UNIVERSITÀ DEGLI STUDI DI TRIESTE
Sede Amministrativa del Dottorato di Ricerca

XX CICLO DEL
DOTTORATO DI RICERCA IN FISICA

Walks in the Lyman
and Metal-line Forests

Settore scientifico-disciplinare
FIS/05 ASTRONOMIA E ASTROFISICA

DOTTORANDO:
FRANCESCO SAITTA

COORDINATORE DEL COLLEGIO DEI DOCENTI:
PROF. GAETANO SENATORE
Dipartimento di Fisica Teorica Università di Trieste

TUTORI:
PROF. STEFANO CRISTIANI
INAF - Osservatorio Astronomico di Trieste

DOTT. PIERLUIGI MONACO
Dipartimento di Astronomia Università di Trieste

CO-TUTORI:
PROF. SANDRO D’ODORICO
European Southern Observatory, Garching bei München

DOTT. VALENTINA D’ODORICO
INAF - Osservatorio Astronomico di Trieste
Contents

0 Abstract ......................................................... 1

1 The physics of the Inter Galactic Medium .................. 6
  1.1 Cosmological Context .................................... 7
    1.1.1 Basic Equations ................................... 7
      1.1.1.1 Power Spectrum ................................ 9
    1.1.2 The Standard Cosmological Model ................... 10
      1.1.2.1 The Cosmological Constant .................... 11
      1.1.2.2 Cosmological parameters ....................... 12
    1.1.3 Structure Formation ................................ 13
      1.1.3.1 Linear regime .................................. 13
      1.1.3.2 Non-linear regime: Dark Matter ................. 14
      1.1.3.3 Non-linear regime: Baryonic gas ............... 15
  1.2 Gunn-Peterson test and the Ionization of the IGM ....... 16
    1.2.1 Constraining Reionization with QSO absorption spectra 18
  1.3 The UV background from the Proximity Effect ............. 19
  1.4 Absorption Line Definition ................................ 23
    1.4.1 The curve of growth ................................ 25
1.5 Lyα forest evolution as a function of redshift ......................................... 27
  1.5.1 Line number density evolution ......................................................... 28
  1.5.2 Evolution of the HI mean opacity ..................................................... 30
1.6 The Differential Density Distribution Function ........................................... 30
1.7 The Width of the Lines and the Temperature of the IGM ............................. 33
1.8 Two point correlation function of the Lyα lines ........................................ 34
1.9 Flux Statistics of the Lyα forest ............................................................ 36

2 New Sample, Old Tools ........................................................................... 38
  2.1 Simulated Data Sample ......................................................................... 38
  2.2 Observed Data Sample ................................................................. 41
    2.2.1 Creation of the line lists ............................................................. 42
  2.3 Line Number density evolution ............................................................ 43
  2.4 TPCF of the Lyα lines ........................................................................ 44

3 Overcoming the line counts: the FLO approach ........................................... 49
  3.1 Introducing FLO ............................................................................... 50
  3.2 Testing FLO with Cosmological Simulations .......................................... 52
    3.2.1 Reconstruction of the δ field ....................................................... 52
    3.2.2 Two point statistics of the δ field ............................................... 54
  3.3 FLO Applied to the observed data sample ............................................ 56
    3.3.1 Tracing the gas at redshifts ∼ 1.7-3.5 ........................................... 56
      3.3.1.1 Two point statistics of the δ field ....................................... 58
    3.3.2 The distribution of matter close to QSO from the proximity effect .... 59
3.3.3 QSO emission redshift determination .................................. 60
  3.3.3.1 QSO luminosity determination ..................................... 60
  3.3.3.2 The proximity effect .................................................. 63
  3.3.3.3 The density structure around QSOs ................................. 64
3.4 Future Perspectives .......................................................... 64

4 Tomography of the IGM ......................................................... 66
  4.1 Data Sample ........................................................................ 66
    4.1.1 The simulated spectra ................................................... 71
  4.2 The flux auto-correlation function ........................................ 71
    4.2.1 Contribution of the stronger absorptions ......................... 73
    4.2.2 Selection effects in the observed spectra ......................... 73
  4.3 Clustering of the flux across the line of sight ....................... 74
    4.3.1 The cross-correlation function in redshift space ............... 75
    4.3.2 The cross-correlation coefficient .................................... 77
  4.4 Discussion .......................................................................... 81
    4.4.1 Exploring the parameter space of the simulations .............. 82
      4.4.1.1 Effective optical depth ............................................ 82
      4.4.1.2 Temperature and $\gamma$ ........................................... 84
  4.5 Future Perspectives .......................................................... 85

