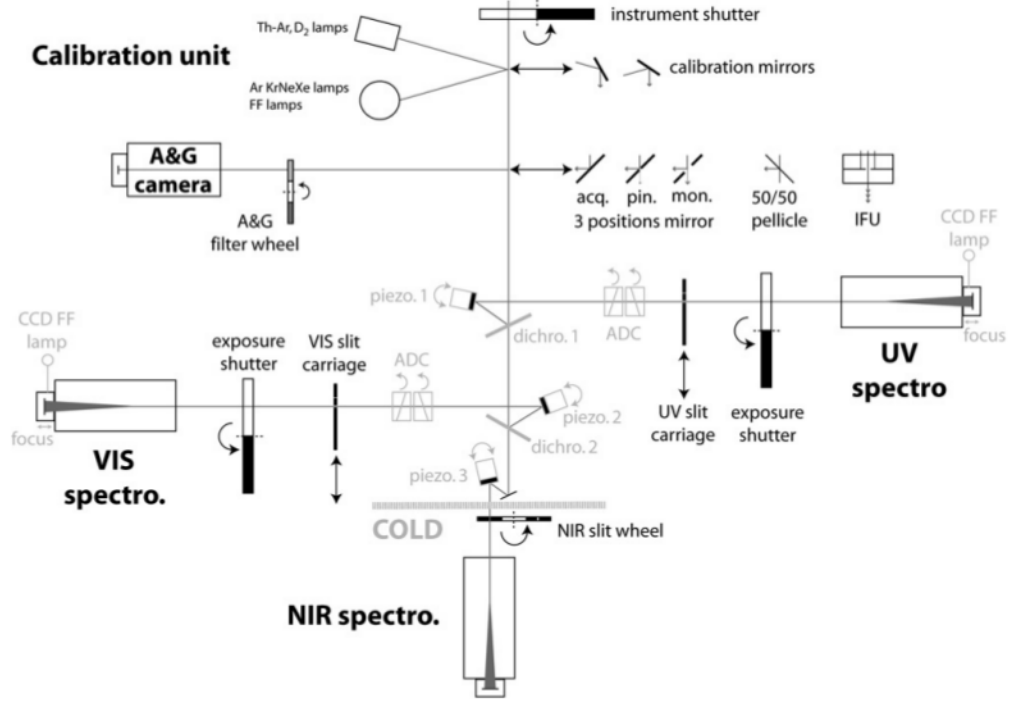


Figure 3.1: The structure of the X-Shooter spectrograph from [1]. The light is led through two dichroics that sends first the UVB/Blue band light into the UVB arm, then send the visual/red light into the VIS arm before the rest enters the cooled NIR arm.

3.2 OBSERVATION CONDITIONS

The data for this project are the same as were used in [11], targeting the $z = 3.07$ quasar SDSS J 091826.16+163609.0 (hereafter Q 0918+1636). This quasar was chosen for X-Shooter spectroscopy because it was known to contain the DLA treated in this work, located at a redshift $z_{abs} = 2.412$, but what is treated in [11] is a second, even more metal rich DLA system at a redshift of $z_{abs} = 2.583$, unexpectedly found in the same spectrum and also fulfilling the sample selection criteria as described in sect. 2.2. While the $z = 2.583$ system also showed emission from a galaxy counterpart in the spectrum, this is not the case for the system treated in this work.

The quasar was observed with the X-Shooter on the VLT on February 16th, 2010, on three different position angles in order to maximize the probability of catching any nearby galaxy counterparts of the DLA systems in the slit. It has been shown by Fynbo et al. [10] that at this arrangement, slightly more than 90% of galaxy counterparts will be found within the slit in at least one of the exposures, assuming the model in [12] is correct. Therefore, three exposures has been

POSSIBLE X-SHOOTER CONFIGURATIONS			
Arm	Mode	Binning	Read-out
UVB	Slit/IFU	1×1 , 1×2 , 2×2	100k/400k
VIS	Slit/IFU	1×1 , 1×2 , 2×2	100k/400k
NIR	Slit/IFU	1×1	-

Table 3.1: Possible instrument settings for the X-Shooter spectrograph

chosen as the best compromise between chance of finding the galaxy counterpart and the constraints on observing time imposed by a tight time budget. At each of the position angles the QSO was exposed for 3600s through a 1.3 slit for the UVB arm, and a 1.2 slit for the VIS and NIR arms.

3.2.1 RESOLUTION

The spectral resolution $R = \lambda / \delta\lambda$ of any spectrum is - all else equal - ultimately determined by the extent of the source along the direction of dispersion. This is most often simply determined by the width of the slit, since this will in most cases be the smaller of the two.

The nominal resolving power determined by the slit width of the X-Shooter arms is $R_{\text{UVB}} = 4000$, $R_{\text{VIS}} = 6700$ and $R_{\text{NIR}} = 4300$ for the UVB, VIS and NIR arms with slit widths of 1.3'', 1.2'' and 1.2'', respectively. However, the seeing was, at a mere $\varphi = 0.67$ in all three exposures, considerably better than the slit width would suggest. The true resolution can be measured directly from the width of the telluric absorption lines as sketched below, and was determined to be $R_{\text{VIS}} = 11900$ in the VIS arm.

The resolution of the NIR and UVB arms was then found by assuming the same ratio between expected and true resolution for all three arms. This gave resolutions for the remaining two arms of $R_{\text{UVB}} = 6400$ and $R_{\text{NIR}} = 8800$.

The situation is sketched qualitatively in figure 3.2. The box in the upper left represents the slit geometry, with the axis of dispersion oriented horizontally. The large gray disc represents the 1σ radius

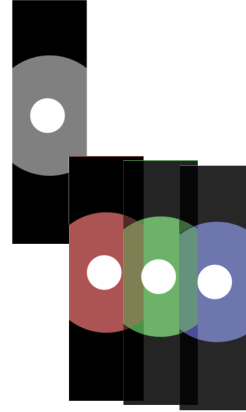


Figure 3.2: Sketch of how seeing better than the slit width can improve the true spectral resolution. The box represents the slit, the gray and white disks represent a point source at different seeing conditions.

of the point spread function, or the **seeing disc**, of a point source under conditions in which the seeing is wider than the slit width. The smaller, white disk is the seeing disk of the same object at seeing conditions better than the slit width. The images in the lower right shows the refracted image of the slit and object at three different wavelengths. While the slit images are already overlapping and blending into each other, the small seeing disks are still fully resolved. The limit width at which the images of the the large seeing disks are blending is exactly the width of the slit. In this case, the spectral resolution is determined by the slit width, which for the VIS arm is $1.2''$. The limit width for the she smaller seeing disk, however, is given by the radius of the disk itself. Here, the spectral resolution is determined by the seeing, which in our case was $0.67''$.

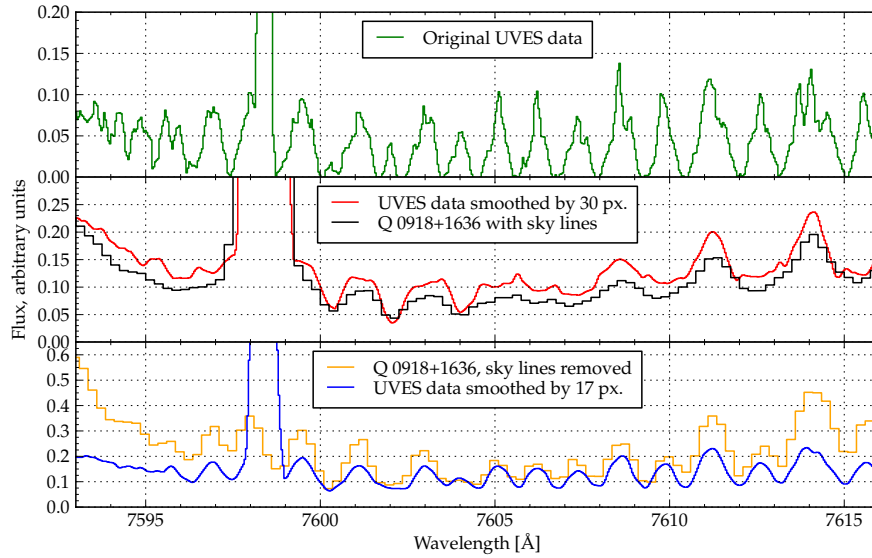


Figure 3.3: *plot of high-resolution sky and telluric lines, smoothed to match the slit and the seeing, respectively. Skylines and telluric lines for our data are shown for comparison. Only the width of the lines is accurate, the strength is altered for illustrational purposes.*

Figure 3.3 can help illustrate how this can help determine the effective resolution from the data themselves. A region is chosen in which can be found a single sky emission line, a sky continuum and telluric absorption lines. Since the sky lines are due to glow of the entire sky, their width are determined by the slit width. The telluric lines, on the other hand, are due to absorption in the source spectrum, and their spectral resolution is therefore determined by the size of the seeing disc, if this is smaller than the slit. In the upper panel is plotted the UVES sky line map [15], a high resolution atlas of sky emission lines at optical wavelengths. The resolution of the UVES data is high enough to serve as a point source

for this purpose. In the middle panel, a spectrum of Q 0918+1636 is plotted, in which the sky line removal of the X-Shooter sky line removal has been disabled. With it is shown the UVES data smoothed with a boxcar kernel that gives it a resolution similar to the sky lines - trial and error gave a best width of 30 pixels. It is plainly visible that the UVES data convolved with this resolution is a good description of the actual data seen. Note, especially, how the very strong sky line at $\sim 7597\text{-}7598 \text{ \AA}$ fits well with this resolution.

In the lower panel is plotted the now sky-cleaned quasar spectrum, along with the UVES data, this time smoothed with the kernel width that best imitated the quasar data. Similar to how the 30 pixels kernel corresponded to the width of the slit, this kernel corresponds to the resolution of the quasar spectrum, which is the smaller of the slit width and the seeing disc. In the figure above, I estimated a best smoothing kernel width of 17 pixels' width. The ratio of these widths then corresponds to the ratio of seeing disc and slit width, which again equals the ratio of nominal and true resolution:

$$\begin{aligned} \frac{17}{30} &= \frac{\varphi}{1.2''} & \Rightarrow \\ 0.57 \times 1.2 &= \varphi = 0.68'' \end{aligned}$$

and, for the resolution

$$30/17 \times 6700 = 11824,$$

all in good agreement with the values found in [11].

3.3 EXTRACTION AND NORMALIZATION

This section describes the steps required to prepare the data for analysis of the absorption line features relative to the quasar continuum emission. The main steps are cleaning, reduction and rectification with the X-Shooter pipeline, stacking, σ -clipping and extraction of the spectra into a one-dimensional (λ)-space, and normalization by aid of a standard quasar spectra and modeling of various extinction mechanisms. The result is a spectrum expressed in flux relative to the quasar emission.

3.3.1 THE X-SHOOTER PIPELINE

Raw science frames from the X-Shooter consist of a number of curved orders with tilted sky lines, giving a non-trivial relation between wavelength bins and

Chapter 3. Data Reduction

CCD pixels. Apart from normal image reduction routines like sky line and bias subtraction, flat field reduction etc., it is necessary to straighten and merge the orders.

Software and calibration files for this are provided in the **X-Shooter pipeline**. The pipeline consists of a number of so-called *recipes* which are callable either by the graphical tool `gasgano` or, as was the case in this work, the command-line tool `esorex`. An extensive manual for the X-Shooter pipeline exists [1], on which the following short outline of the reduction process is mainly based.

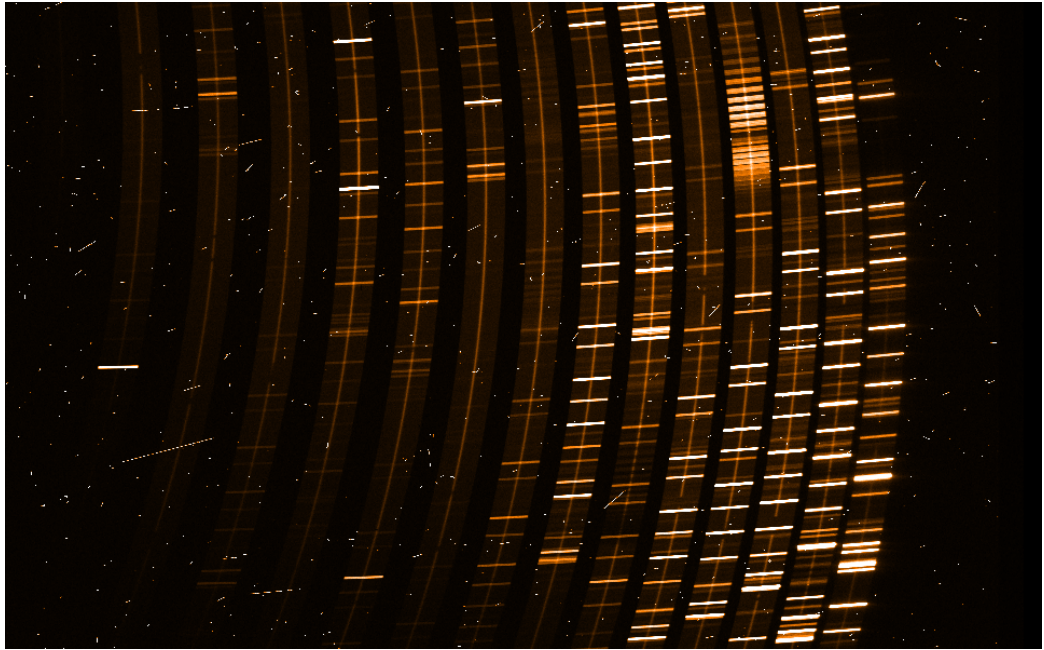


Figure 3.4: Section of a raw science *frt*am from the X-shooter. The orders are bent and overlapping, and since the science object is very faint, the sky lines are strong in comparison.

Before running the pipeline, the data must be classified by data-type (i.e. BIAS, DARK), arm (i.e. UVB, VIS or NIR) and group (i.e. specific observation group or calibration template). File names and data types are then listed into a *set-of-files* `.sof` file for each recipe for each arm. The data type of each file is stored as keywords in the FITS header. Data type sorting is handled automatically in `gasgano`, but it is also a fairly easily scriptable task, which in this case was done as a Bash-script calling the ESO tools `dfits` and `fitsort`.

For obtaining the spectra for Q 0918+1636, the instrument was run with the slit configuration, stare mode, and the pipeline ran in the *poly* mode, in which the instrument is modeled with polynomial fits. Here is a listing of the recipes that were used for this project

- xsh_util_physmod** Takes as input a table of values for the X-Shooter model configuration and a reference table of known lines and generates a table of expected positions for these lines.
- xsh_mbias** Takes as input a number of raw *bias* frames and, optionally, a frame mapping bad pixels, and generates a master bias frame. Each of the bias frames is a readout of the CCD at zero exposure time. Bias frames should match any images they are subtracted from in binning and readout mode, so a master bias frame is to be generated for any number of combinations relevant. In this case, a master bias was made for the combinations of binning 1×2 ; readout mode 100k and binning 1×1 ; readout mode 400k. No bias frame is created for the NIR arm, since the dark current is dominating there.
- xsh_mdark** Creates the master *dark* frames, which are used to correct for the signal generated by the current within the CCD. The raw DARK images are taken as an exposure the shutter closed. Dark frame generation is only done for the NIR arm, since the contribution for the other arms is negligible.
- xsh_predict** Makes a preliminary prediction of the position of the orders and overlaps areas. It uses the input guesses *th* recipe, and at were also fed to the *xsh_util_physmod* recipe and the output of the same, and outputs a model (either a polynomial fit or an enhanced physical model, depending on the reduction mode) for finding the position in (x, y) coordinates on the CCD for for the n 'th dispersion order of wavelength λ for a point in the center of the slit. This will serve as the basis for mapping of the entire slit at order n , wavelength λ at a later stage.
- xsh_mflat** This recipe creates the master *flat* frame, which is used to correct for the varying efficiency of the detector. Five flat field frames, created by exposure to a continuum lamp, are prepared in order to eliminate statistical outliers. As in standard procedure, master *bias* and *dark* frames are corrected for, before applying the flat field correction. For the UVB arm, two different lamps are used, in order to cover the entire spectral range of the UVB arm with a continuum.
- xsh_2dmap** This recipe uses the wavelength maps from earlier, together with a frame exposed through a multi-pinhole mask, to create a 2-dimensional mapping of the slit onto the CCD in (x, y) coordinates for the n 'th order of wavelength λ .
- xsh_scired_slit_stare** This is the only receipt that works on the science frame(s). After having prepared all necessary tables, frames and maps, here is where

the reduction and correction is done to the science frame. The maser bias is subtracted (for UVB and VIS arms), the master dark is subtracted for the NIR arm. The science frame is divided by the master flat field, and the sky background is estimated and removed if this option is passed. Finally, based on the 2-dimensional order and wavelength map, the measured fluxes are remapped from pixel (x, y) space and into a regular grid in (λ, s) space.

For each arm, a script was made in the bash scripting language to carry out the pipeline recipes. An example script, `makeitso-VIS.sh`, is included in appendix B.1. The script assumes that the `.sof` files already exist (their file names are hard-coded into the script), but to change this is fairly easy, as would a `.sof` creating addition.

The script is called by a number of switches, each of which calls one recipe of the pipeline, and an argument telling on which science frame(s), including the telluric standard star frame, to operate. A few convenience aliases makes it possible to easily run all preparation steps only, the science step only, and whether to turn on or off sky subtraction (since both options were needed at some point). The recipes were called with a set of switches and parameters for fine-tuning the processes, found in a chain file provided by Johan Fynbo; these switches and options are shown in the source code of the script, included in appendix B.1.

3.3.2 WAVELENGTH CORRECTION

After pipeline reduction, there is a slight wavelength offset between the UVB and VIS arms. This is corrected by creating a version of the reduced science frames from the arms involved with no sky background removed. These spectra are then compared with the high-precision data from the UVES Sky Atlas [15], and if they show any discrepancy from this, value of the reference pixel keyword `CRVAL1` in the FITS header of the relevant spectra is changed to match the UVES data.

As was visible in figure 3.3, the raw UVES data do not resemble the data at hand very much, due to the different resolutions of the spectra. To emulate the significantly lower resolution of the X-Shooter, the UVES data are convoluted with a boxcar kernel with the same width as the slit image would have on the CCD of UVES. Qualitatively speaking, the choice of a boxcar kernel with this width can be explained as what the image of the sky emission on the CCD would look like in the λ direction; a uniform brightness the width of the slit, abruptly cut off at the sides. To model the resolution of the QSO image, a Gaussian kernel would be more accurate, as the PSF of the quasar is a Gaussian, not a boxcar function. The difference is shown in figure 3.5.

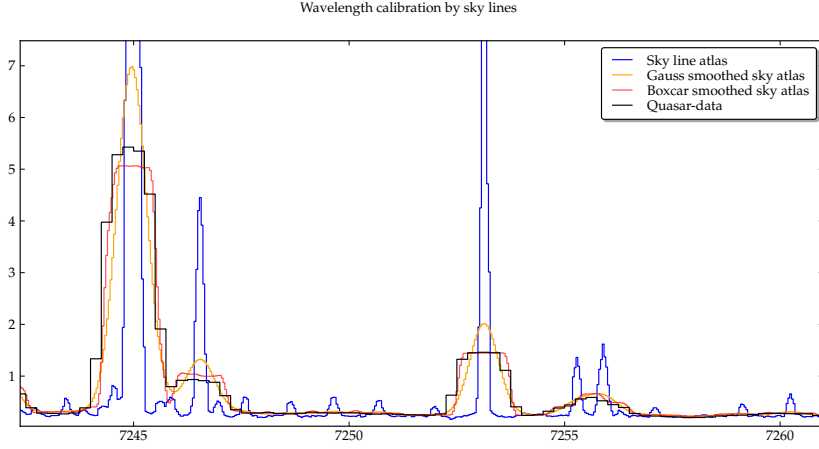


Figure 3.5: Section of VIS arm of Q 0918+1636, with the UVES Sky Atlas in its raw form and smoothed with a boxcar kernel and a Gaussian kernel for comparison.