5 Metal Enrichment of the IGM ................................................. 88
  5.1 QSO absorption lines as probes of galactic winds ................... 89
  5.2 Observations, Data Reduction, Preliminary Results and Work in Progress ........................... 90
5.2.1 New Observations ........................................... 91
5.3 Future Perspectives ......................................... 92

6 X-shooter .......................................................... 94

6.1 The instrument and the IGM Science Case ............... 94

6.2 Laboratory Measurements of Wavelength Calibration Sources in the near IR 98

6.2.1 Wavelength Calibration during X-shooter operations ... 98

6.2.2 Wavelength Calibration Sources for X-shooter .......... 99

6.2.2.1 Pen ray lamps ....................................... 99

6.2.2.2 Th-Ar Hollow cathode lamps ......................... 100

6.2.2.3 Experimental Setup .................................. 100

6.2.2.4 Measurements ....................................... 101

6.2.3 Physical Modeling and Simulated Data ................. 103

6.2.3.1 Physical Model ..................................... 103

6.2.3.2 Simulated 2D Echellograms ......................... 106

6.2.3.3 Extracted 1D spectra ............................... 108

6.2.4 Analysis and Results .................................... 109

6.2.4.1 Line Identifications ................................. 109

6.2.4.2 Order-by-order analysis ............................ 109

6.2.4.3 Summary of two lamp solutions ..................... 110

6.2.4.4 Summary of three lamp solutions ................. 113

6.2.4.5 Relative Intensities ................................. 113

6.2.4.6 Operation Current .................................. 113

6.2.4.7 Th-Ar .............................................. 114
6.2.4.8 Results ............................................. 116
6.2.4.9 Remaining Issues .................................. 117
6.2.4.10 Lessons Learned ................................... 118

6.3 Building up a database of spectro-photometric standard stars from UV to NIR ........................................... 118

6.3.1 Choice of the Instrument ................................ 119
6.3.2 Observational Strategy .................................. 119
6.3.3 Status of the Project ................................... 122
  6.3.3.1 Efficiency of SINFONI ......................... 123
  6.3.3.2 Photometry ........................................ 123

7 Conclusions .................................................. 126

A APPENDIX A: An Order by order analysis ................. 132
List of Tables

2.1 Observed Data Sample .................................................. 40

2.2 Detailed budget of the number of lines used to compute TPCFs. The first column refers to the selection done, in terms of emission redshift of the considered objects. .................................................. 48

3.1 Relevant properties of the QSOs forming our sample. See text for further details. .................................................. 62

4.1 Characteristics of the observed QSO spectra for the tomography of the IGM  68

4.2 ESO Large Program QSOs used as reference sample in the study of the tomography of the IGM .................................................. 69

4.3 Effective optical depth to build mock spectra .................................................. 82

5.1 Observed QSOs characteristics .................................................. 90

6.1 X-shooter characteristics .................................................. 95

6.2 Th-Ar lamp current and exposure times .................................................. 103

6.3 Pen ray lamps observations .................................................. 105

6.4 Usage time of the lamps .................................................. 105

6.5 Summary of the number of lines observed on the FTS spectra .................................................. 109

6.6 Requirement and goals for number of calibration lines provided .................................................. 110

6.7 Summary of the number of lines provided by each pen ray lamp for a given order .................................................. 111
6.8 Summary of the lines provided by a combination of Ne and Ar pen ray lamp for a given order .................................................. 111

6.9 Summary of the lines provided by a combination of Ne and Kr pen ray lamp for a given order .................................................. 112

6.10 Summary of the lines provided by a combination of Ar and Kr pen ray lamp for a given order .................................................. 112

6.11 Summary of the lines provided by a combination of Ne, Ar and Kr pen ray lamp, the best three lamp solution. ............................. 113

6.12 Relative overall intensities of the lamps based on the integrated line flux . 114

6.13 List of the standard stars to be observed ............................................. 122

6.14 Atmospheric transmission windows used to compute photometric points . 123

7.1 Codex characteristics ................................................................. 130

7.2 ESPRESSO characteristics .......................................................... 131

A.1 Order #11 ............................................................................. 133

A.2 Order #12 ............................................................................. 134

A.3 Order #13 ............................................................................. 135

A.4 Order #14 ............................................................................. 136

A.5 Order #15 ............................................................................. 137

A.6 Order #16 ............................................................................. 138

A.7 Order #17 ............................................................................. 139

A.8 Order #18 ............................................................................. 140

A.9 Order #19 ............................................................................. 141

A.10 Order #20 ........................................................................... 142

A.11 Order #21 ........................................................................... 143

A.12 Order #22 ........................................................................... 144

viii
List of Figures

1.1 Gunn-Peterson test ................................................. 17
1.2 Gunn-Peterson optical depth ...................................... 20
1.3 UV background, QSOs & galaxies ............................... 22
1.4 Curve of growth ................................................... 26
1.5 Redshift evolution of the Lyα forest ............................ 27
1.6 Lyα Lines number density evolution ............................ 28
1.7 UVB contribution to $dN/dz$ ..................................... 29
1.8 H I mean opacity evolution with redshift ....................... 31
1.9 H I column density distribution ................................. 32
1.10 Two Point Correlation Function (TPCF) of Lyα lines .......... 34
1.11 TPCF of Lyα lines with $\log N(HI) \geq 13.8$ .................. 35
2.1 Lyα forest redshift coverage of the considered QSO sample. 40
2.2 Portion of the Lyα forest for an observed (lower panel) and a simulated (upper panel) line of sight in our sample. ............... 42
2.3 Number density evolution of the Lyα forest lines, $13.64 < \log N(HI) < 17$ cm$^{-2}$ .................................................. 43
2.4 Number density evolution of the Lyα forest lines, $13 < \log N(HI) < 14$ cm$^{-2}$, $14.5 < \log N(HI) < 17$ cm$^{-2}$ ....................... 44
2.5 TPCF for the observed Lyα lines, $12 < \log N(HI) < 17$ cm$^{-2}$ ........... 45
4.9 Auto & cross-correlation function comparison ....................... 81
4.10 Auto-correlation function \& \( \tau_{eff} \) ................................. 83
4.11 Cross-correlation function \( \tau_{eff} \) .................................. 83
4.12 Cross-correlation function \& \( T_{IGM} \) ................................. 84
4.13 Distribution of pairs of QSOs in the observable universe .......... 86
5.1 Si IV absorption in the observed spectra ............................... 91
5.2 C IV absorptions in the observed spectra ............................. 92
6.1 X-shooter simulated spectrum ........................................... 96
6.2 Expected SNR for the UV arm of Xshooter ............................ 97
6.3 Fourier Transform Spectrometer (FTS) ................................. 101
6.4 Th-Ar lamp ............................................................. 102
6.5 Th-Ar spectrum .......................................................... 102
6.6 Pen ray lamps ........................................................... 103
6.7 Pen ray lamps spectra .................................................... 104
6.8 Simulated Pen ray lamps 2D spectra .................................. 107
6.9 Simulated Th-Ar lamp 2D spectra ...................................... 107
6.10 Blaze function .......................................................... 108
6.11 Simulated 1D spectrum of X-shooter, Order #20 .................... 108
6.12 Th-Ar lamp: Intensity vs current plot ................................ 114
6.13 Pen ray lamps: Intensity vs current plot .............................. 115
6.14 Th-Ar spectrum at X-shooter resolution ............................... 116
6.15 Telluric absorption in the J band ...................................... 120
6.16 Telluric absorption in the H+K band .................................. 120
6.17 Status of the standard star project ........................................... 122
6.18 SINFONI efficiency ................................................................. 124
6.19 Spectrum of Hiltner600 .......................................................... 124
6.20 Flux measurements for Hiltner600 ........................................... 125
A.1 Order #11 ................................................................. 133
A.2 Order #12 ................................................................. 134
A.3 Order #13 ................................................................. 135
A.4 Order #14 ................................................................. 136
A.5 Order #15 ................................................................. 137
A.6 Order #16 ................................................................. 138
A.7 Order #17 ................................................................. 139
A.8 Order #18 ................................................................. 140
A.9 Order #19 ................................................................. 141
A.10 Order #20 ............................................................... 142
A.11 Order #21 ............................................................... 143
A.12 Order #22 ............................................................... 144
A.13 Order #23 ............................................................... 145
A.14 Order #24 ............................................................... 146
A.15 Order #25 ............................................................... 147
A.16 Order #26 ............................................................... 148
List of acronyms