3.3.3 STACKING & EXTRACTION

The X-Shooter pipeline outputs a 2D spectrum for each of the input files and can, if fed more than one science frame, also do a stacking of the science frames. However, the stacking method of the pipeline is taking a pixel-wise median from the three images. This method ensures a very nice and “clean” image, since any pixel is extremely unlikely to have random noise in more than one of the frames; but it does discard information, resulting in a poorer signal-to-noise ratio than necessary. The S/N ratio of a pixel grows as the square root of the counts in said pixel, so stacking all (good) pixels, utilizing every photon possible, makes for better certainties. A simple per-pixel average would on the other hand not be desirable either, since cospics are weighed equally along with good quality pixels.

Instead was used the so called **sigma clipping method**, in which any pixel deviating more than 3σ from the per-pixel median is discarded, and the per-pixel average is taken of the rest. If (x, λ) denotes the coordinates of a pixel in the directions perpendicular and parallel to dispersion, respectively, then $I_{x\lambda}$ and $\sigma_{x\lambda}$ are the measured flux and standard error in the x, λ 'th pixel. The approach will then be as follows:

1. Construct the median image by finding the per-pixel median of the three frames.
2. In each of the original images, find all pixels that deviates from the corresponding median pixel by more than 3σ .

Chapter 3. Data Reduction

3. For these pixel coordinates, the per-pixel average of the two remaining images are calculated and saved, as is the per-pixel average of the variance $v = \sigma_{x\lambda}^2$.
4. For pixel indexes (x, λ) where no values are rejected, an average of pixel value and variance of all three images is calculated and saved.

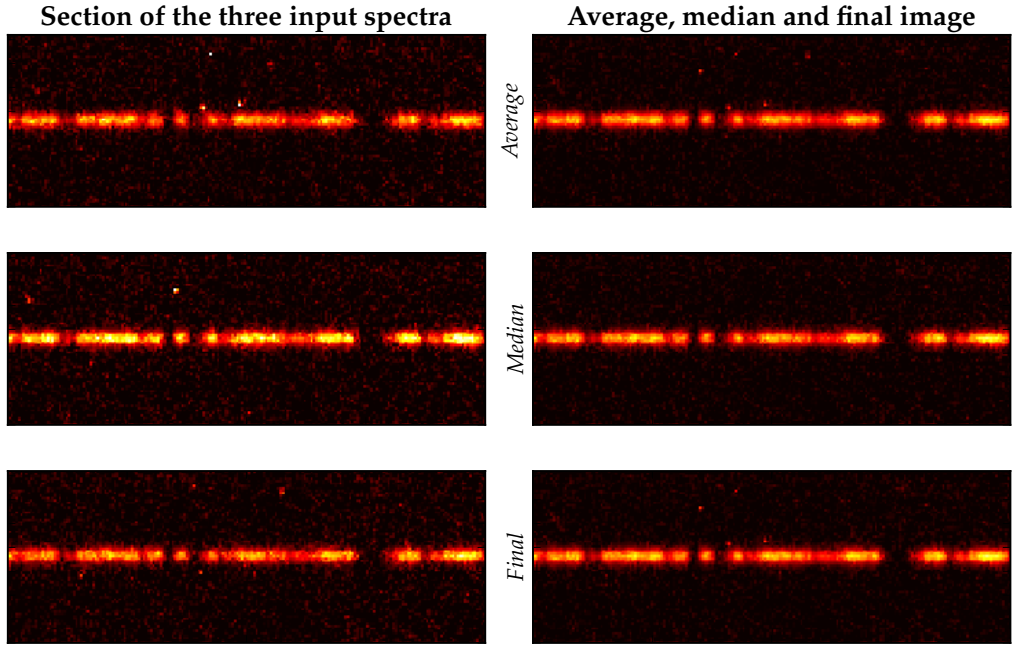


Figure 3.6: Steps in the stacking process. Left, a section of the three 2D spectra of the X-Shooter pipeline output. Right, the same region stacked with different approaches. Units on all axes are pixels, the shown section is from the UVB arm.

The variance is used due to the simple law of error propagation saying that, assuming the errors are random and independent, errors can be added in quadrature, so $\delta(a + b) = \sqrt{\delta a^2 + \delta b^2}$. The square root is not taken yet, as the variances will be needed during extraction as well. The result is shown in figure 3.6, where on the left side is shown a small section of the UVB spectrum. On the right side is shown the simple average and median images, as well as the result of the sigma clipping. As is seen, the simple average shows the most cosmics and artifacts, but it has a good S/N ratio. The median image is cleaner and smoother, but it has a significantly poorer S/N ratio. Finally, as a compromise between these two, the

sigma clipped spectrum that mostly has an S/N ratio as the simple average, but where the poorest values are discarded at the cost of a slightly degraded S/N.

EXTRACTION

The 2D spectrum resulting from the stacking is a sum of contributions from the object, which is only found in the central rows, and the sky background which is present in all rows of the image. In order to get as little sky background contamination as possible, the image is weighted by the strength in the QSO signal, so that regions with no object flux are completely masked out, while the image rows are weighted by their average flux, so rows that are weak in QSO emission and hence have a relatively large contribution from the sky background are given less weight, while central rows with a relatively smaller contribution from the sky background are given stronger weight.

In the following, columns are along the spatial axis, and rows are along the wavelength axis of the image. A weighting profile column is created by taking a row-wise flux average over 4000 columns, which is then normalized through division by the integral over the profile. Every column in the 2D spectrum is then multiplied by the profile column. To preserve S/N ratio, the variance - carried over from the stacking - is multiplied by the square of the weighting profile

Next, the rows of the spectrum and variance frames are summed in the spatial direction, leaving a row of bins for signal and for variance. Finally, the square root of the variance is taken to find the bin-wise σ .

The extraction steps are shown in figure 3.7, showing a section of the UVB arm including the two DLA lines and the Ly β emission line of the quasar. The stacked spectrum is shown in the upper panel. In the middle panel is shown the spectrum after being masked and weighted, and in the lower panel a plot of the extracted 1D spectrum with errors.

3.3.4 CONTINUUM MODELING

Since the science in the work present is based entirely on *relative* fluxes, the fraction of QSO light absorbed by the DLA system, there is no need to perform flux calibration on the data. Instead, the data need to be normalized, so that all fluxes are expressed in terms of the (interpolated) QSO continuum flux. This operation programmed in a Python program, the source code of which is included in appendix B.2.

In some places, it is not difficult to estimate the continuum level of the QSO emission, but in certain areas it is quite a bit harder. That is, for instance, the case around the DLA lines, as is seen from figure 3.7. What the continuum levels would have been with no intervening DLA systems is not obvious; and here,

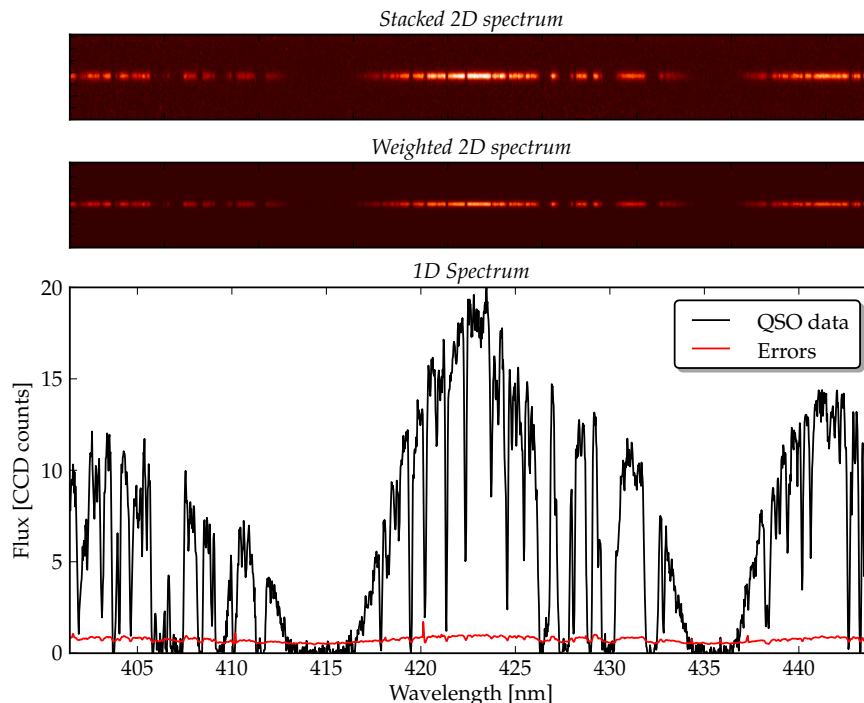


Figure 3.7: Extraction of 1D spectrum. The upper panel is the “raw” image from the stacking process, the second is the weighted image, and the bottom panel shows the extracted 1D spectrum with errors.

maybe more than anywhere else, a good continuum estimate is crucial. If only there were a way to figure out, what the original quasar, sans intervening clouds, *should* look like...

COMPOSITE QUASAR DATA

Quasar spectra, with their high temperature continuum emission and their strong, broad emission lines, are fairly “standardized”; and in the work [33], Telfer et al. have utilized this to create a composite “standard QSO” spectrum out of 332 exposures of 184 QSOs from the Hubble Space Telescope, which has already been shown in figure 1.1. The probability for any two spectra to contain absorption features from intervening clouds at the same redshifts is very small and will be strongly outweighed by the emission of the other spectra. It is, however, clear from the plot that the noise gets stronger towards the blue end, as an increasing fraction of the averaged spectra will show a Lyman α Forest.

As mentioned in section 1.3, quasar spectra are composed of a continuum

contribution and a number of strong emission lines. One of the findings of [33] is that while the strength and width of the strong lines can vary quite a bit, the continuum mostly lies fairly close to a power law $f_\nu \propto \nu^\alpha$ and, as we shall see, the composite spectrum also matches the one of Q 0918+1636 well after extinction correction.

Since the composition technique of [33] will tend to cancel out any peculiarities of the individual spectra, the final result is a “standard quasar” as its spectrum would look right at emission. To have it look comparable to Q 0918+1636, it is necessary to take a few steps to model the transformation of the quasar spectrum as it travels through the Cosmos.

First off, the $z = 2.58$ DLA system is rich on dust and metals, and to correct for the extinction caused by this, it was found that the parametrization of Pei et al. [28] for a dwarf galaxy with a composition resembling that of the Small Magellanic Cloud, best modeled the effect.

Second, it is necessary to correct for dust- and gas-extinction withing the Milky Way Galaxy. For this, the polynomial fits of Cardelli et al. [6] with updated coefficients for the Near-UV by O’Donnell et al. [27] as found in the IDL program CCM_UNRED.PRO was used in a Python implementation written as part of the project.

Finally, modeling of the effects of atmospheric extinction and instrument response was done and applied “backwards” to mimic the effects on the composite data rather than perform a flux calibration on the spectrum of Q 0918+1636.

The modeling of these effects is shown in details in appendix A.

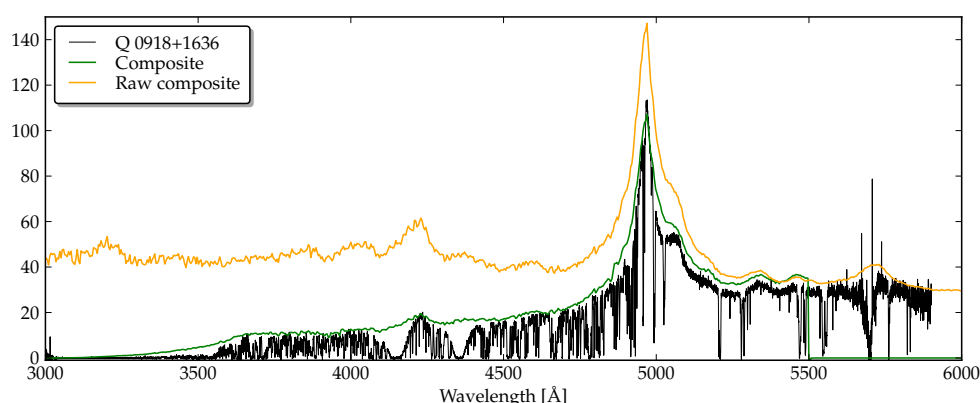


Figure 3.8: UVB arm of the Q 0918+1636 shown along with the composite spectrum of [33] in its raw form and after modelling of SMC-like extinction, Galactic reddening and instrument response.

Chapter 3. Data Reduction

The result of these steps is shown in figure 3.8, in which is plotted the spectrum of Q 0918+1636, along with the raw composite spectrum, and the composite spectrum after imitated SMC-like, Galactic and atmospheric extinction as well as instrument response. As can be seen, the composite data and the combination of models mimics the actual QSO data well, and it is now going to be an aid for the next step: estimating the continuum of Q 0918+1636.

3.3.5 NORMALIZATION

The composite spectrum makes for a pretty good model of what is a likely QSO continuum. This is now drawn in an interactive graphics window, as shown in figure 3.9.

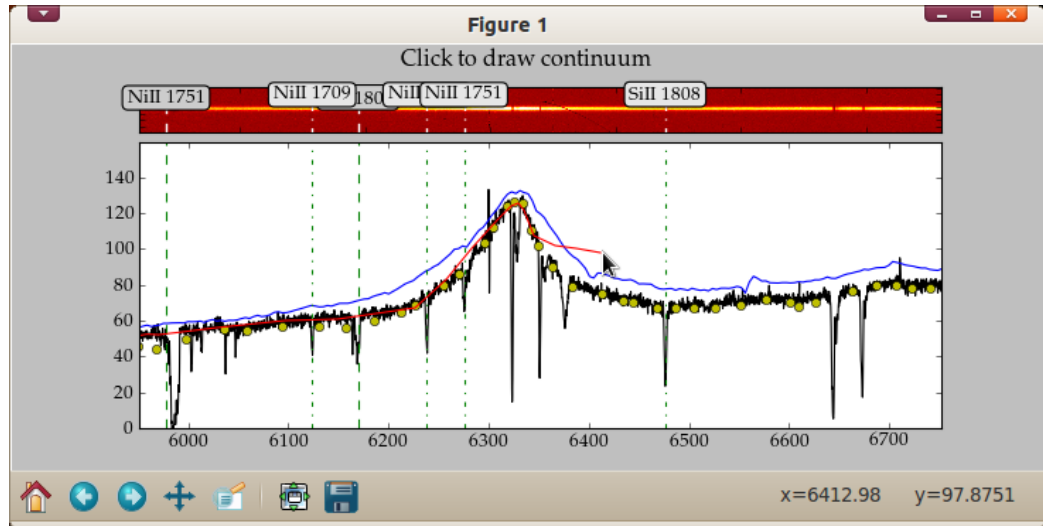


Figure 3.9: Screenshot of the `normalizeVIS.py` python script in action, here working on the VIS arm. Darker labels are wavelengths of transitions for the low-redshift system, lighter labels are for the high-redshift system.

By clicking on a number of points, the estimated continuum is now drawn in the window. The program automatically sorts the coordinates by x-value, allowing the user to add extra points in between already drawn points. A right-click removes the last set of coordinates on the list, and a scrolling upward and downward on the scroll wheel moves to the next and former section of the spectrum, respectively.

As an aid to tell whether a feature is a part of the science object or due to background subtraction, pipeline artifacts etc., the two-dimensional spectrum image is shown above the drawing window, along with markers showing the wavelengths of some important metal absorption lines. The Python `matplotlib`

libraries allow for panning and zooming the data, but due to a limitation in the `imshow()` routine that plots the 2D spectrum, this will only follow the wavelengths of the drawing window if changed by scroll wheel.

When done drawing, the user can close the drawing window with a middle click or by simply closing it in the window manager, and a continuum function is then created by interpolating the selected points with a cubic spline function. The QSO spectrum is then divided by the interpolated continuum function, as is the error spectrum.

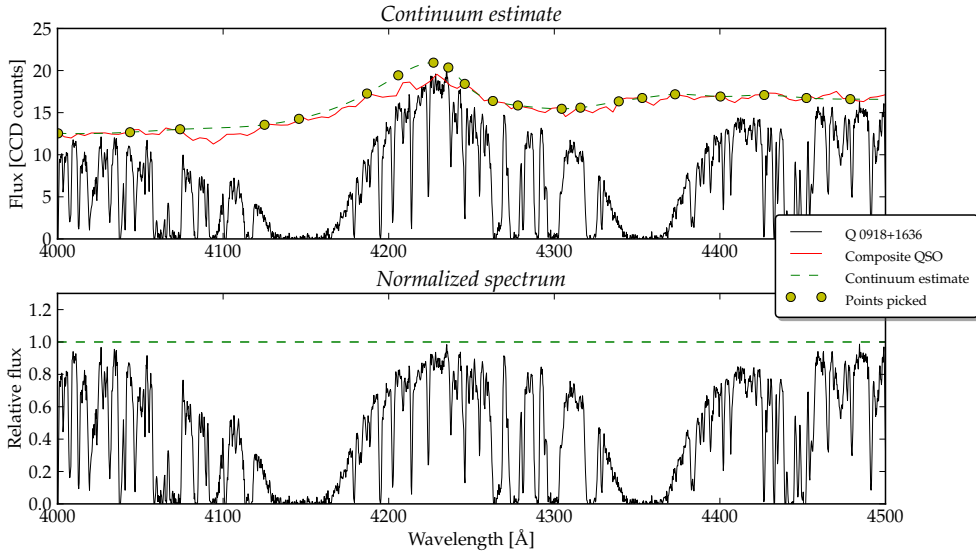


Figure 3.10: Normalization. Upper panel: QSO data shown along with the composite QSO data, the selected points and a cubic spline interpolation. Lower panel: The resulting normalized spectrum, the continuum now having the value 1 everywhere.

Estimating the right continuum is fairly easy in e.g. the region shown in figure 3.9, but less so in other regions, for instance in the $\text{Ly}\alpha$ forest and around the DLA lines, as can be seen in figure 3.10. Especially the fact that the DLA lines are blended with the quasar $\text{Ly}\beta$ emission line meant that estimating the continuum in this area was non-trivial, and an insufficient estimate means it will not be possible to fit the DLA line to a Voigt profile as described in the next section. A couple of iterations were necessary, in which a continuum was estimated, and a Voigt profile fit then attempted. To aid the correction and adjustment of the continuum estimate, the file with the last chosen continuum point coordinates are shown in the drawing window as yellow, filled circles, which can then serve as a scaffolding for a new estimate, as it allows the user to pick the same points in areas that turn out to be fine, and help fine tuning in areas in need of this.

In figure 3.10 is shown a segment of the Q 0918+1636 spectrum before normalization, along with the composite QSO data, the points that were picked in the draw window and the interpolated continuum estimate. Around 4200 Å, the estimated continuum is seen to be a bit above both the Q 0918+1636 and the composite spectrum. This is the estimated Lyman β emission of the QSO without blending from the DLA systems. The Lyman α emission line of science spectrum is, as evident from figure 3.8, stronger and narrower than in the composite spectrum, which fits well with the estimate in figure 3.10. In the lower panel is shown the resulting normalized spectrum of Q 0918+1636, the continuum now being rescaled to the value 1 everywhere. This segment clearly illustrates the utility of the composite quasar spectrum and the difficulty of precise continuum estimation without this modeled continuum - and even with it.