AC       Alternate Current
AIT      Assembly, Integration and Testing
BB       Big Bang
CDM      Cold Dark Matter
CMX      Cosmic Microwave Background
CODEX    COsmic Dynamics EXperiment
DIT      Detector Integration Time
DM       Dark Matter
ELT      Extremely Large Telescope
ESO      European Southern Observatory
FF       Flat Field
FFT      Fast Fourier Transform
FLO      From Lines to Overdensities
FoF      Friend-of-Friend
FRW      Friedmann-Robertson-Walker
FTS      Fourier Transform Spectrometer
GR       General Relativity
HCL      Hollow Cathod Lamp
IFU      Integral Field Unit
IGM      Inter-Galactic Medium
ISM      Inter-Stellar Medium
LJ       Jeans Length
LOS      Line of sight
LP       Large Programme
LSS      Large Scale Structure
Lyα      Lymanα
NDIT     Number of Integrations Time
NIR      Near-Infrared
NIST     National Institute of Standards and Technology
Prims    Primary Stars
PSU      Power Supply Unit
QSO      Quasi Stellar Object
QTH      Quartz Tungsten Halogen
RA       Right Ascension
Secs     Secondary Stars
SNR      Signal to Noise Ratio
TPCF     Two Point Correlation Function
UVB      UV cosmic background
VLT      Very Large Telescope
WD       White Dwarf
WMAP     Wilkinson Microwave Anisotropy Probe
ΛCDM     Λ-Cold Dark Matter
The aim of this PhD thesis is the study of the properties of the Inter-Galactic Medium (IGM) through Quasi Stellar Objects (QSOs) absorption lines at redshifts $z \sim 2-4$. The study of the IGM provides a unique picture of the Universe in the early phases of structure formation and is crucial to test cosmological models. Furthermore, the metal content of the IGM can be used to constrain the nature of the enrichment mechanisms at high redshifts. Two main mechanisms have been proposed predicting different distribution of metals in the IGM: a late enrichment, due to the observed galaxies at redshift $\sim 3$, and an early enrichment, due to a population of very massive first stars (the so called popIII stars), not yet observed, at higher redshift. The existing observations are not sufficient to discriminate between the two mechanisms. QSO spectra show interesting features to study the IGM:

- The Lyman $\alpha$ (Ly$\alpha$) forest: the forest of absorption lines blue-ward of the Ly$\alpha$ emission arising in large scale neutral hydrogen density fluctuations of moderate amplitude in the warm photo-ionized IGM. Since the dynamical state of the low density IGM is governed mainly by the Hubble expansion, gravitational instabilities and photo-ionization, the involved physics is quite simple and mildly non-linear. The Lyman forest is then a fair tracer of the underlying matter density field.

- Metal absorptions: many absorption features due to ionic transitions in chemical elements heavier than He ("metals") are clearly present red-ward of the Ly$\alpha$ emission. Some of them are associated with the QSO itself but the majority are tracers of intervening metals belonging to structures of different nature, from diffuse gas to galaxies.

Chapter 1 of the thesis presents in detail how the physics of the IGM can be understood studying QSO absorption lines, and the classical results obtained in this field. Then the thesis focuses on the work done during the PhD, which has addressed three research areas:

1. Study of lines of sight (LOSs) to isolated QSOs (Chapter 2,3):
a sample of 22 high-resolution QSO spectra has been studied. Classical statistics have been applied: all the absorption features in the spectra have been fit with Voigt profiles to extract physical parameters associated with the absorbing material; quantities like the evolution in redshift of the number density of the lines and the two-point correlation function of the line distribution along the spectra have been studied and compared with measurements available in the literature;

- a new method to analyze the Lyα forest has been implemented. Traditionally, absorption spectra are resolved in a collection of discrete absorption systems. This method instead reconstructs the underlying density field processing the lines on the basis of the physical properties of the IGM. Therefore, a continuous density field is built and the main drawbacks of the Voigt fitting approach are overcome.