3.3.6 TELLURIC ABSORPTION

Like sky emission, the atmosphere also leaves absorption lines in the spectrum. Even though none of the data used in the end were affected by this, the process is described here for completeness. In order to correct for telluric absorption lines, a spectrum is taken of a so called *telluric standard star*, a hot star with few spectral absorption/emission features of its own and a high S/N ratio. This spectrum is reduced with the X-Shooter pipeline along with the science frames, and the spectrum is then extracted in the same fashion as the science spectrum.

The spectrum is then normalized like the science data, except the relative intensity of the telluric standard spectrum is set to 1 everywhere but in the areas of telluric absorption, as shown in figure 3.11. If taken under the same observing conditions as the science data, it should now be possible to correct for these lines by simply dividing the reduced and normalized science spectrum by the reduced and normalized telluric standard spectrum. However, the telluric standard frame for these observations was several days old, resulting in different weights of different line compounds. This could be manually corrected for by means of splitting the telluric spectrum up into regions which were then multiplied by a corrective factor before dividing the science data.

The UVB arm is unaffected by telluric absorption. Only a few regions in the VIS are affected, while telluric absorption is a strong effect in the NIR arm. However, the NIR data have not been fitted and normalized in this work, so only the VIS arm was corrected for telluric data; and it was later found that only one line was affected, which was after the correction found to blend with other, unidentified absorption features and hence not suitable for inclusion in the fits.

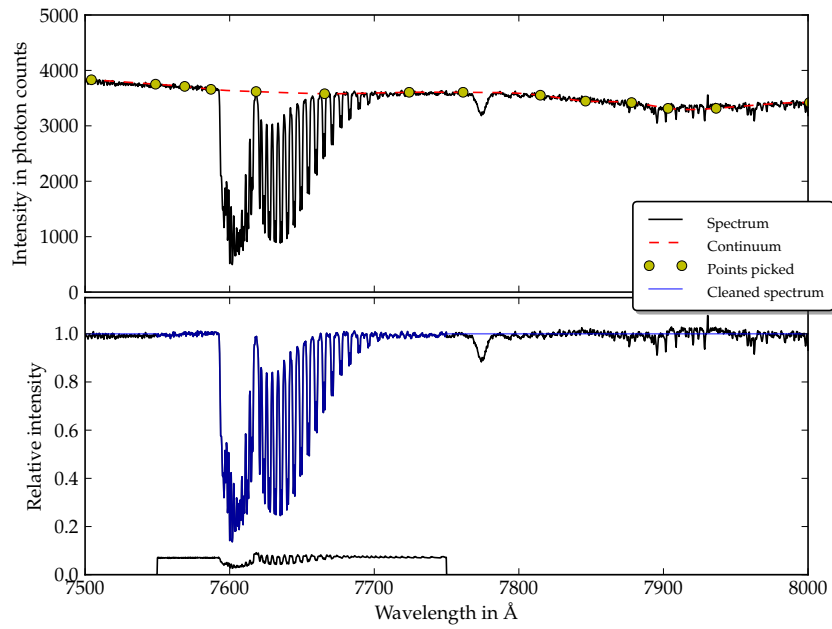


Figure 3.11: Normalization and cleaning of telluric standard spectrum. The relative intensity is set to 1 everywhere but in areas of known telluric absorption.

CHAPTER 4

ANALYSIS AND RESULTS

4.1 DLA LINE FITTING

Since the DLA absorption features are riddled with Lyman α forest absorption lines, they cannot easily be modeled by a rigorous χ^2 -fit. Instead, we must fit them by eyeballing - that is, drawing a Voigt profile along with the data from some reasonably guessed values of the parameters N and z - the DLA lines are strongly saturated and thus the equivalent width and the shape does not depend significantly on b , as explained in section 2.1.3. The process is iterative in two senses; the values of N and z are continually adjusted to give an optimal fit, but at the same time, this fitting also serves as a quality control of the continuum estimate, since it is not possible to make a satisfactory fit if the continuum estimation is not done correctly.

Error estimation is likewise done by hand, by choosing a value of $\delta \log N$, plotting the Voigt Profile with a column density value corresponding to $\log N \pm \delta \log N$, and varying $\delta \log N$ to the lowest value at which the column density is clearly too high or too low. This is obviously a somewhat subjective method and could be seen as unsatisfactory; but the value of the column density is to be used to find a metallicity of the system, the uncertainty of which is clearly dominated by those of the Voigt profile fits of the metal lines, which *are* carried out as proper χ^2 fits, as explained below. It is probably possible to find a more rigorous way of estimating the errors for the HI column density, but it will most likely not be worth the effort.

For drawing the Voigt profile is done an implementation of the approximation by Thorsten Tepper García [34], an approximation that has an accuracy of 10^{-4} for the entire range of column densities and damping parameters characteristic for all Lyman transitions in HI absorption systems, including Damped Lyman α Absorber systems along with Lyman Limit systems and Ly α forest clouds.

Q 0928+1636		
System	Redshift z	Col. dens. $\log(N[\text{cm}^{-2}])$
DLA 1	2.4124	21.26 ± 0.06
DLA 2	2.5832	20.96 ± 0.05

Table 4.1: Best fit values of redshift and HI column density for the two DLA systems in the spectrum of Q 0918+1636. The values for DLA 2 are from [11].

The best fit reached is shown in figure 4.1, drawn as the bold red line. The dotted red lines are the line profiles of the column densities corresponding to $\log N \pm \delta \log N$. The blue line is a tentative fit of the other DLA system, treated in [11], included to indicate their relative location.

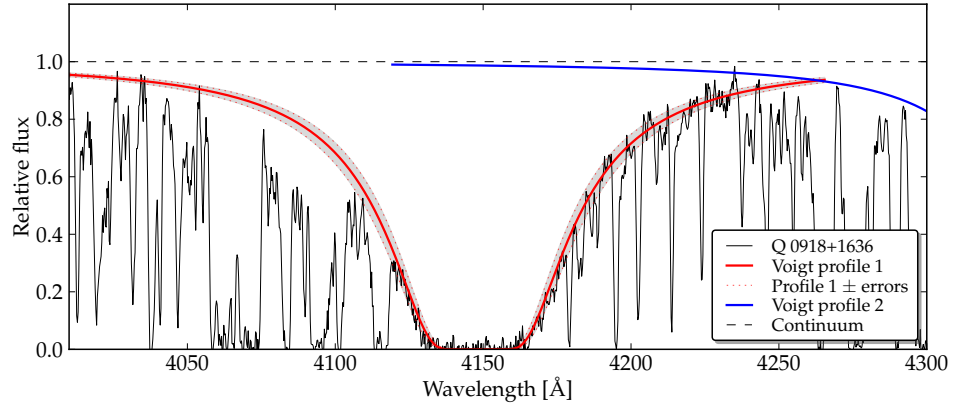


Figure 4.1: The $z = 2.412$ DLA absorption feature fitted to a Voigt profile, shown as the bold red curve. Best fit values are $z = 2.4124$ and $\log N = 21.2624$. In this strongly saturated regime, N does not depend on b . The blue line is a tentative fit of the DLA treated in [11].

The best fit values of redshift and HI column density of the two DLA systems in the spectrum of Q 0928+1636 are given in table 4.1, with the values for the highest redshift system taken from [11].

4.2 ROBUST LOWER LIMITS

Lower limits for the metal column densities and an inferred metallicity are found as described in section 2.1 by finding the equivalent width w as the width of a rectangle of height $h = 1$ and the area of the integral $\int_{\lambda_1}^{\lambda_2} 1 - f(\lambda) d\lambda$, with λ_1 and λ_2 being the start and end wavelengths of the line feature. The column

density is then found by assuming the line to be within the optically thin regime and using equation 2.13.

The metallicity is found as $\log(\frac{N_{ion}}{N_{HI}})$, and then expressed in terms of the Solar metallicity for the given ion, the values of which are found in [2], in which an extensive table is given for the abundances of a number of metals. The areas are shown in figure 4.2, and the inferred values of the equivalent width and metallicity are shown in table 4.2.

The comparison to solar metallicity levels rests on the assumption that the majority of a given metal is in the low-ionized phase; an assumption that might not be entirely correct. However, both the assumption of being in the optically thin limit and of the ions to be constituting the entire abundance of the metal will both lead to underestimating the metallicity, stressing that the limits are indeed both robust and lower. Likewise, no error estimates are given for these limits, since the uncertainties are likely dominated by the validity of these assumptions.

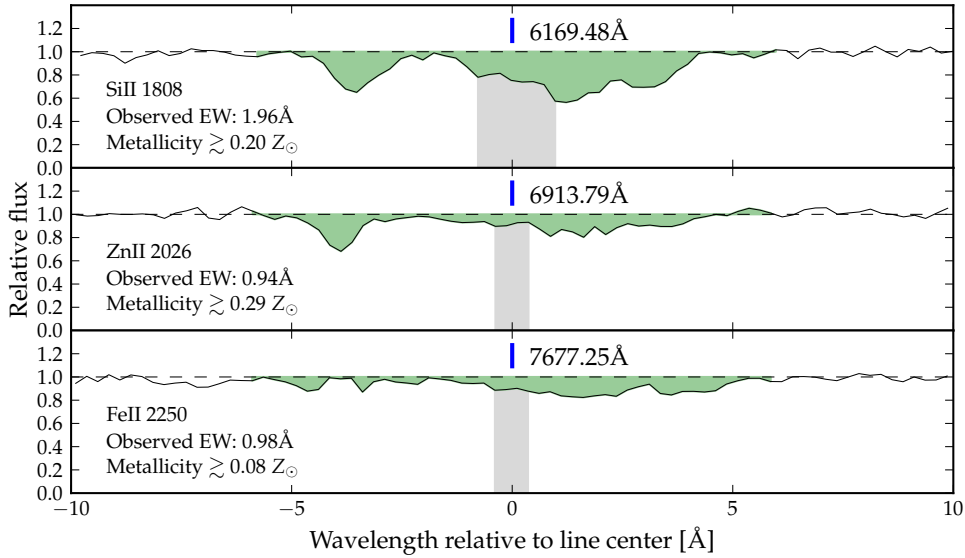


Figure 4.2: Robust lower limits of the column density and inferred metallicity for the lines SiII 1808, ZnII 2026 and FeII 2050. Note how the line of FeII, which sticks to dust grains easily, shows significantly lower metallicity than SiII and ZnII which are much less prone to adhere to dust.

ROBUST LOWER LIMITS					
Ion (transition)	$w_{obs}[\text{\AA}]$	$w_{em}[\text{\AA}]$	$\log N_{ion}$	$\log N_{HI}$	$Z_{lim} [Z_{\odot}]$
SiII (1808 \AA)	1.96	0.57	15.98	21.26	0.20
ZnII (2026 \AA)	0.94	0.28	13.18		0.29
FeII (2250 \AA)	0.98	0.29	15.46		0.08

Table 4.2: Robust lower limits for equivalent widths and inferred metallicity for the ion transitions SiII 1808, ZnII 2026 and FeII 2250. Values of N are given in units of cm^{-2}

4.3 LINE PROFILE FITTING

A number of selected lines are fitted to a line profile as described in section 2.1; with tied b parameters reflecting the assumption that ions of low ionization (0-1 missing electron) all belong to a common phase, and ions of higher ionization belonging to a second, highly-ionized phase. Originally, purely turbulent motion was assumed, yielding a simple tying of the Doppler parameters to a single value for each component, but also a fit of thermal broadening was tested, which turned out to yield considerably better χ^2 values as described below.

The lines to be fitted were selected for being good tracers of metallicity (especially SiII and MnII), for being optically thin to minimize the risk of hidden saturation while still having a well-defined component structure. The structure of the absorption line feature is quite complex, at a first glance consisting of at least 3-4 components. These components are interpreted as physical subsystems with each a peculiar velocity within the main system, adding to the cosmological redshift (and possibly Doppler redshift due to comoving velocity) a Doppler redshift which shifts them slightly from the redshift of the entire system as measured from the DLA line.

The main part of this work has assumed a five or six component system; the final choice of the number of lines is discussed below.

4.3.1 VPFIT

The fits were done using the software package `vpfit` [7], a powerful software package specialized in fitting Voigt profiles in astronomical absorption line systems. It is free software and released under the *GNU General Public License v3*. The software package contains a number of fitting functions written in FORTRAN, and uses the PLPLOT library for data visualization. It has a large number of features and optional settings, of which only the ones used for this work shall be treated here, a full documentation can be found in the manual [7]. As input, `vpfit` takes an initial guess of the line profile characteristics; the most impor-

tant ones being the data set and wavelength intervals to include in the fitting, first-guess values of z , b and N for each of the line components and, optionally, which parameters are to be considered fixed at a certain value or tied together with other parameters. It features a terminal-based interface which can either interactively take graphic input from the plot window, or it can be fed at setup text file which can be manually edited or, as seen below, edited by third-party software.

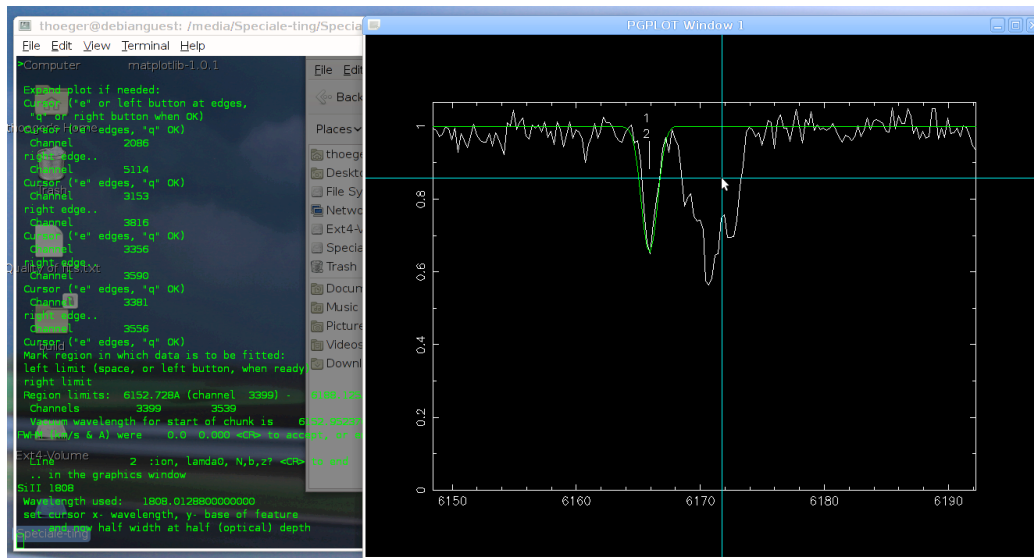


Figure 4.3: The interactive `vpfit` interface in action; prompting the user to click on the “half width at half (optical) depth”. A challenge to the imagination and spatial intelligence.

`vpfit` allows the option to fix any of the line parameters to a certain value, as well as tying them together for different lines and line components. When tying the Doppler parameter, an option can be set to control whether they should be fitted assuming thermal or turbulent motion, or a combination thereof[7].

4.3.2 VPGUESS

While `vpfit` provides excellent fitting capability and a wide range of features, the built-in user interface for providing the initial guesses is rather crude and can be frustrating to work with. Providing the guesses consists of a long sentence of mouse clicks and keyboard presses, of which none can be undone, so that just one wrong key press or mouse click can render the preceding series of input actions invalid. Furthermore, the graphical user interface of `vpfit` provides little by the means of aid to visualize the guesses made. It is not possible to see, how two line guesses blend together in order to check how well they fit the actual data,

and it is only possible to view the different lines one by one. This poses quite a challenge if a line profile contains more than 2-3 blending components, as it is quite difficult, unless one is very experienced, to “eyeball” a deconvolution of the line components. At the same time, with an increasing number of components comes an increasing sensitivity to the quality of the initial guesses.

A very helpful piece of software in this regard is `vpguess` which, as the name suggests, helps providing the initial guesses for line profile fitting. The software can be used as a stand-alone program but is designed to be used as an enhancement and front-end to `vpfit`. A screenshot of `vpguess` in action can be seen in figure 4.4.

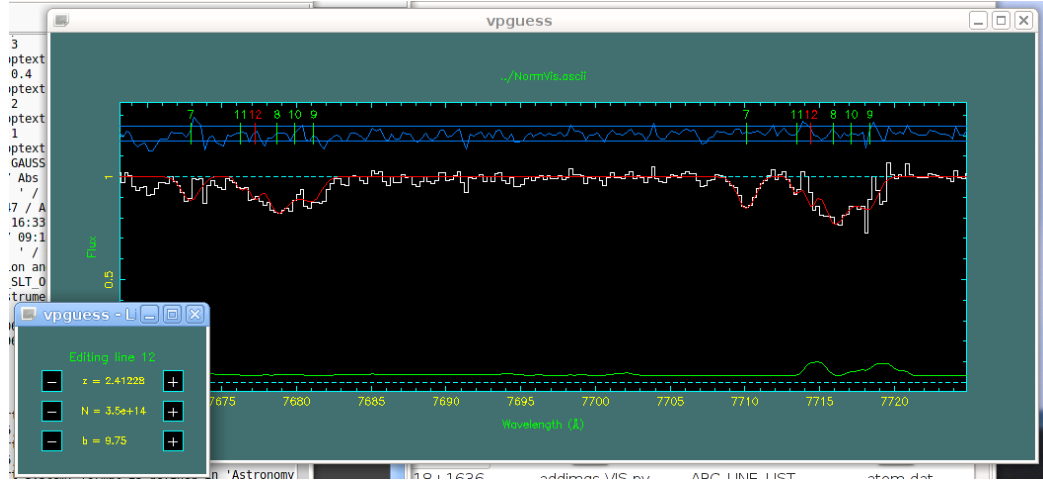


Figure 4.4: The software package `vpguess` in action. Here working on the FeII lines of rest frame wavelengths 2250 Å and 2261 Å.

The `vpguess` software allows the user to add a number of line components, provides an interactive window to adjust z , b and N , and allows the user to fix, tie and untie the values of these parameters between selected lines, in which case the values of said parameter are locked and shown in red in the line editing window for all but the selected reference line. All this can be done while viewing the resulting model on the screen along with the data, as shown on the figure. It also provides an option to show residuals and error spectrum on the plot if so wanted. One very neat feature of `vpguess` is the option to show an arbitrary number of data/model segments on the screen, helping to see if an adjustment to the input model is an improvement in other lines than the one currently worked on. When a satisfactory result has been reached, `vpguess` can either output the guessed model in a `vpfit` readable format, or feed it directly to `vpfit`, which then does the rigorous χ^2 fit of the model. The fitted parameters are then read back into the `vpguess` window, where the model is generated using a Voigt profile

approximation with slightly better accuracy than the one provided by `vpfit`.

For this work, the initial guesses are results of a combination of `vpguess` and manual editing, and the output model is generated by `vpguess` based on the parameters fitted by `vpfit`.

4.3.3 VELOCITY COMPONENTS

Figuring out how many line components are constituting the line profile is, for a large part, a matter of trial-and-error, and some assumptions have had to be made, partly based on physical conjecture, partly based on the system's morphology in velocity space. It seems, therefore, that a short note on the process is in order here.

When looking at any one of the absorption profiles in e.g. figure 4.3, it is clear that it consists of a number of components with various degree of blending, depending on their differences in redshifts and their Doppler parameters. Each different redshifts is interpreted as a peculiar velocity of a physical subsystem within the enveloping HI cloud. For low- or high-ionized elements, respectively, we are assuming that they all belong to one, common phase; a fundamental assumption for the curve-of-growth-method that is the very basis for determining metallicities.

While there is no reason *per se* that the high- and low-ionized phases should be found in the same subsystems, the morphology of the lines does suggest that they do, which was assumed throughout the guessing and fitting process. Although the phases were assumed to constitute the same subsystems, they were still allowed to have different redshifts. Based on knowledge of interstellar medium in the Galaxy, there could be several reasons that phases of high and low ionization, though parts of the same system, would have differing peculiar velocities and hence different redshifts; e.g. in regions of star formation where a highly ionized medium could be part of a shock front propagating outwards from the center of star formation. Hence, the two phases were given similar but not identical values of redshift in the initial guess model, and their redshifts were tied within the phases but not between them; as was also the case with the Doppler parameters.

Selecting the number of components for `vpfit` input is a matter of finding the lowest possible number of line components that can constitute a satisfying model for all line components. After having tried with four and five, components, I eventually settled on six, as that sixth line gave a significant improvement in both visual resemblance and fit χ^2 .

Column densities were kept as a free parameter everywhere, as the abundance of any ion in one subsystem is assumed to have no influence of the abundance of said ion in any other subsystem.

4.3.4 STATES OF IONIZATION

The metallicities found below are, unless otherwise stated, based on the assumption that practically the entire abundance of a metal has the given ionization. For hydrogen, however, it is an important assumption that practically all is neutral (see e.g. Howk et al. [17] for a further discussion of ionization in DLA systems).

This is based on the difference in ionization energies of the various elements. If an element has an ionization energy lower than that of H I, it can be easily ionized if any photons at right energies are around, which will be the case in young, star-forming galaxies. If, however, the ionization energy of an ion is similar to or larger than that of H I, the vastly more abundant H I atoms will have a much larger probability of absorbing the photon and becoming ionized. Further ionization of the metal will cease, and it has become **Hydrogen shielded**. The abundance of neutral hydrogen, however, is so large that the number of ionizing photons will not change the ionization ratio significantly, as they could for the much less abundant metals. Ionization energies for the elements can be found in various catalogs. As an example, the ionization energy of Si I is $E_I \approx 8.15$ eV, while the ionization energy of Si II is $E_{II} \approx 16.35$ eV. Since the ionization energy of hydrogen is $E_{HI} \approx 13.6$ eV, it does leave unaffected a large part of the photons capable of ionizing Si I, while any photons capable of ionizing Si II will most likely have been absorbed by a hydrogen atom before meeting a Si II atom; and further ionization does not happen. Similar numbers can be found for the other metals, except Mg I which, as seen below, has a very low abundance, because most of the Mg probably is in a state of higher ionization. However, Mg was not within the scope of this project, and the Mg I lines were included only because the weak but significant Mg I 2026 line blends with the Zn II 2026 line. The significantly stronger Mg I 2853 line was then added to improve the fit.

4.3.5 TURBULENT MOTION

Originally, the physical working assumption of this project has been that all internal motion in the subsystems would be purely turbulent.

The values of z and b given as input to `vpfit` are shown in table 4.3 along with their respective results with uncertainties. Since the parameters are tied, values are given for the reference ion only. For the highly ionized components, `vpfit` has consistently dropped one component, meaning that it was too faint to be fitted with a satisfactory uncertainty. This was also the case when assuming thermal internal motion, so it seems that at least one of the components has a very low degree of ionization.

What seems a bit more disconcerting is that `vpfit` has also dropped the tying of redshifts for one of the lines for the low ionizations, the one marked with a

(^{*}). Physically, this is a quite dubious assumption, as it means that each ion in this specific component should have slightly different peculiar velocities, which seems unphysical. However, the values found do not seem to be in disagreement with what was found when assuming that the internal motion is instead thermal, see section 4.3.6. The values of z for the different ions in the offending line component are shown in table 4.4.

INITIAL GUESS AND FIT RESULTS, TURBULENT MOTION				
Reference Ion	Guess value		Best fit	
	z	b [km/s]	z	b [km/s]
SiII	2.410245	22.84	2.410328 ± 0.000007	38 ± 1
	2.412974	30.92	2.41289 ± 0.00008	37 ± 11
	2.413992	26.03	2.41404 ± 0.00003	28 ± 3
	2.413415	46.76	2.4134 ± 0.0002	45 ± 28
(*)	2.412377	28.81	2.41231 ± 0.00004	22 ± 6
	2.411792	10.72	2.41179 ± 0.00003	20 ± 3
C IV	2.410436	51.06	2.41039 ± 0.00002	33 ± 3
	2.414308	15.60	2.4142 ± 0.0001	21 ± 10
	2.412786	19.26		
	2.411939	46.78	2.409 ± 0.002	274 ± 176
	2.413445	37.41	2.4130 ± 0.0001	84 ± 8
	2.411825	5.27	2.4110 ± 0.0003	88 ± 22

Table 4.3: Initial guesses and best fit values for z and b under assumption of purely turbulent motion. `vpfit` has dropped the tying of redshift for one of the components, the values found are shown in table 4.4. For high ionizations, `vpfit` has dropped a line, which is done if it is too weak and/or has too low Doppler parameter, or no satisfactory fit can be made for this line component.

VALUES OF z FOR THE LINE ([*]) FOR LOW IONIZATION ELEMENTS:						
Ion	SiII	MnII	ZnII	FeII	CrII	MgI
z	2.412377	2.412691	2.412344	2.412370	2.412211	2.412403

Table 4.4: The values of redshift found by `vpfit` under assumption of turbulent internal motion for the line (^{*}) in table 4.3.

In table 4.5 is shown in the same fashion the initial guess values of N along with the `vpfit` results, this time for all ions and all redshifts, since none of these

INITIAL GUESSES AND FIT RESULTS FOR $\log N$, TURBULENT MOTION					
Ion	$\log N$, input	$\log N$, result	Ion	$\log N$, input	$\log N$, result
SiII	15.352	15.40 ± 0.05	CrII	13.0228	12.8 ± 0.1
	15.4879	15.4 ± 0.5		13.2545	13.2 ± 0.3
	15.2918	15.2 ± 0.2		12.9783	13.0 ± 0.1
	15.0876	15.3 ± 0.7		13.0485	12.9 ± 0.8
(*)	15.1624	14.9 ± 0.2	(*)	12.6381	12.6 ± 0.2
	14.8564	15.0 ± 0.1		11.9628	12.3 ± 0.4
MnII	12.765	12.79 ± 0.05	MgI	12.5359	12.60 ± 0.05
	12.8642	12.6 ± 0.6		12.1545	12.4 ± 0.6
	12.5153	12.5 ± 0.2		11.8787	11.5 ± 1.0
	12.244	12.6 ± 0.7		12.3155	12.4 ± 0.6
(*)	12.1161	12.5 ± 0.2	(*)	12.3643	12.1 ± 0.1
	11.6618	11.9 ± 0.2		12.1519	12.34 ± 0.07
ZnII	12.7247	12.77 ± 0.05	C IV	15.0244	15.53 ± 0.08
	12.5038	12.5 ± 0.5		16.0244	12.0 ± 1.1
	12.3264	11.9 ± 0.4		17.0244	
	12.329	12.4 ± 0.7		18.0244	13.7 ± 0.3
(*)	12.2567	12.0 ± 0.3	(*)	19.0244	13.4 ± 0.2
	11.9628	12.3 ± 0.1		20.0244	13.9 ± 0.3
FeII	14.6921	14.70 ± 0.05	AlIII	21.0244	13.75 ± 0.05
	14.8819	14.9 ± 0.5		22.0244	12.3 ± 0.2
	14.6432	14.7 ± 0.2		23.0244	
	14.8348	14.8 ± 0.7		24.0244	12.5 ± 0.4
(*)	14.7374	14.6 ± 0.1	(*)	25.0244	13.22 ± 0.08
	14.1305	14.29 ± 0.08		26.0244	12.8 ± 0.3

Table 4.5: Initial guess and `vpfit` results for column densities assuming purely turbulent motion. The values of N are given in units of cm^{-2}

METALLICITIES, TURBULENT MOTION				
Ion	$\log N$	Z_{DLA}	Z_{\odot}	$Z - Z_{\odot}$
SiII	16.0 ± 0.2	-5.1 ± 0.2	-4.5	-0.6
MnII	13.3 ± 0.2	-7.8 ± 0.2	-6.6	-1.3
ZnII	13.2 ± 0.1	-8.0 ± 0.2	-7.4	-0.6
FeII	15.5 ± 0.2	-5.7 ± 0.2	-4.5	-1.2
CrII	13.7 ± 0.2	-7.5 ± 0.2	-6.4	-1.2
MgI	13.1 ± 0.2	-8.1 ± 0.2	-4.4	-3.7
C IV	15.55 ± 0.08	-5.6 ± 0.1	-3.6	-2.0
AlIII	13.92 ± 0.04	-7.2 ± 0.07	-5.6	-1.69

Table 4.6: Metallicities, purely turbulent motion. The values of N are given in units of cm^{-2} .

values are tied. Again, the component with the untied redshifts is marked with (*).

The modeled spectrum with the fitted parameter values is shown in figure 4.5. The figure is showing the 11 data regions with the modeled spectrum in red superimposed on the data in black. The line center of each component is shown as a vertical tick above the graph, and between line centers and the spectrum+model is shown the residuals in terms of uncertainties, shown on top of a black dashed line showing the zero level of residuals and full black lines for $\pm 1\sigma$. The error spectrum, which has been multiplied by 2 in the figure to visually enhance its shape, is shown as shaded gray at the spectrum axes.

METALLICITY

By adding the column densities of an ion for each of the Doppler components, an overall column density is found for this ion which can then be compared to the total column density of HI in the system to determine its metallicity. This is done in the script `Metallicities.py`, which is included in the appendices. This script reads the output of `vpfit`, sees how many components are present, adds them up for each ion and intelligently finds the solar abundance for the relevant ions and subtract their solar metallicity from the measured one.

The results of these calculations, assuming purely turbulent motion, are shown in table 4.6.

The metallicities found are compared to Solar levels as found by Asplund et al. [2] and written in terms of these. Uncertainties of Z_{\odot} are ignored, as these values merely serve as a frame of reference and their uncertainties provide no information about the qualities of the present work. The highest metallicities

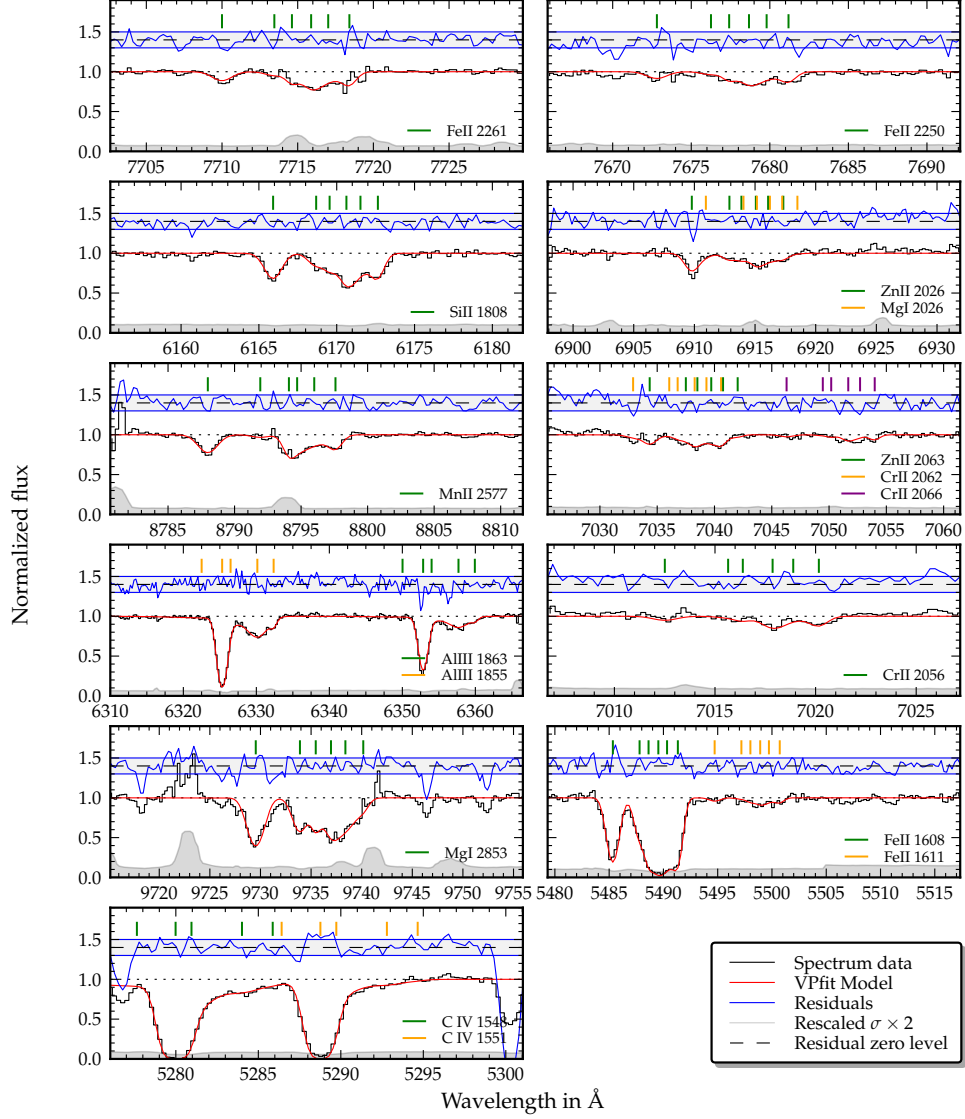


Figure 4.5: Fitted model of Q 0918+1636 under the assumption of turbulent internal motion. Data are shown in black, and the model shown in red. Uncertainties, doubled for illustrative purposes, are shown as shaded gray. The blue line above the model shows the residuals in terms of the errors, with the gray area denoting $\pm 1\sigma$. Due to rebinning in the X-Shooter pipeline, the uncertainties are likely overestimated by a factor of $\sim \sqrt{2}$.

INITIAL GUESSES AND BEST FITS, MIXED MOTION				
Reference ion	Guess value		Best fit	
	z	b [km/s]	z	b [km/s]
SiII	2.410361	25.38	2.410328 ± 0.000007	38 ± 1
	2.412991	33.83	2.41289 ± 0.00008	37 ± 11
	2.413992	26.03	2.41404 ± 0.00003	28 ± 3
	2.413415	35.97	2.4134 ± 0.0002	45 ± 28
	2.412277	36.25	2.41231 ± 0.00004	22 ± 6
	2.411825	12.54	2.41179 ± 0.00003	20 ± 3
C IV	2.410436	51.06	2.410380 ± 0.000009	47 ± 1
	2.414308	15.60	2.4141 ± 0.0001	52 ± 14
	2.412786	19.26	2.4124 ± 0.0001	69 ± 21
	2.411939	46.78	2.4110 ± 0.0002	71 ± 20
	2.413445	37.41	2.41324 ± 0.00007	55 ± 19
	2.411825	5.27		

Table 4.7: Initial guesses and best fit values for redshift and Doppler parameter when assuming thermal internal motion.

are found for SiII and ZnII, which are largely indifferent to the presence of dust. Other ions, like e.g. FeII and MnII, are more prone to adhere to dust grains and will thus be depleted in the presence of dust [25].

4.3.6 THERMAL MOTION

In table 4.7 is shown, similar to table 4.3, the initial guess values of and best fit results for b and z , this time when assuming internal motion is thermal. The corresponding column densities are found in table 4.8.

In figure 4.6 is shown the modeled spectrum when allowing for a thermal part of the Doppler broadening. Though not visibly much of an improvement, this meant significantly lower errors in the fit.

When comparing the fits for the two possibilities, it is seen that there is a few line components that are quite persistent regardless of assumption of turbulent or thermal motion, while others vary quite a bit more. This could suggest that there are in fact even more subsystems present than the six already fitted. For instance, `vpfit` quite consistently fits the deep component with the lowest redshift as consisting of two subcomponents for the high ionizations, while sticking with the input guess of one component in the low ionization lines. However, when looking at this component in the low ionization case, the is quite consistently either a bit too shallow and wide, a bit too deep and narrow, or simply too

INITIAL GUESSES AND BEST FIT RESULTS FOR N , WITH THERMAL MOTION

Ion	log N , input	log N , result	Ion	log N input	log N , result
SiII	15.3978	15.43 ± 0.03	CrII	13.0228	12.79 ± 0.10
	15.5739	15.5 ± 0.3		13.2545	13.2 ± 0.2
	15.2918	15.3 ± 0.1		12.9783	13.03 ± 0.08
	15.0084	15.3 ± 0.5		13.0485	12.9 ± 0.4
	15.1134	14.9 ± 0.2		12.6381	12.6 ± 0.2
	14.7757	14.90 ± 0.09		11.9628	12.3 ± 0.2
MnII	12.765	12.78 ± 0.03	MgI	12.5853	12.63 ± 0.02
	12.8642	12.8 ± 0.2		12.4422	12.4 ± 0.4
	12.5153	12.56 ± 0.08		11.8831	11.5 ± 1.1
	12.244	12.5 ± 0.4		12.3021	12.5 ± 0.4
	12.1161	11.8 ± 0.4		12.4685	12.1 ± 0.2
	11.6618	11.6 ± 0.3		11.9628	12.23 ± 0.06
ZnII	12.6593	12.74 ± 0.03	C IV	15.0244	14.97 ± 0.02
	12.5038	12.4 ± 0.2		12.7881	12.9 ± 0.3
	12.3264	12.0 ± 0.2		13.0412	13.4 ± 0.3
	12.329	12.4 ± 0.3		13.4598	13.8 ± 0.2
	12.2567	12.1 ± 0.1		13.2601	13.2 ± 0.4
	11.9628	12.13 ± 0.10		11.9628	
FeII	14.6921	14.74 ± 0.02	AlIII	13.8688	13.75 ± 0.02
	14.9229	14.9 ± 0.2		12.626	12.6 ± 0.1
	14.7559	14.78 ± 0.08		12.8543	12.9 ± 0.1
	14.7552	14.8 ± 0.3		12.8882	12.8 ± 0.2
	14.6071	14.61 ± 0.10		12.7666	12.9 ± 0.2
	14.3159	14.21 ± 0.07		11.9628	

Table 4.8: Initial guesses and `vpfit` results when assuming thermal internal motion. Values of N are given in units of cm^{-2}

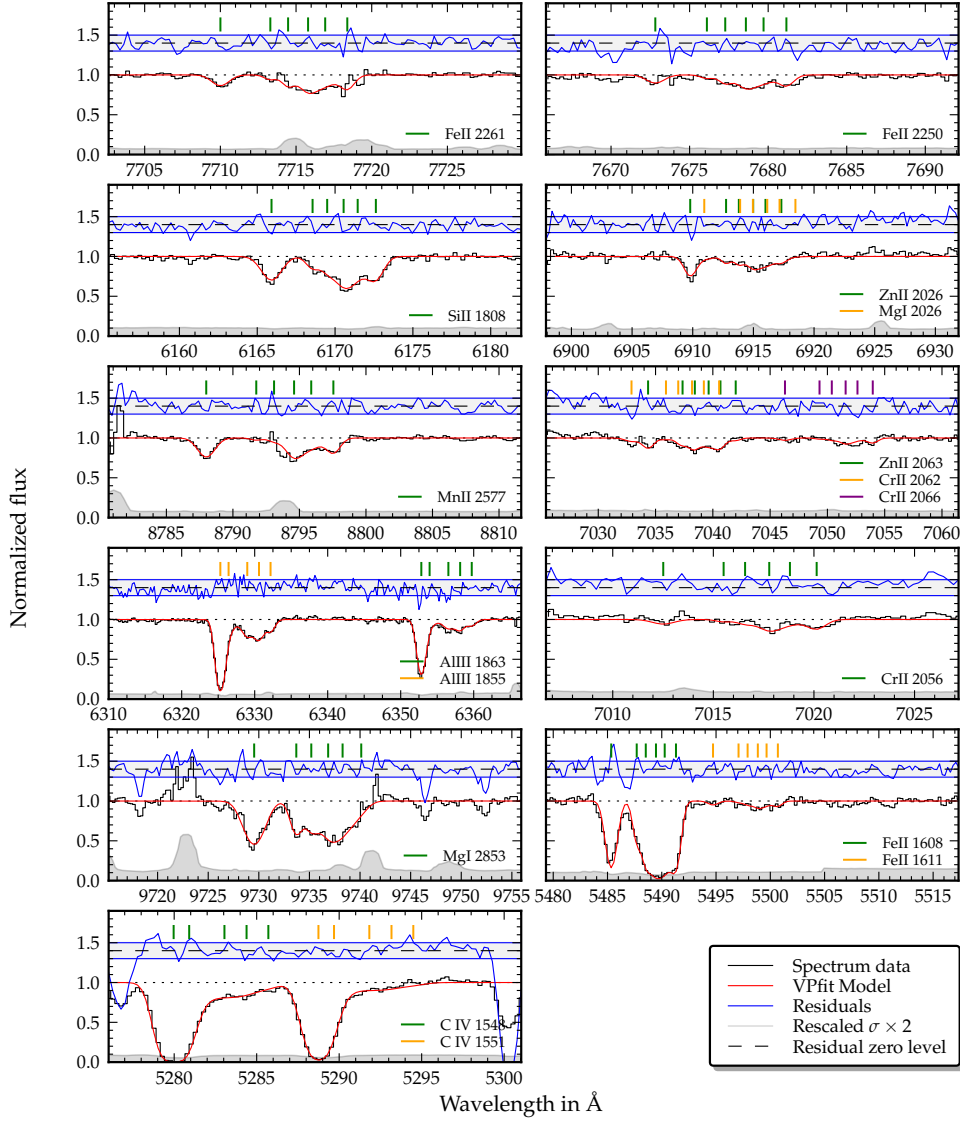


Figure 4.6: Fitted model of the absorption features also seen in figure 4.5, this time assuming thermal internal motion. See figure 4.5 for explanation.

METALLICITIES, TURBULENT MOTION				
Ion	$\log N$	Z_{DLA}	Z_{\odot}	$Z - Z_{\odot}$
SiII	16.1 ± 0.1	-5.1 ± 0.1	-4.49	-0.62
MnII	13.30 ± 0.09	-7.9 ± 0.1	-6.57	-1.29
ZnII	13.15 ± 0.07	-8.01 ± 0.09	-7.44	-0.57
FeII	15.50 ± 0.09	-5.7 ± 0.1	-4.5	-1.16
CrII	13.7 ± 0.1	-7.5 ± 0.1	-6.36	-1.15
MgI	13.1 ± 0.1	-8.1 ± 0.1	-4.4	-3.65
C IV	15.02 ± 0.03	-6.15 ± 0.07	-3.57	-2.58
AlIII	13.91 ± 0.03	-7.25 ± 0.07	-5.55	-1.70

Table 4.9: *Metallicities as found when fitting for thermal motion. Values of N are given in units of cm^{-2}*

symmetric, suggesting that this component in fact does consist of two subcomponents as is fitted for high ionizations. This component is very strong in high ionizations, which could explain why the discrepancies introduced by the oversimplification of only having one sub component would be so large it affected the fit result significantly for the high ionization case, urging the program to correct for this, while it would be smaller in the low ionization case and less important for the overall χ^2 .

Likewise, the line components on the redward side if the absorption feature are significantly stronger in the low ionization phase than in the high ionization phase. This could explain why `vpfit` has dropped lines in this segment in the high ionization case, while it had a larger effect on the total χ^2 of the fit to get these lines right in the low ionization case. Hence, modeling the input with a seventh component in the red end of the absorption feature could possibly increase the fit results.

4.3.7 METALLICITIES

While the distribution on subcomponents varies in the fits above, it is reassuring to see that the metallicities found are in good agreement; meaning that the discrepancies between the fits are of little importance to the gross results of this work. Metallicities for thermal motion fits are shown in table 4.9.

COMPARISON WITH ROBUST LOWER LIMITS

The fit results are now compared to the robust lower limits of section 4.2; partly to make sure that they are in accordance, partly to see how well the assumption

FITTED METALLICITIES AND ROBUST LOWER LIMITS			
Line	Robust limit	Metallicity, turb.	Metallicity, therm.
SiII 1808	-0.70	-0.63 ± 0.19	-0.62 ± 0.13
ZnII 2026	-0.54	-0.55 ± 0.16	-0.57 ± 0.09
FeII 2250	-1.10	-1.17 ± 0.20	-1.16 ± 0.11

Table 4.10: Comparison of robust lower limits with metallicity calculations based on the best fit values of N and b for the two scenarios of internal motion. Silicon is seen to have a slightly higher metallicity than its robust lower limit, though still within a 1σ limit. FeII and ZnII are, slightly surprising, seen to have slightly lower best fit metallicities than the robust lower limits; but they are still well within the 1σ limit.

of optically thin conditions seems to be in accordance with the values calculated from best fit values of `vpfit`. The values for robust limits and metallicities are calculated from the best fit values of N and b from the two different scenarios for internal motion. The values are shown in table 4.10.

Overall, the values are seen to be in agreement within a 1σ error, most of them even well within it. The metallicity for SiII seems to be slightly above the lower limit, while the other two lines, perhaps a bit surprisingly, are slightly lower than the lower limit. This is in good agreement with the SiII line having the largest optical thickness and hence the largest likelihood of not living up to the assumption of being in the optically thin regime. But since all values are within one standard deviations, this is just speculation; the robust lower limits and the fits match very well for all the three lines considered. Therefore, it seems unlikely that Any significant hidden saturation has been present.

4.4 DYNAMICS

The dynamic velocity for a transition is calculated by integrating the optical depth of the line, and measuring the distance in velocity space between the wavelengths at which 5% and 95% of the integral has accumulated, respectively. The Doppler velocity found is the dynamic velocity of the system.

In figure 4.7 is shown the dynamic velocity for the two lines FeII 2261 and SiII 1808. In the upper panel of each plot is shown the accumulated optical depth as integrated between the limiting wavelengths λ_1 and λ_2 , and the transition and the calculated dynamic velocity is written. In the lower panel is shown the spectral data, with the integrated optical depth shown in green, and a bar marking the distance between the 5% and the 95% limits. The same procedure was tried with a couple more lines, showing very similar results.

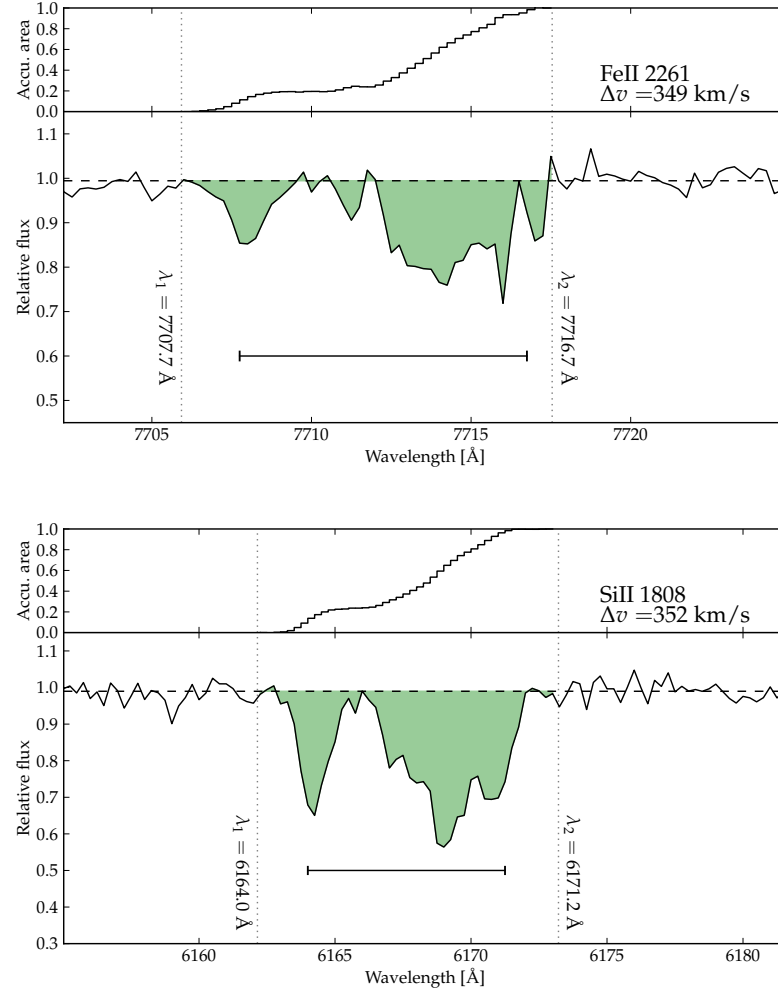


Figure 4.7: Dynamic velocity for the lines FeII 2261 and SiII 1808. Top panel in each is the cumulative integral of the optical depth over wavelength, bottom panel shows the line with the wavelength interval of integration and the dynamic width marked as well.

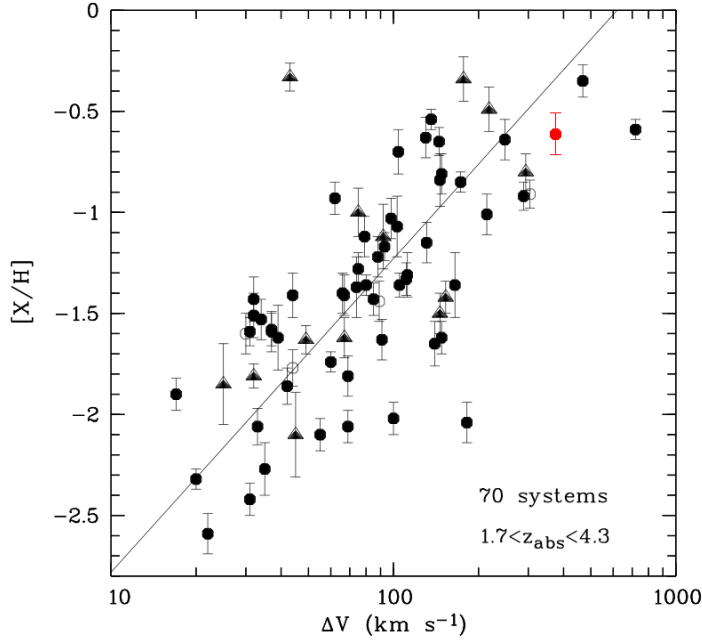


Figure 4.8: The plot of dynamic velocity-metallicity correlation from [22] which is also shown in figure 2.5 with the $z = 2.4124$ DLA system added in red.

In figure 4.8, the values found for SiII 1808 have been used to place the $z = 2.4124$ DLA system in the spectrum of Q 0918+1636 in the correlation found by Ledoux et al. in [22]. The plotting of the numbers is rough and should not be taken too literally, as it is done by hand, and is meant to only roughly show the systems properties in context. It is seen that the system is indeed comparatively metal rich and has a comparatively large dynamic velocity. It fits very nicely into the general picture of the correlation stated by Ledoux et al.

4.5 DISCUSSION

It was a bit surprising to find out that a thermal fit was more accurate than a turbulent fit, as a turbulent fit would normally be expected ¹, but the results were quite clear. While the metallicities found were in good agreement, the uncertainties were significantly better for a thermal fit, and `vpfit` dropped the tying of redshifts for one line component, which is hard to explain what even means physically. It should be noted here that the `vpfit`'s final result for redshifts and Doppler parameters seems to be quite sensitive to the initial values of the guesses,

¹J. Fynbo, conversation

Chapter 4. Analysis and results

but at least for this set of data, the inferred column densities have been quite robust to these changes.

The fitting software has consistently found two components in the high-ionization phase for the components of lowest redshift, while the higher redshift component has only one. When looking more closely, however, it looks like that same component actually consists of two components in the low-ionization phase as well, as the fits seem to be either too broad and shallow or too narrow and deep; and the line feature looks asymmetric. Whether there is much fit quality to be one is a different question.

It should also be noted that the initial guesses given for the two fits were not identical, even though the difference was small. It is very easy to adjust and polish the guess a bit using `vpguess`, which I apparently have done without thinking about it. I did however, after having finished the data reduction work, try to run a fit in a separate folder using the exact same parameters for the turbulent fit as is given for the thermal fit here. The result was a nicer looking fit, but degraded uncertainties compared to the turbulent fit presented here, and a metallicity practically indistinguishable from it.

The Robust lower limits and the metallicities derived from the fits are in a very good agreement, suggesting that there is little or no hidden saturation present and hence that the assumption of being in the optically thin limit is justified for at least the FeII 2250 and ZnII 2026 lines, and close to being so for SiII 1808.

POSSIBLY CORRELATED ERRORS

Looking at the residual plots in figures 4.5 and 4.6, there is visibly fewer than the expected 32% of the points lying outside the 1σ line. Looking at the output of `vpfit`, the χ^2/ndf is ~ 0.65 for both fits, significantly less than one which is unlikely for uncorrelated errors. A possible explanation for this could be that the data are rebinned as part of the X-Shooter pipeline, which could correlate the errors, resulting in an overestimate of the uncertainties of a factor of $\sqrt{2}$. What this signifies for the final results is not completely clear, as I have not had the time to try to propagate a simulated error, but it does mean that the errors of the metallicities found are overestimated by a factor of around $\sqrt{2}$, but at the worst of times, that pushes the agreement between each of the plots and the robust lower limit just outside the 1σ limit, and it does not affect the very good agreement between the nominal values of the two fits.

4.6 SUMMARY AND OUTLOOK

I have found that the $z = 2.4124$ DLA system seen in the spectrum of the quasar Q0918+1636 has a high metallicity for a DLA with a best fit for Zinc, which is the one metal least prone to dust depletion [25], of $Z_{\text{ZnII}} - Z_{\odot} = -0.57 \pm 0.09$, with a possible overestimation of the errors due to rebinning in the X-Shooter pipeline. The system shows evidence of dust depletion. It fits well into the correlation of [22], having a high-end metallicity and dynamic velocity, which also suggests a high mass, although this cannot be readily confirmed.

4.6.1 POSSIBLE FUTURE WORK

Even though it was clear from an early stage that there was no Lyman α emission to be found from the host galaxy of this DLA system, it is still possible that there could be emission from the system in other lines, e.g. $\text{H}\alpha$ and $\text{H}\beta$ at 6562 Å and 4862 Å rest frame wavelengths, respectively, or from OII or OIII I tried to do a tentative fit for H_2 , but it looked like the Lyman α Forest was too strong to get any consistent results. However, I would not consider the possibilities exhausted for that. Furthermore, it seems very possible that the data could be useful as part of a statistical sample for mapping the Lyman α forest itself.

Finally, the prospects of the work of Fynbo et al [12] [10] [11] are very promising, as the fraction of observed spectra that show Ly α emission from the DLA host when selected for high metallicity is indeed higher than for a random sample of DLA systems.

Appendices

APPENDIX A

REVERSE FLUX-CALIBRATION OF THE COMPOSITE QUASAR SPECTRUM

As described in section 3.3.3, in order to get the composite quasar data to resemble the spectrum of Q 0918+1636 as closely as possible, it is necessary to imitate the various extinction effects to which the light of the QSO has been exposed. These steps are described below:

A.1 EXTRAGALACTIC EXTINCTION

On its course through the Cosmos, the light of the quasar is subject to two different kinds of transformation. One is absorption features due to HI and metal transitions in intervening clouds, features that we do not want to replicate, since they are exactly the features of the QSO spectrum against which we want to hold a likely continuum.

Dust extinction and reddening, on the other hand, results in a large scale transformation of the entire spectrum, and a reconstruction of this is necessary in order to make the composite spectrum look like our data.

At least one of the two DLA systems on the light's path is sufficiently dust- and metal-rich to impose a considerable dust reddening on the QSO light. This reddening is reproduced by using the parametrization model developed by Pei et al. [28]. In this work, they have used the Magellanic Clouds as models for early dwarf galaxies. This model is in turn built on a Galactic dust composition with the abundances of graphite and silicates adjusted to reproduce the empirical extinction curves of the MCs. The Pei et al. extinction curve is given as:

Appendix A. Reverse flux-calibration of the composite quasar spectrum

SMALL MAGELLANIC CLOUD				
a_i	$\lambda_i[\mu\text{m}]$	b_i	n_i	K_i
185.	0.042	90.	2.0	2.89
27.	0.08	5.50	4.0	0.91
0.005	0.22	-1.95	2.0	0.02
0.010	9.7	-1.95	2.0	1.55
0.012	18.	-1.80	2.0	1.72
0.030	25.	0.0	2.0	1.89

Table A.1: The parameters for an SMC-like extinction curve, as given in [28].

$$f'(\lambda) = f(\lambda) \times 10^{-0.4 \times \zeta(\lambda) \times A_B}; \quad (\text{A.1})$$

$$\zeta(\lambda) = \sum_{n=1}^6 \frac{a_i}{(\lambda/\lambda_i)^{n_i} + (\lambda_i/\lambda)^{n_i} + b_i} \quad (\text{A.2})$$

with $f(\lambda)$ and $f'(\lambda)$ denoting the flux as a function before and after the reddening, respectively. The parameter values are empirical and given in [28] for the Galaxy, the Large Magellanic Cloud and the Small Magellanic Cloud. The reddening is applied as:

The data of Q 0918+1636 are best fitted by a SMC-like extinction curve with a total extinction of $A_B = 0.25$ and a redshift of $z_{\text{abs}} = 2.583$, telling that the DLA system treated in [11] is containing the bulk of dust encountered.

The parameters for the SMC is given in [28], table 4, and are shown in table A.1.

A.2 GALACTIC EXTINCTION AND REDDENING

Besides the dust extinction of the SMC-like system at $z = 2.583$, the Q 0918+1636 spectrum is further reddened by matter within the Galaxy, which also needs modeling. This is done using the results of Jason Cardelli et al. [6] with improved coefficients for the optical and near-infrared by O'Donnell et al. [27].

The Galactic extinction is given as:

$$f'(\lambda) = f(\lambda) \times 10^{0.4 \times A(\lambda)}; \quad (\text{A.3})$$

$$A(\lambda) = A_V \times \left(a + \frac{b}{R_V} \right) \quad (\text{A.4})$$

$$A_V = R_V \times E_{B-V} \quad (\text{A.5})$$

where R_V is the Galactic reddening, of which the standard value of $R_V = 3.1$ has been adopted; E_{B-V} is the color excess in magnitudes of the QSO light on its way through the Galaxy, for which a value of $E_{B-V} = -0.025$ has been used. The quantities a and b are an empirical set of polynomials in different (inverse) wavelength intervals, which are as follows:

INFRARED: $[0.3\mu\text{m}^{-1} \leq x \leq 1.1\mu\text{m}^{-1}]$
 $a(x) = 0.574x^{1.61} \quad (\text{A.6})$

$$b(x) = -0.527x^{1.61} \quad (\text{A.7})$$

NIR/OPTICAL: $[1.1\mu\text{m}^{-1} \leq x \leq 3.3\mu\text{m}^{-1}]$
AND: $y = x - 1.82$
 $a(x) = 1 + 0.104y - 0.609y^2 + 0.701y^3 + 1.147y^4$
 $\quad - 1.718y^5 - 0.827y^6 + 1.647y^7 - 0.505y^8 \quad (\text{A.8})$

$$b(x) = 1.952y + 2.908y^2 - 3.989y^3 - 7.985y^4$$

$$\quad + 11.102y^5 + 5.491y^6 - 10.805y^7 + 3.347y^8 \quad (\text{A.9})$$

ULTRAVIOLET: $[3.3\mu\text{m}^{-1} \leq x \leq 8\mu\text{m}^{-1}]$
 $a(x) = 1.752 - 0.316x - \frac{0.104}{(x - 4.67)^2 + 0.341} + F_a(x) \quad (\text{A.10})$

$$b(x) = -3.090 + 1.825x + \frac{1.206}{(x - 4.62)^2 + 0.263} + F_b(x) \quad (\text{A.11})$$

WITH:

$$F_a(x) = -0.04473(x - 5.9)^2 - 0.009779(x - 5.9)^3 \quad [5.9 \leq x \leq 8]$$

$$F_b(x) = 0.2130(x - 5.9)^2 + 0.1207(x - 5.9)^3 \quad [5.9 \leq x \leq 8]$$

$$F_a(x) = F_b(x) = 0 \quad [x < 5.9]$$

FAR-UV: $[8\mu\text{m}^{-1} \leq x \leq 10\mu\text{m}^{-1}]$

AND: $y = x - 8$

$$a(x) = -1.073 - 0.628y + 0.137y^2 - 0.070y^3 \quad (\text{A.12})$$

$$b(x) = 13.670 + 4.257y - 0.420y^2 + 0.374y^3 \quad (\text{A.13})$$

Appendix A. Reverse flux-calibration of the composite quasar spectrum

The parametrization is found in the IDL-program `CCM_UNRED.PRO`, which was rewritten in Python for this project as the script `unred.py`.

A.3 ATMOSPHERIC EXTINCTION AND INSTRUMENT RESPONSE

Finally, atmospheric extinction and the instrument response function need to be modeled. The extinction function is given as:

$$\varepsilon(\lambda) = 10^{0.4 \cdot b(\lambda)a} \quad (\text{A.14})$$

where $b(\lambda)$ is the atmospheric extinction per unit air mass at wavelength λ . Values for $b(\lambda)$ is given as a data file along with the instrument response function $R(\lambda)$. Since these are sharing the resolution of the X-Shooter and not the composite data, they need to be interpolated before being applied to the composite data as follows:

$$f'(\lambda) = \frac{f(\lambda)}{R(\lambda)} \cdot 10^{-0.4 \times b(\lambda) \times 1.4} \quad (\text{A.15})$$

with $f'(\lambda)$ denoting the response/extinction-corrected flux in the λ 'th wavelength bin, etc.; and the airmass set to 1.4 from the FITS headers of the data files.

The result of these steps is shown in figure 3.8

APPENDIX B

SOURCE CODE

In this appendix is printed a list of the pieces of code I’ve written for use in this work, and some example scripts are included for some of the more important operations. All the code will be available at <http://bit.ly/Thesis-Scripts>. Typically, slightly different versions of the scripts are used for each of the arms, although some of the scripts are carried out only once, typically in the VIS arm. When more than one version of a script exists, one version is shown and the differences from any other versions explained.

The code I’ve written for this work is listed below, with “ARM” denoting that a version was written for each arm, though mostly they only differ slightly:

makeitso-ARM.sh A convenience script for calling `esorex` to run the recipes of the X-Shooter pipeline, written in the Bash scripting language. This script is printed and further described below.

addings-ARM.py A Python script to perform sigma clipping and stacking of the 2D spectra output by the X-Shooter pipeline.

spextract-ARM.py A Python script to perform the extraction described in section 3.3.3.

normalizeARM.py This is where the normalization process of section 3.3.5 happens. This script is printed and introduced further below.

unred.py A Python implementation of the IDL program `CCM_UNRED.PRO`, which in turn is an implementation of the procedures published in [6], with improved coefficients for the near-UV by O’Donnell et al., 1994 [27]. This script is called from the normalization scripts to perform a simulated Galactic reddening of the composite quasar data.

Appendix B. Source Code

SkyLines.py Small, simple Python script to plot the spectral data of the non-sky-reduced spectrum along with the smoothed data of the UVES sky atlas to fine tune wavelength calibration, which was then manually changed in the FITS header in the script `prepforfitARM.py`.

DLAfit.py Included and described further below. This is a script that implements Torsten Tepper Garcías Voigt profile approximation from [34] to do the fitting of the DLA line described in section 4.1.

RemoveTellurics.py A script to use the telluric standard star to correct for telluric absorption. Only the VIS arm was corrected.

prepforfitARM.py A few small tasks collected in one script, including reference pixel correction based on comparison with sky lines, addition of the AIR keyword in the FITS header (Necessary for VPFIT to perform correctly) and other small, similar tasks.

RobLowe.py Named with a little hat tip to a hero of the cinema, this script reads the spectrum and calculates the robust lower limits as described in section 4.2. The lines and their wavelengths are hard-coded into the script.

Metallicities.py As the name suggests, calculates the metallicities of the lines fitted in `vpfit`. This script is significantly smarter than e.g. `RobLowe.py`, as it reads the list of fitted lines and line components and automatically looks up the relevant information about Solar abundances in the file `atom.dat`, which is distributed with `vpfit`.

PlotModel.py Like the name suggests, it plots the data along with the output model from `vpguess`. It reads the limits of the fitted regions from the `vpfit` input file and figures out which ions are present in each region, then finds the line centres of each component in the `vpfit` output file and plots these above the data and model spectrum. A few minor alterations of the code could make it quite usable.

B.1 PIPELINE-SCRIPT

The XShooter pipeline is made up of a number of so-called “recipes” callable by the esorex tool. The recipes take care of creating bad pixel maps, master flat field images etc., as described in section 3.3.1; only in the last recipe is the reduction done on the science images.

This is a script to run all or a subset of the recipes required to reduce my data. I have made switches available to run each recipe separately, and by listing switches the relevant recipes are run. For the science reduction recipe, it is possible to select one of the three science images or to have it use them all to create a combined spectrum from a median value of each pixel across the images.

The script does *not* yet check whether the files required to run a recipe are present, and neither does it yet have any error handling if an esorex run fails.

```

1  #!/bin/bash
2
3  ###=====
4  #                               Read options set
5  ###=====
6
7  HELP="Usage: $0 [--options] [number].                               \n
8  Accepted options:  \t --physmod   | -p: run the xsh_util_physmod routine.   \n
9  \t\t\t\t\t --bias      | -b: run the xsh_mbias receipt.             \n
10 \t\t\t\t\t --dark      | -d: run the xsh_mdark receipt.             \n
11 \t\t\t\t\t --predict  | -p: run the xsh_predict receipt.             \n
12 \t\t\t\t\t --orderpos | -o: run the xsh_orderpos receipt.            \n
13 \t\t\t\t\t --flat      | -f: run the xsh_mflat receipt.             \n
14 \t\t\t\t\t --2dmap    | -m: run the xsh_2dmap receipt.             \n
15 \t\t\t\t\t --scired   | -s: run the xsh_scired receipt.            \n
16 \t\t\t\t\t --help     | -h: print this help message and exit.      \n
17 Accepted numbers: 0, 1, 2, 3, 4, all, tell, none.                   \n
18 \t\t\t\t\t 1 means perform scired on file 1 only,                 \n
19 \t\t\t\t\t 2 means on file 2 only,                                   \n
20 \t\t\t\t\t 3 on file 3 only, and                                       \n
21 \t\t\t\t\t 4 or 'all' means on all.                                     \n
22 \t\t\t\t\t 'tell' means that the telluric standard frame is reduced instead of \n
23 \t\t\t\t\t the science frames.                                         \n
24 Only first number is read, any subsequent numbers passed are ignored. \n
25 None is synonymous with 0. Default is to perform all operations and perform \n
26 the Scired routine of all frames into one combined frame."
27
28
29 args=("$*")                  # Define an array of arguments passed.
30 echo "Args=("${args[@]}")"  # For debugging only.
31 options=(0 0 0 0 0 0 0 0)  # Option switches All off from the start:
32 optioncheck=0

```

Appendix B. Source Code

```
33 files=4                                # Files to operate on: all by default.
34
35 #Check which options are passed, switch on the relevant switches:
36 ###=====
37 #                                Check if help option is set:
38 ###=====
39
40 for i in $args; do
41     case "$i" in
42         --help | -h) echo -e $HELP; exit;;
43     esac
44 done
45
46 while [[ $1 == -* ]]; do
47     case "$1" in
48         --physmod | -p) options[0]=1; optioncheck=$optioncheck+1;
49             echo "Modeling receipt selected."; shift;;
50         --bias|-b) options[1]=1; optioncheck=$optioncheck+1;
51             echo "Bias receipt selected.";shift;;
52         --flat|-f) options[2]=1; optioncheck=$optioncheck+1;
53             echo "Flat field receipt selected.";shift;;
54         --dark|-d) options[3]=0;
55             echo "Dark receipt not available for visual arm.";exit;;
56         --predict|-e) options[4]=1; optioncheck=$optioncheck+1;
57             echo "Predict receipt selected.";shift;;
58         --2dmap|-m) options[5]=1; optioncheck=$optioncheck+1;
59             echo "2D Map receipt selected.";shift;;
60         --orderpos | -o) options[6]=1; optioncheck=$optioncheck+1;
61             echo "Orderpos receipt selected.";shift;;
62         --scired | -s) options[7]=1; optioncheck=$optioncheck+1;
63             echo "Scired receipt selected.";shift;;
64         *) echo "Invalid option: $1 "; echo $HELP; exit;;
65     esac
66 done
67 while [[ -n $1 && $1 != -* ]]; do
68     case "$1" in
69         1 ) files=1; break;;
70         2 ) files=2; break;;
71         3 ) files=3; break;;
72         "tell" | "Tell" | "telluric" | "Telluric" ) files=5; break;;
73         "all" ) files=4; break;;
74         "none" | "None" | "test" | "Test") filescheck=20;
75             files=0; echo "No files/test mode chosen"; break;;
76         * ) echo $i":  invalid option";;
77     esac
78 done
79
80 echo "optioncheck = " $optioncheck
81
```

```

82  ###=====
83  #           If no options are selected; select all,
84  #           except if none or test are passed, then switch all off.
85  ###=====
86
87
88  if [[ $optioncheck -eq 0 ]]; then
89      for i in ${!options[*]}; do
90          options[$i]=1
91      done
92  fi
93
94  physmod=${options[0]}
95  bias=${options[1]}
96  flat=${options[2]}
97  dark=${options[3]}
98  predict=${options[4]}
99  map=${options[5]}
100  orderpos=${options[6]}
101  scired=${options[7]}
102
103  #echo ${options[*]}
104  echo "File(s) chosen (4 means all):"$files
105  echo
106  echo "Physmod switch:      "$physmod
107  echo "Bias switch:         "$bias
108  echo "Flat field switch:    "$flat
109  echo "Dark switch:         "$dark
110  echo "Predict switch:      "$predict
111  echo "2D Map switch:       "$map
112  echo "Orderpos map         "$orderpos
113  echo "Science Reduction:   "$scired
114
115  if [[ $filescheck -eq 20 ]]; then
116      echo "Oh! testmode with no actual runs is set. Switching them all off now!";
117      for i in ${!options[*]}; do
118          options[$i]=0
119      done
120  fi
121
122  physmod=${options[0]}
123  bias=${options[1]}
124  flat=${options[2]}
125  dark=${options[3]}
126  predict=${options[4]}
127  map=${options[5]}
128  orderpos=${options[6]}
129  scired=${options[7]}
130

```

Appendix B. Source Code

```
131 echo -e ${options[*]}
132 echo -e "File(s) chosen (4 means all):"$files
133 echo -e
134 echo -e "Physmod switch:      \t "$physmod
135 echo -e "Bias switch:           \t "$bias
136 echo -e "Flat field switch:     \t "$flat
137 echo -e "Dark switch:         \t "$dark
138 echo -e "Predict switch:       \t "$predict
139 echo -e "2D Map switch:        \t "$map
140 echo -e "Orderpos map           \t "$orderpos
141 echo -e "Science Reduction:    \t "$scired
142
143
144
145 ###=====
146 #      Model instrument
147 ###=====
148 if [[ $physmod -eq 1 ]]; then
149     esorex xsh_util_physmod xsh_util_physmod_vis.sof
150     mv out_0000.fits prod/THEO_TAB_MULT_VIS.fits
151     mv out_0001.fits prod/THEO_TAB_IFU_VIS.fits
152     mv out_0002.fits prod/THEO_TAB_SING_VIS.fits
153     #      rm *fits *log *paf
154 fi
155
156 ###=====
157 #      Make master biases for OBJECT and
158 #      FMTCHK/ORDERDEF, respectively
159 ###=====
160
161 if [[ $bias -eq 1 ]]; then
162     #OBJECT
163     esorex xsh_mbias xsh_mbias_vis_100_1x2.sof
164     mv out_0000.fits prod/MASTER_BIAS_VIS_100_1x2.fits
165     #      rm *fits *log *paf
166     #FMTCHK/ORDERDEF:
167     esorex xsh_mbias xsh_mbias_vis_400_1x1.sof
168     mv out_0000.fits prod/MASTER_BIAS_VIS_400_1x1.fits
169     #      rm *fits *log *paf
170 fi
171
172 ###=====
173 #      Predict order positions from
174 #      qualified guesses
175 ###=====
176
177 if [[ $predict -eq 1 ]]; then
178     esorex --time xsh_predict
179     --keep-temp=yes
```

```

180         --detectarclines-fit-window-half-size=5      \
181         --detectarclines-search-window-half-size=2    \
182         --detectarclines-min-sn=2                     \
183         xsh_predict_vis.sof
184     mv out_0001.fits prod/WAVE_TAB_GUESS_VIS_1x1.fits
185     mv out_0002.fits prod/ORDER_TAB_GUESS_VIS_1x1.fits
186     #   rm *fits *paf *log
187 fi
188
189 ###=====
190 #       Make table of line order positions
191 ###=====
192
193 if [[ $orderpos -eq 1 ]]; then
194     esorex --time xsh_orderpos                        \
195           xsh_orderpos_vis.sof
196     cp -f out_0000.fits prod/ORDER_TAB_CENTR_VIS_1x1.fits
197     #   \rm *fits *log *paf
198 fi
199
200 ###=====
201 #       Create master flats
202 ###=====
203
204 if [[ $flat -eq 1 ]]; then
205     esorex xsh_mflat                                \
206           --keep-temp=no                            \
207           --detectorder-edges-flux-thresh=0.2       \
208           --background-radius-x=1                   \
209           --detectorder-min-sn=10                    \
210           xsh_mflat_vis_100_1.2_1x2.sof
211     cp -f out_0000.fits prod/ORDER_TAB_EDGES_SLIT_VIS_100_1.2_1x2.fits
212     cp -f out_0001.fits prod/MASTER_FLAT_SLIT_VIS_100_1.2_1x2.fits
213     #   \rm *fits *log *paf
214 fi
215
216
217 ###=====
218 #       Create wavelength calibration maps
219 ###=====
220
221 if [[ $map -eq 1 ]]; then
222     esorex xsh_2dmap                                \
223           --keep-temp=no                            \
224           xsh_2dmap_vis.sof
225     cp -f out_0001.fits prod/WAVE_TAB_2D_VIS_1x1.fits
226     cp -f out_0003.fits prod/WAVE_MAP_VIS_1x1.fits
227     cp -f out_0005.fits prod/DISP_TAB_VIS_1x1.fits
228     #   \rm *fits *log *paf

```

Appendix B. Source Code

```
229 fi
230
231 ###=====
232 #       Perform reduction of science data
233 ###=====
234
235 if [[ $scired -eq 1 ]]; then
236     case $files in
237         1) sofname="xsh_scired_slit_stare_vis1.sof";
238            foldername="final_1";
239            number="1";;
240         2) sofname="xsh_scired_slit_stare_vis2.sof";
241            foldername="final_2";
242            number="2";;
243         3) sofname="xsh_scired_slit_stare_vis3.sof";
244            foldername="final_3";
245            number="3";;
246         4) sofname="xsh_scired_slit_stare_vis.sof";
247            foldername="final_all";
248            number="all";;
249         5) sofname="xsh_scired_slit_stare_vis_tell.sof";
250            foldername="final_tell";
251            number="tell";;
252     esac
253
254     if [[ ! -d end_results/$foldername ]]; then
255         mkdir end_results/$foldername
256         echo "Created dir end_results/"$foldername
257     fi
258
259     esorex --time xsh_scired_slit_stare \
260            --keep-temp=no \
261            --rectify-bin-lambda=0.025 \
262            --rectify-bin-slit=0.15 \
263            --rectify-consume-flux=FALSE \
264            --rectify-radius=1 \
265            --background-nb-y=40 \
266            --background-radius-y=50 \
267            --background-method=poly \
268            --sky-slit-edges-mask=1.5 \
269            --sky-subtract=TRUE \
270            --sky-method=MEDIAN \
271            --mergeord-method=0 \
272            --removecrhsingle-niter=10 \
273            --compute-map=TRUE \
274            --extract-method=FULL \
275            --localize-slit-position=-0 \
276            --localize-slit-hheight=1 \
277            $sofname
```



```
278 #           $sofname
279 #           --localize-slit-position=-2      \
280 #           --localize-slit-hheight=1        \
281
282 cp out_0000.fits end_results/$foldername/
283 cp out_0001.fits end_results/$foldername/
284 cp out_0002.fits end_results/$foldername/
285 cp out_0003.fits end_results/$foldername/
286 #   \rm *fits *paf *log
287 fi
```

B.2 NORMALIZATION OF VIS ARM

This is a script to normalize the one-dimensional spectrum once it is extracted, and is written in Python. The script starts out drawing the extracted spectrum along with a composite quasar spectrum by [33], corrected for reddening, instrument response function etc. The user is then prompted to draw what is believed to be a likely continuum spectrum by setting off a number of points with the mouse. A left click sets a point, a right click removes the latest point, an upwards scroll moves the spectrum to the left downward moves it to the right, and once the user is satisfied with the continuum drawn, a middle click sends a signal to the program that the continuum drawing is finished. The program then interpolates the drawn points by a cubic spline function and divides the original spectrum by this continuum.

It can sometimes be difficult to see whether a feature is indeed a spectrum feature, or it is instead noise, reduction artifacts or similar unwanted features. However, the two dimensional image is often very helpful, since spectral features are only present in the thin spectrum line in the middle, while sky lines and pipeline artifacts are often visible in the full height of the image. For this reason, above the main plot is shown the 2D spectrum image for the corresponding wavelengths. Likewise, the wavelengths of a number of select lines are marked with vertical lines and name tags containing the transition ion name and laboratory wavelength.

```

1  #!/usr/bin/env python
2  # The LineBuilder class is built on the 20 lines example code snippet
3  # LineBuilder in the online Matplotlib documentation at http://goo.gl/RLGS .
4  # I have modified and extended it to serve my needs.
5
6  import matplotlib as mpl
7  mpl.rc('font',**{'family':'sans-serif','sans-serif':['Helvetica']})
8  mpl.rc('font',**{'family':'serif','serif':['Palatino']})
9  mpl.rc('text', usetex=True)           #For pretty LaTeX fonts in plots
10 import matplotlib.pyplot as plt
11 import numpy as np
12 import scipy as sp
13 import scipy.interpolate as ipl
14 import math
15 import unred
16 import pyfits as pf
17 import os
18 import matplotlib.gridspec as grisp
19
20 #####=====

```

```

21  ### The function that builds the line. Left clicking adds a point on
22  ### the line, right clicking removes the one last added.
23  ### Points are automatically ordered, so a point can be added
24  ### between two already existing points.
25  #####
26
27  class LineBuilder:
28  ###====
29  # Draws piecewise linear function from mouse click,
30  # returns clicked point coordinates both as ordered array and
31  # in the order clicked (the latter draws nice doodles, is only
32  # returned because it was easier than not to)
33  #
34  ###=====
35  def __init__(self, line):
36      global VISSpec
37      global l
38      ax.axis([lstart, lstart+steps1, 0, annmax])
39      print lstart, l.min()
40      self.line = line
41      self.xs = np.array(line.get_xdata())
42      self.ys = np.array(line.get_ydata())
43      self.axes = np.array(ax.axis())
44      self.imglims = (self.axes[0:2]-SpecHead['CRVAL1'])/lstep
45      self.cid = line.figure.canvas.mpl_connect('button_press_event', self)
46      self.did = line.figure.canvas.mpl_connect('scroll_event', self)
47
48  def __call__(self, event):
49      print 'click', event.button, event.xdata, event.ydata
50
51      if (event.button == 1):
52          if event.inaxes!=self.line.axes: return
53          self.xs = np.append(self.xs, event.xdata)
54          self.ys = np.append(self.ys, event.ydata)
55          self.line.set_data(self.xs[self.xs.ravel().argsort()],
56                             self.ys[self.xs.ravel().argsort()])
57          self.line.figure.canvas.draw()
58      if (event.button == 2):
59          self.xsord = self.xs[self.xs.ravel().argsort()]
60          self.ysord = self.ys[self.xs.ravel().argsort()]
61          plt.close()
62      if (event.button == 3):
63          self.xs = np.delete(self.xs, -1)
64          self.ys = np.delete(self.ys, -1)
65          self.line.set_data(self.xs[self.xs.ravel().argsort()],
66                             self.ys[self.xs.ravel().argsort()])
67          self.line.figure.canvas.draw()
68      if (event.button == 'up'):
69          self.axes[0] = self.axes[0]+550

```

Appendix B. Source Code

```
70         self.axes[1] = self.axes[1]+550
71         self.axes[2] = 0.
72         self.axes[3] = VISSpec[np.where((1 > self.axes[0]) &
73             (1 < self.axes[1]))].max()*1.2
74         ax.axis(self.axes)
75         self.imglims = self.imglims+550/lstep
76         img.axis([self.imglims[0], self.imglims[1], 0, 72])
77         annmax = self.axes[3]*0.9
78         self.line.figure.canvas.draw()
79     if (event.button == 'down'):
80         self.axes[0] = self.axes[0]-550
81         self.axes[1] = self.axes[1]-550
82         self.axes[2] = 0.
83         self.axes[3] = VISSpec[np.where((1 > self.axes[0]) &
84             (1 < self.axes[1]))].max()*1.2
85         ax.axis(self.axes)
86         self.imglims = self.imglims-550/lstep
87         img.axis([self.imglims[0], self.imglims[1], 0, 72])
88         annmax = self.axes[3]*0.9
89         self.line.figure.canvas.draw()
90
91
92 #####=====
93 #####           And now for the actual script
94 #####=====
95
96 ##     For testing purposes, switch the slow and tedious
97 ##     point selecting on and off.
98 pickpoints = False
99 pickpoints = True
100
101 ## Open image data files
102 specimghdu = pf.open('sumVIS.fits')
103 specimg = specimghdu[0].data
104 VISSpecFile = pf.open('SpecVIS.fits')
105 VISSpec = VISSpecFile[0].data
106 VISErrFile = pf.open('SpecVIS.err.fits')
107 VISErr = VISErrFile[0].data
108 SpecHead = VISSpecFile[0].header
109 ErrHead = VISErrFile[0].header
110
111 ### The pixel-to-wavelength conversion
112 lstep = SpecHead['CD1_1']
113 lZero = SpecHead['CRVAL1']-(SpecHead['CRPIX1']-1)*SpecHead['CD1_1']
114
115 ### How many of the following vars are actually necessary?
116 ###     Might want to trim them...
117 nbins = VISSpec.shape[0]
118 l = lZero+lstep*np.arange(0, nbins, dtype='float')
```

```

119 lstart = l.min()+100
120 lend = l.max()
121 steps1 = 800.
122
123
124 #####=====
125 ### Experimental: how about we also plot the
126 ### composite spectrum in the visual range?
127 #####=====
128
129 tab = np.genfromtxt('compo.data')
130 wl = tab[:, 0]*4.09
131 flux = tab[:, 1]*6.3
132 fluxred = unred.ccm(wl,flux,-0.025)
133
134 #####=====
135 ### Pei et al. parametrization curve:
136 #
137 # Now: the Pei et al. parametrization
138 # for SMC-like extinction curve.
139 # Formula:
140 #
141 #
142 #
143 #
144 # Or in \LaTeX:
145 #
146 #
147 #
148 #####=====
149
150 ai = np.array([185.,27.,0.005,0.010,0.012,0.030])
151 wli = np.array([0.042,0.08,0.22,9.7,18.,25.])
152 bi = np.array([90.,5.50,-1.95,-1.95,-1.80,0.0])
153 ni = np.array([2.0,4.0,2.0,2.0,2.0,2.0])
154 Ki = np.array([2.89,0.91,0.02,1.55,1.72,1.89])
155 AB = 0.25
156
157 Alambda = np.zeros((flux.shape[0]))
158 wlr = wl/3.583/1e4
159
160 for n in range(0, 5):
161     Alambda = Alambda + ai[n]/((wlr/wli[n])**ni[n]+(wli[n]/wlr)**ni[n]+bi[n])
162
163 Alambda = Alambda*AB
164 flux = fluxred*10**(-0.4*Alambda)
165
166

```

Appendix B. Source Code

```
167 #####=====
168 ### Correct for response and atmospheric extinction:
169 #####=====
170
171 extinc = np.genfromtxt('../Response/wlext.dat', unpack=True)
172 RespVIS = np.genfromtxt('../Response/respVIS_interp.dat', unpack=True)
173 RespVIS[0, :] = RespVIS[0, :]*10
174 RespVIS[1, :] = RespVIS[1, :]*1e14
175
176 # Because response and extinction have different
177 # resolution than data: interpolation.
178
179 resptck = ipl.splrep(RespVIS[0, :], RespVIS[1, :])
180 response = ipl.splev(wl, resptck, der=0)
181
182 extitck = ipl.splrep(extinc[0, :], extinc[1, :])
183 extinction = ipl.splev(wl, extitck)
184
185 # Use the above to correct the flux data:
186 # (Theory behind formula...?)
187
188 flux = flux/response * 10.**(-0.4*extinction*1.4)
189
190
191
192 #####=====
193 ### Plot spectrum, composite spectrum and draw
194 ### new continuum by eye.
195 #####=====
196
197 ### If-loop to be able to turn the line drawing off.
198
199 if(pickpoints==True):
200     fig1 = plt.figure(1)
201     gs = grisp.GridSpec(2, 1, height_ratios=[1, 6])
202
203     # Upper plot: 2d spectrum to help identify
204     # residual sky lines and other
205     # features of noise.
206     imgstart = (lstart-SpecHead['CRVAL1'])/lstep
207     imgstep = steps1/lstep
208     img = fig1.add_subplot(gs[0])
209     plt.imshow(specimg, cmap='hot', vmin=-40, vmax=140, aspect=2.5)
210     tickls=img.get_xticklabels()+img.get_yticklabels()
211     plt.setp(tickls, visible = False)
212
213     # Draw approximate positions of ion lines
214     # to hint at where extra care is needed.
215     if os.path.isfile('linestofind.data'):
```

```

216         lineguides = np.loadtxt('linestofind.data', dtype=[
217             ('ion', 'S4'), ('labwl', 'S4'),
218             ('rs1', 'float32'), ('rs2', 'float32')])
219
220         linenum = lineguides['ion'].shape[0]
221
222         for n in lineguides['rs1']:
223             plt.axvline(x=(n-SpecHead['CRVAL1'])/lstep, linewidth=1.3,
224                 ymin=0, ymax=0.9, linestyle='--', color='w')
225
226         for n in lineguides['rs2']:
227             plt.axvline(x=(n-SpecHead['CRVAL1'])/lstep, linewidth=1.3,
228                 ymin=0, ymax=0.9, linestyle='-.', color='w')
229
230     img.axis([imgstart, imgstep, 0, 72])
231
232
233     # Lower plot: where the action goes on.
234     ax = fig1.add_subplot(gs[1])
235     fig1.suptitle(r'\rm{\Large{Click to draw continuum}}')
236     plt.plot(1, VISSpec, 'black')
237     plt.plot(wl, flux)
238     uerr = VISSpec+VISSpec*VISErr
239     lerr = VISSpec-VISSpec*VISErr
240     annmax = 1.2*VISSpec[lstart+50:lstart+50+steps1].max()
241     plt.subplots_adjust(hspace=0.05)
242
243
244     ###=====
245     ## Show any old chosen points to help figuring out
246     ##   where to put new ones...
247     ##   Also, put approximate wavelebgths of atomic
248     ##   lines in there
249     ###=====
250
251     if os.path.isfile('pointstable.data'):
252         oldpoints = np.genfromtxt('pointstable.data')
253         print 'Shape of oldpoints array: ', oldpoints.shape
254         plt.plot(oldpoints[0, :], oldpoints[1, :], 'yo')
255     #     if os.path.isfile('normtabVIS.tab'):
256     #         johanspoints = np.genfromtxt('normtabVIS.tab')
257     #         plt.plot(johanspoints[:, 0], johanspoints[:, 1], 'r~')
258
259     if os.path.isfile('linestofind.data'):
260         lineguides = np.loadtxt('linestofind.data', dtype=[
261             ('ion', 'S4'), ('labwl', 'S4'),
262             ('rs1', 'float32'), ('rs2', 'float32')])
263
264         linenum = lineguides['ion'].shape[0]

```

Appendix B. Source Code

```
265         for n in np.arange(linenum):
266             plt.axvline(x=lineguides['rs1'][n],
267                         ymin=0, ymax=1, linestyle='--', color='g')
268             plt.axvline(x=lineguides['rs2'][n],
269                         ymin=0, ymax=1, linestyle='-.', color='g')
270
271             string = lineguides['ion'][n]+' '+lineguides['labwl'][n]
272             print string
273
274             img.annotate(string, bbox=dict(boxstyle="round", fc="0.8"),
275                         xy=((lineguides['rs1'][n]-SpecHead['CRVAL1'])/lstep, 45),
276                         xycoords='data', color='black', size='medium', ha='center')
277             img.annotate(string, bbox=dict(boxstyle="round", fc="0.9"),
278                         xy=((lineguides['rs2'][n]-SpecHead['CRVAL1'])/lstep, 50),
279                         xycoords='data', color='black', size='medium', ha='center')
280
281
282         line, = ax.plot([lstart], [1.], 'red') # empty line
283         ax.axis([lstart, lstart+steps1, 0,
284                 annmax])
285         img.axis([(lstart-SpecHead['CRVAL1'])/lstep,
286                  (lstart-SpecHead['CRVAL1']+steps1)/lstep, 0, 72])
287         linebuilder = LineBuilder(line)
288         plt.show()
289
290         xcoords = np.array(linebuilder.xs)
291         ycoords = np.array(linebuilder.ys)
292         xord = np.array(linebuilder.xsord)
293         yord = np.array(linebuilder.ysord)
294         points = np.vstack( (xord, yord) )
295         np.savetxt('pointstable.data', points)
296
297         points = np.genfromtxt('pointstable.data')
298         xord = points[0, :]
299         yord = points[1, :]
300         tck = ipl.splrep(xord, yord)
301         xxx = ipl.splev(l, tck)
302         norm = VIISpec/xxx
303         errnorm = VISErr/xxx
304
305         if os.path.isfile('NormVIS.fits'):
306             os.remove('NormVIS.fits')
307         pf.writeto('NormVIS.fits', norm, SpecHead)
308
309         if os.path.isfile('NormVIS.err.fits'):
310             os.remove('NormVIS.err.fits')
311
312         pf.writeto('NormVIS.err.fits', errnorm, SpecHead)
313
```



```

314
315 #####=====
316 ### The following was just to check the functionality of my
317 ### script. Actual visualisation for thesis to be decided.
318 ### But all necessary arrays and data are loaded
319 ### at this point.
320 #####=====
321
322 plt.figure(2)
323 plt.suptitle(r'\rm{\Large{Normalization of QSO spectrum}}')
324 plt.subplots_adjust(hspace=0.02)
325
326 plot1 = plt.subplot(211)
327 specplot = plt.plot(l, VISSpec, 'black', label=r'\rm{\small{Spectrum}}')
328 mycont = plt.plot(l, xxx, 'r--', label=r'\rm{\small{Continuum}}')
329 mypoints = plt.plot(xord, yord, 'yo', label = r'\rm{\small{Points picked}}')
330 plt.axis([6000, 6500, 0, 170])
331 plt.figlegend((specplot, mycont, mypoints),
332              (r'\rm{\small{Spectrum}}',
333               r'\rm{\small{Continuum}}', r'\rm{\small{Points picked}}'),
334              loc = 'upper right', shadow=True, fancybox=True, labelspace=0.,
335              bbox_to_anchor=[0.94, 0.94])
336 xticklabels = plot1.get_xticklabels()
337 plt.setp(xticklabels, visible=False)
338 plt.ylabel(r'\rm{Intensity in photon counts}')
339
340 plot2 = plt.subplot(212, sharex=plot1)
341 plt.plot(l, norm, 'black', label=r'\rm{Reduced spectrum}')
342 plt.axhline(y=1, xmin=0, xmax=1, linewidth=1.5, color='red',
343            label=r'\rm{Continuum level}', linestyle = '--' )
344 plt.axis([6000, 6500, 0, 1.3])
345 plt.xlabel(r'\rm{Wavelength in \AA}')
346 plt.ylabel(r'\rm{Relative intensity}')
347 plt.savefig('Normalization.pdf')
348 plt.show()

```

B.3 DLA FITTING ROUTINE

This script implements the Voigt profile approximation by García [34], taking as input column density and Doppler parameter and returning a Voigt absorption profile.

The guessed values are written into the script along with the estimated error as described in section 4.1 and plots the QSO spectrum along with the Voigt profiles of best fit and upper/lower 1σ levels. The input values can then be varied until a satisfactory result has been reached.

```

1  #!/usr/bin/env python
2  # -*- coding: utf-8 -*-
3  import numpy as np
4  import scipy as sp
5  import scipy.constants as con
6  import pyfits as pf
7  import matplotlib as mpl
8  mpl.rc('text', usetex=True) #For pretty LaTeX fonts in plots
9  mpl.rc('font',**{'family':'sans-serif','sans-serif':['Helvetica']})
10  mpl.rc('font',**{'family':'serif','serif':['Palatino']})
11  mpl.rc('legend', shadow=True,
12         fancybox=True, borderpad=.5,
13         labelspace=.1,
14         fontsize='medium')
15  mpl.rcParams['xtick.labelsize'] = 'large'
16  mpl.rcParams['ytick.labelsize'] = 'large'
17  mpl.rcParams['lines.linewidth'] = .5
18  import matplotlib.pyplot as plt
19  import os
20
21  ###=====
22  # Voigt function:
23  #
24  # From Tepper García:
25  #
26  #  $\tau(\lambda) = (C N a) \cdot H[a, x(\lambda)]$  , where  $H[a, x(\lambda)]$ 
27  # is the Voigt function, of which the mathematical description and computer
28  # implementation is given in García.
29  # From same;  $C N a = \frac{e^2}{m_e c} \frac{\sqrt{\pi^3} f \lambda}{\pi b} \cdot a$ 
30  #
31  #
32  ###=====
33
34  def H(a, x):
35      P = x**2
36      H0 = np.e**(-(x**2))
37      Q = 1.5/x**2

```

```

38     H = H0-(a/np.sqrt(np.pi)/P*(H0*H0*(4*P*P+7*P+4+Q)-Q-1))
39     print P, H0, Q
40     return H
41
42
43     ###=====
44     #     Load data:
45     ###=====
46
47     specHDU = pf.open('NormUVB.fits')
48     spectrum = specHDU[0].data
49     head = specHDU[0].header
50     print spectrum.shape
51
52
53     ###=====
54     #     Misc. parameters,
55     #     wavelength calibrations etc.:
56     ###=====
57
58     #     CGS units:
59     cgsc = 2.998e10
60     cgsme = 9.1095e-28
61     cgse = 4.8032e-10
62
63     x = np.arange(spectrum.shape[0])
64     wavlen = x*head['CD1_1']+head['CRVAL1']
65     w10 = 1215.67*1.e-10 #in m
66
67
68     ###=====
69     #     Resonator strength and
70     #     Doppler parameter.
71     ###=====
72
73     f = 0.516
74     b = 2.e1 #Johans orig.
75     b = 2.e1
76     dwld = b/con.c*w10
77     gamma = 6.265e8 #Original
78     a = w10*gamma/(4.*np.pi*b)
79
80
81     ###=====
82     #     The (C N a) part:
83     ###=====
84
85     Ca = cgse**2/(cgsme * cgsc)*np.sqrt(np.pi**3)/np.pi*f*w10/b
86     N1 = 14.5e20 # This is our main knob to turn.

```

Appendix B. Source Code

```
87 N2 = 8.3e20                                # Same for the other line...
88
89 deltalogN1 = 0.06
90 logN1 = np.log10(N1)
91 hilogN1 = logN1+deltalogN1
92 lologN1 = logN1-deltalogN1
93 loN1 = 10.**lologN1
94 hiN1 = 10.**hilogN1
95 wl = (np.arange(10000)/100.+1150.)*1.e-10
96 x = (wl-wl0)/dwld
97 binsize = wl[11.]-wl[10.]
98
99
100 ###=====
101 # Set different parameters plus
102 # the extinction function:
103 ###=====
104
105 z1=2.4123 #Original
106 z1=2.4124
107 z2=2.582
108
109 extfun1 = np.e**(-Ca*N1*H(a, x))
110 extfun2 = np.e**(-Ca*N2*H(a, x))
111 loextfun1 = np.e**(-Ca*loN1*H(a, x))
112 hiextfun1 = np.e**(-Ca*hiN1*H(a, x))
113 wavlen21 = wl*1.e10*(1+z1)
114 wavlen22 = wl*1.e10*(1+z2)
115
116
117 ###=====
118 # Create figure of the plots fitted line profiles.
119 ###=====
120
121 fig1 = plt.figure(1, figsize=(7, 4.2))
122
123 plt.subplot(111)
124 specplot = plt.plot(wavlen, spectrum, 'black', label='Q 0918+1636')
125 voigt1 = plt.plot(wavlen21, extfun1, linestyle='-', linewidth=1.5,
126                  label='Voigt profile 1', color='red')
127 errdots = plt.plot(wavlen21, loextfun1, 'r', linestyle=':', alpha=.7)
128 plt.plot(wavlen21, hiextfun1, 'r', linestyle=':', alpha=.7,
129          label=r'Profile 1 \pm errors')
130 voigt2 = plt.plot(wavlen22, extfun2, 'b-', linewidth=1.5,
131                  label='Voigt profile 2')
132
133 # Because Tepper García algorithm is weird, only fill between wings,
134 # not line center
135 plt.fill_between(wavlen21[:6500], loextfun1[:6500], hiextfun1[:6500],
```

```

136         color='grey', alpha=.3)
137 plt.fill_between(wavlen21[6600:], loextfun1[6600:], hiextfun1[6600:],
138                 color='grey', alpha=.3)
139 plt.axis([4010, 4300, 0, 1.2])
140 plt.xlabel(r'\Large{Wavelength [\AA]}')
141 plt.ylabel(r'\Large{Relative flux}')
142 contline = plt.axhline(y=1, xmin=0, xmax=1, linestyle='--', color='black',
143                       label='Continuum')
144 plt.legend(loc='lower right')
145 lt.subplots_adjust(right=.92, left=.08)
146
147 # Subplot2:
148 plt.subplot(212)
149 plt.plot(wavlen, spectrum, 'black', label='Spectrum')
150 voigt2 = plt.plot(wavlen22, extfun2, 'b-', label='Voigt profile fit 2')
151 plt.axhline(y=1, xmin=0, xmax=1, linestyle=':', color='black',
152            label='Continuum')
153 plt.plot(wavlen21, extfun1, 'r-', label='Voigt profile fit 1')
154 plt.plot(wavlen21, loextfun1, 'r-', linestyle=':')
155 plt.plot(wavlen21, hiextfun1, 'r-', linestyle=':')
156 plt.xlabel(r'Wavelength [\AA]')
157 plt.axis([4230, 4480, 0, 1.3])
158
159 # Add figure wide legend + save figure
160 plt.figlegend((specplot, voigt1, errdots, voigt2, contline),
161              (r'\rm{Q 0918+1636}', r'\rm{Voigt profile 1}',
162               r'\rm{Profile 1 \pm errors}', r'\rm{Voigt profile 2}',
163               r'\rm{Continuum}'),
164              'lower right', fancybox=True, labelspring=0., shadow=True)
165 plt.savefig('doublevoigt.pdf')
166 plt.show()
167
168
169 ###=====
170 # Write fitted parameters to file:
171 ###=====
172
173 if os.path.isfile('dlafit.params'):
174     os.remove('dlafit.params')
175
176 paramsfile = open('dlafit.params', 'w')
177 paramsfile.write('Value of N1: \t\t'+str(N1)+'\n')
178 paramsfile.write('Value of log(N1): \t'+str(np.log10(N1))+'\n')
179 paramsfile.write('Error of log(N1): \t\t'+str(deltalogN1)+'\n')
180 paramsfile.write('Value of N2: \t\t'+str(N2)+'\n')
181 paramsfile.write('Value of log(N2): \t'+str(np.log10(N2))+'\n')
182 paramsfile.write('Value of N2: \t \t'+str(N2)+'\n')
183 paramsfile.write('z1 and z2: \t'+str(z1)+' , '+str(z2)+'\n')
184 paramsfile.close()

```

B.4 METALLICITY CALCULATION

This is the script in which metallicities are calculated for the ions fitted in `vpfit`. The script reads the output file from the fit, sorts the ion names, calculates the appropriate column densities and error propagations, looks up the relevant solar abundances and calculates the metallicity in terms of solar abundances. Error propagation is done in two ways; one is by an implementation of the procedure described in section 2.3; the other is the Python package `uncertainties`. Reassuringly, they reach identical results.

```
1  #!/usr/bin/env python
2  # -*- coding: utf-8 -*-
3  # Import standard libraries:
4
5  import numpy as np
6  import os.path as op
7  import sys
8  import uncertainties as uncs
9  import uncertainties.umath as umc
10 import uncertainties.unumpy as unp
11
12 ###=====
13 #          Put constants etc. here, and for compatibility with other systems:
14
15
16 NeuHyd = 14.5E20
17 for line in open('../UVB/dlafit.params'):
18     if 'log(N1)' in line:
19         temp = line.split()
20         print float(temp[-1])
21         if 'Error' in line:
22             DelNeuHyd = float(temp[-1])
23         else:
24             LogNeuHyd = float(temp[-1])
25
26 print 'NeuHyd before: ', NeuHyd
27 print 'Log NeuHyd before: ', np.log10(NeuHyd)
28 NeuHyd = 10.**LogNeuHyd
29 HiNeuHyd = 10.** (LogNeuHyd+DelNeuHyd)
30 LoNeuHyd = 10.** (LogNeuHyd-DelNeuHyd)
31
32 print NeuHyd, ' + ', HiNeuHyd, ' + ', LoNeuHyd
33 uLogNeuHyd = uncs.ufloat((LogNeuHyd, DelNeuHyd))
34 print uLogNeuHyd
35 uNeuHyd = 10**uLogNeuHyd
36 print uNeuHyd
37
38 # Johans vardi:
```

```

39 #NeuHyd = 1.499684835502368e+21
40
41 if not op.isfile('./fort18-tail.data'):
42     print 'Please make sure the VPfit output'
43     print 'file is placed in the working directory'
44     sys.exit()
45
46
47 ###=====
48 #         Load VPfit results from fort.26 type file and extract the lines with
49 #         actual data into a numpy array.
50
51 ResultList = list([])
52 n = 0
53 #for line in open('./vpguess-turb/fort18-tail.data'):
54 for line in open('./vpguess-therm/fort18-tail.data'):
55 #for line in open('./vpguess-experimental/fort18-tail.data'):
56     n += 1
57
58     if line.startswith(('%', '!')):
59         print 'Ignored comment line', n
60     else:
61         temp = line.split()
62
63         if len(temp[0]) <= 1:
64             temp[0] = temp[0]+' '+temp[1]
65             temp.pop(1)
66
67         temp[1] = ''.join(i for i in temp[1] if not i.isalpha())
68         temp[3] = ''.join(i for i in temp[3] if not i.isalpha())
69
70         ResultList.append(temp[:7])
71
72 ResultList = np.asarray(ResultList)
73 IonList = list([])
74
75 ###=====
76 #         Figure out how many different ions are present, list them.
77
78 for ion in ResultList[:, 0]:
79     if ion not in IonList:
80         IonList.append(ion)
81
82
83 ###=====
84 #         Now, pair ion and atom and calculate metallicities in terms
85 #         of Z_Solar.
86 #         Error propagation:
87 #         
$$\delta Z_{\text{ion}} = \sqrt{(\delta \log N_H)^2 + (\delta \sum \log(x_i))^2}$$


```

Appendix B. Source Code

```
88
89 Metalicity = list(['Ion ', 'Log N ', 'dlog N ', 'Zdla ', 'Z0 ', 'Z-Z0 ',
90                  'd(Z - Z0) '])
91 print ' '
92 print ' '
93 print ' '
94 print 'Metallicities: '
95 print IonList
96 print ResultList
97 for ion in IonList:
98     ThisIon = np.where(ResultList[:, 0] == ion)
99     LogN = ResultList[ThisIon][:, 5].astype(np.float)
100    dLogN = ResultList[ThisIon][:, 6].astype(np.float)
101    ColDens = 0.
102    counter = 0
103    dColDens = 0.
104
105    # The uNumpy way:
106    uLogN = unp.uarray((LogN, dLogN))
107    # print ion, uLogN.T
108    uNion = 10**uLogN
109    # print uNion
110    uNtot = uNion.sum()
111    # print uNtot
112    uLogNtot = umc.log10(uNtot)
113    # print 'With python uncertainties package:'
114    # print 'Log N, total, for ion '+ion+':\t', uLogNtot
115
116
117
118    for i in LogN:
119        dens = 10.**float(i)
120        # print 'dens=', dens
121        ColDens += dens
122        counter += 1
123    for i in np.arange(dLogN.shape[0]):
124        # Propagate thru exp10 function:
125        unc = float(dLogN[i])
126        # print unc
127        dNi = np.log(10.)*10.**float(LogN[i])*unc #Seems more like it
128        # Prepare for pyth summing:
129        dNi = dNi**2
130        dColDens += dNi
131
132    dColDens = np.sqrt(dColDens)
133    dLogColDens = 1./np.log(10.)*1./ColDens*dColDens
134
135    ###=====
136    # Find solar metallicities for the different ions:
```



```

137
138     for i in open('./SolarMetal.data'):
139         if ion.startswith(i.split()[0]):
140             if len(i.split()[0]) >= 2:
141                 # They match!
142                 Zsolar = i.split()[1]
143                 atom = i.split()[0]
144                 break
145             elif ' ' in ion:
146                 # They match!
147                 Zsolar = i.split()[1]
148                 atom = i.split()[0]
149                 break
150
151     M = np.log10(float(ColDens))-np.log10(float(NeuHyd))-\\
152         (float(Zsolar)-12.)
153     Zdla = np.log10(float(ColDens)/float(NeuHyd))
154     uZ = uLogNtot-umc.log10(uNeuHyd)
155     uZsolar = np.float(Zsolar)
156     Zsolar = str(float(Zsolar)-12.)
157     uZZ0 = uZ-(uncs.ufloat((uZsolar, 0.0))-uncs.ufloat((12., 0.)))
158     print ion+'\\t', "%.4f" % (uLogNtot.nominal_value), '+/-' ,\\
159         "%.4f" % (uLogNtot.std_dev()), '\\t', \\
160         "%.4f" % (uZ.nominal_value), '+/-' ,\\
161         "%.4f" % (uZ.std_dev()), '\\t', \\
162         "%.4f" % (uZZ0.nominal_value), '+/-' ,\\
163         "%.4f" % (uZZ0.std_dev())
164
165     # Error:
166     dM = np.sqrt(DelNeuHyd**2+dLogColDens**2)
167     Metallicity.append([ion, np.log10(ColDens), dLogColDens, Zdla,
168         Zsolar, M, dM])
169
170 print np.asarray(Metallicity)
171 print 'Value of log(n) for neutral H used: ', "%.5f" % np.log10(NeuHyd)
172 np.savetxt('Metallicities.txt', np.asarray(Metallicity), fmt='%10s', delimiter='\\t')

```

Appendix B. Source Code

BIBLIOGRAPHY

- [1] *X-Shooter Pipeline User Manual*, 2009.
- [2] Martin Asplund, Nicolas Grevesse, A. Jacques Sauval, and Pat Scott. The chemical composition of the sun. *ANNUAL REVIEWS OF ASTRONOMY AND ASTROPHYSICS*, 47:522, 2009.
- [3] Robert H. Becker, Xiaohui Fan, Richard L. White, Michael A. Strauss, Vijay K. Narayanan, Robert H. Lupton, James E. Gunn, James Annis, Neta A. Bahcall, J. Brinkmann, A. J. Connolly, Istvan Csabai, Paul C. Czarapata, Mamoru Doi, Timothy M. Heckman, G. S. Hennessy, Zeljko Ivezic, G. R. Knapp, Don Q. Lamb, Timothy A. McKay, Jeffrey A. Munn, Thomas Nash, Robert Nichol, Jeffrey R. Pier, Gordon T. Richards, Donald P. Schneider, Chris Stoughton, Alexander S. Szalay, Anirudda R. Thakar, and D. G. York. Evidence for reionization at $z \sim 6$: Detection of a Gunn-Peterson trough in a $z=6.28$ quasar. *Astronomical Journal*, 122:2850, 2001.
- [4] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolte, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results. *The Astrophysical Journal, Supplement Series*, 148:1–27, sep 2003.
- [5] R. J. Bouwens, G. D. Illingworth, I. Labbe, P. A. Oesch, M. Trenti, C. M. Carollo, P. G. van Dokkum, M. Franx, M. Stiavelli, V. González, D. Magee, and L. Bradley. A candidate redshift $z \sim 10$ galaxy and rapid changes in that population at an age of 500 myr. *Nature*, 469:504–507, jan 2011.
- [6] Jason A. Cardelli, Geoffrey C. Clayton, and Jon S. Mather. The relationship between infrared, optical and ultraviolet extinction. *The Astrophysical Journal*, (345):245–256, 1989.
- [7] R. F. Carswell, J. K. Webb, A. J. Cooke, and M. J. Irwin. *vpfit9.5*. <http://www.ast.cam.ac.uk/rfc/vpfit.html>, 9.5 edition, May 2010.

Bibliography

- [8] Sandro D’Odorico, Hans Dekker, Ruben Mazzoleni, Joel Vernet, Isabelle Guinouard, Paul Groot, Francois Hammer, Per K. Rasmussen, Lex Kaper, Ramon Navarro, Roberto Pallavicini, Celine Peroux, and Filippo M. Zerbi. X-shooter, uv to k band, intermediate resolution, high efficiency spectrograph for the vlt: status report at the final design review. *Proc. of SPIE*, Vol. 6269(626933-1), 2006.
- [9] A. S. Fruchter, A. J. Levan, L. Strolger, P. M. Vreeswijk, S. E. Thorsett, D. Bersier, I. Burud, J. M. Castro Cerón, A. J. Castro-Tirado, C. Conselice, T. Dahlen, H. C. Ferguson, J. P. U. Fynbo, P. M. Garnavich, R. A. Gibbons, J. Gorosabel, T. R. Gull, J. Hjorth, S. T. Holland, C. Kouveliotou, Z. Levay, M. Livio, M. R. Metzger, P. E. Nugent, L. Petro, E. Pian, J. E. Rhoads, A. G. Riess, K. C. Sahu, A. Smette, N. R. Tanvir, R. A. M. J. Wijers, and S. E. Woosley. Long γ -ray bursts and core-collapse supernovae have different environments. *Nature*, 441:463–468, may 2006.
- [10] J. P. U. Fynbo, P. Laursen, C. Ledoux, A. K. Durgapal, P. Møller, P. Goldoni, B. Gullberg, L. Kaper, J. Maund, P. Noterdaeme, G. Ostlin, M. L. Strandet, S. Toft, P. M. Vreeswijk, and T. Zafar. Galaxy counterparts of metal-rich damped lyman-alpha absorbers - i: The case of the $z=2.35$ dla towards q2222-0946. *Monthly Notices of the Royal Astronomical Society*, (408):2128–2138, 2009.
- [11] J. P. U. Fynbo, C. Ledoux, P. Noterdaeme, L. Christensen, P. Møller, A. K. Durgapal, P. Goldoni, L. Kaper, J. K. Krogager, P. Laursen, J. R. Maund, B. Milvang-Jensen, K. Okoshi, P. K. Rasmussen, T. J. Thorsen, S. Toft, and T. Zafar. Galaxy counterparts of metal-rich damped lyman-alpha absorbers - ii. a solar-metallicity, molecular-rich, and dusty dla at $z_{abs} = 2.58$, 2010.
- [12] J. P. U. Fynbo, J. X. Prochaska, J. Sommer-Larsen, M. Dessauges-Zavadsky, and P. Møller. Reconciling the metallicity distributions of gamma-ray burst, damped $Ly\alpha$, and lyman break galaxies at $z \sim 3$. *The Astrophysical Journal*, 683:321–328, aug 2008.
- [13] Johan Fynbo. *Pieces to the puzzle of galaxy formation*. PhD thesis, Institute of Physics and Astronomy, University of Århus, 2000.
- [14] J. E. Gunn and B. A. Peterson. On the density of neutral hydrogen in intergalactic space. *The Astrophysical Journal*, 142:1633–1641, nov 1965.
- [15] R. W. Hanuschik. A flux-calibrated, high-resolution atlas of optical sky emission from uves. *A & A*, 407(3):1157 – 1164, 2003.

- [16] M. Horrobin, P. Goldoni, F. Royer, P. François, G. Blanc, J. Vernet, A. Modigliani, and J. Larsen. Error propagation in the x-shooter pipeline. In Andreas Kaufer and Florian Kerber, editors, *The 2007 ESO Instrument Calibration Workshop*, volume 47 of *ESO Astrophysics Symposia*, pages 221–223. Springer Berlin / Heidelberg, 2008.
- [17] J. C. Howk and K. R. Sembach. Ionized gas in damped lyalpha systems and its effects on elemental abundance studies. *The Astrophysical Journal*, 523:L141–L144, oct 1999.
- [18] L. S. Sparke J. S. Gallagher. *Galaxies in the Universe*. Cambridge University Press, 2000.
- [19] Willam C. Keel. Quasars as lighthouses: the lyman-alpha forest at low and high redshift. <http://www.astr.ua.edu/keel/agn/forest.html>. Colors changed from original.
- [20] K. Kohler and N. Y. Gnedin. Lyman limit systems in cosmological simulations. *The Astrophysical Journal*, 655:685–690, feb 2007.
- [21] Peter Laursen. *Interpreting Lyman α radiation from young, dusty galaxies*. PhD thesis, DARK Cosmology Center, 2010.
- [22] C. Ledoux, P. Petitjean, J. P. U. Fynbo, P. Møller, and R. Srianand. Velocity-metallicity correlation for high-z dla galaxies: evidence of a mass-metallicity relation? *Astronomy and Astrophysics*, 457(1):71, 2006.
- [23] Jochen Liske. Vpguess 1.5. <http://www.eso.org/~jliske/vpguess/index.html>. Software + manual.
- [24] J.D. Meiring, T. M. Tripp, J. X. Prochaska, J. Tumlinson, J. Werk, E. B. Jenkins, C. Thom, J.M. O’Meara, and K. R. Sembach. The first observations of low redshift damped lyman- α systems with the cosmic origins spectrograph.
- [25] D. M. Meyer and K. C. Roth. Observations of nickel, chromium, and zinc in qso absorption-line systems. *The Astrophysical Journal*, 363:57–65, nov 1990.
- [26] P. Möller, J. P. U. Fynbo, and S. M. Fall. Detection of lyman- α emission from a dla galaxy: Possible implications for a luminosity-metallicity relation at $z = 2-3$. *Astronomy and Astrophysics*, 422:–33, jul 2004.
- [27] J. E. O’Donnell. r_v -dependent optical and near-ultraviolet extinction. *The Astrophysical Journal*, 422:158–163, feb 1994.

Bibliography

- [28] Yichuan C. Pei. Interstellar dust from the milky way to the magellanic clouds. *The Astrophysical Journal*, pages 130–139, 1992.
- [29] Patrick Petitjean. *QSO Absorption Line Systems*. 1998.
- [30] Jason X. Prochaska, Hsiao-Wen Chen, Miroslava Dessauges-Zavadsky, and Joshua S. Bloom. Probing the ism near star forming regions with grb afterglow spectroscopy: Gas, metals, and dust. *The Astrophysical Journal*, (666):267–280, 2007.
- [31] Barbara Ryden. *Introduction to Cosmology*. Addison Wesley, 2003.
- [32] John R. Taylor. *An introduction to error analysis*. University Science Books, second edition, 1997.
- [33] Randal C. Telfer, Wei Zheng, Gerard A. Kriss, and Arthur F. Davidsen. The rest-frame extreme-ultraviolet spectral properties of quasi-stellar objects. *The Astrophysical Journal*, 565(2):773, 2002.
- [34] Thorsten Tepper Garcia. Voigt profile fitting to quasar absorption lines: an analytic approximation to the voigt-hjerting function. *Monthly Notices of the Royal Astronomical Society*, 369(4):2025, 2006.
- [35] Christina Thöne. *Gamma-Ray Burst host Galaxies at high resolution*. PhD thesis, DARK Cosmology Center, 2008.
- [36] Thøger Juul Thorsen. Search for $z \sim 2$ galaxies in the hubble deep field south. B.Sc. project, 2006.
- [37] C. A. Tremonti, T. M. Heckman, G. Kauffmann, J. Brinchmann, S. Charlot, S. D. M. White, M. Seibert, E. W. Peng, D. J. Schlegel, A. Uomoto, M. Fukugita, and J. Brinkmann. The origin of the mass-metallicity relation: Insights from 53,000 star-forming galaxies in the sloan digital sky survey. *The Astrophysical Journal*, 613:898–913, oct 2004.
- [38] The Free Encyclopedia Wikipedia. Voigt profile. Web.
- [39] A. M. Wolfe. New evidence from the lyman-alpha forest concerning the formation of galaxies. *Royal Society of London Philosophical Transactions Series A*, 320:503–515, dec 1986.