- the new algorithm has been tested with N-body hydrodynamical simulations of the IGM;

- the new algorithm has been applied to the observational data sample to study the hydrogen density field (its evolution with redshift and its clustering properties) and the so-called proximity effect of QSO, estimating the overdensity around the object.

2. Study of lines of sight to multiple QSOs (Chapter 4,5):

The thesis presents also the work done studying multiple QSO LOSs, i.e. studying the properties of the IGM not only with the spectrum of a single object, by trying to get transversal informations comparing different LOSs, close both in angular separation and in the emission redshift of the source. A sample of 15 QSO forming 21 pairs have been studied to get the transverse correlation function, using the statistics of the transmitted flux of the objects. Furthermore, we have observed a pair of close QSO during two nights (7-8 August 2007) with UVES, the high resolution spectrograph at the ESO Very Large Telescope (VLT) located at Cerro Paranal, in Chile. These two objects have been observed to study the correspondence between metal absorptions and galaxies in the field; preliminary results and the description of the work in progress are presented in detail in the thesis (Chapter 5).

3. New instrumentation (Chapter 6):

Looking ahead to new possibilities of advance in this field of research, an important role will be played by the high sensitivity and medium resolution spectrograph X-shooter. It will receive first light at the ESO VLT in July 2008 and will start operating in early 2009. When in operation, its wide spectral-range observing capability will be unique at very large telescopes and extremely relevant for the study of QSO spectra. The thesis describes the Science Case "Tomography of the IGM" for the instrument, and the work done at ESO within the X-shooter project. For the operation of this instrument we have carried out laboratory measurements of calibration sources for the Near-InfraRed arm (1–2.5 μm) and I participated to an observational project to build a spectro-photometric flux catalogue of standard stars for the instrument.
Riassunto

Scopo di questa tesi di dottorato è lo studio delle proprietà del mezzo intergalattico (IGM) attraverso le righe in assorbimento negli spettri di quasar (QSO) osservate a redshift $z \sim 2-4$. Lo studio dell’IGM è di fondamentale importanza per verificare modelli cosmologici in quanto rende possibile l’osservazione dell’universo nelle prime fasi del processo di formazione delle strutture cosmiche. Inoltre, le tracce di elementi chimici più pesanti dell’He (detti “metalli”) contenuti nell’IGM possono essere utilizzate per comprendere i meccanismi di arricchimento metallico dell’universo ad alti redshift. La distribuzione osservata di questi metalli è al momento oggetto di un intenso dibattito all’interno della comunità scientifica; recentemente sono stati proposti due scenari per spiegare questa distribuzione: il cosiddetto di “late enrichment”, per cui i metalli osservati nell’IGM sarebbero dovuti principalmente alle galassie osservate a redshift $\sim 3$, e l’“early enrichment”, per cui una popolazione di stelle massicce (“popIII stars”, non ancora osservate) a redshift ancora maggiore, sarebbe responsabile della maggior parte di metalli presenti nell’IGM. Al momento le osservazioni non sono sufficienti per falsificare o confermare i meccanismi proposti. L’IGM viene studiato attraverso l’analisi di spettri di QSOs ed in particolare:

- La foresta Lyman$\alpha$ (Ly$\alpha$): l’insieme delle righe in assorbimento che si osservano a lunghezze d’onda minori rispetto all’emissione Ly$\alpha$ negli spettri di QSOs; questi assorbimenti sono dovuti a piccole fluttuazioni su larga scala del campo di idrogeno neutro presente nell’IGM. La fisica che descrive l’IGM è relativamente semplice. Il gas si trova in regime lineare o moderatamente non-lineare ed in equilibrio di fotoionizzazione. La sua dinamica, alla densità media, è dominata dall’espansione di Hubble e dall’instabilità gravitazionale, perciò la foresta Ly$\alpha$ risulta essere un buon candidato per la descrizione del campo di densità presente a questi redshift.

- Gli assorbimenti metallici: a lunghezze d’onda maggiori dell’emissione Ly$\alpha$ sono chiaramente presenti molti assorbimenti dovuti a transizioni di ioni metallici. Alcuni di essi hanno origine in gas associato al QSO stesso, ma la maggior parte di questi assorbimenti riflette la presenza di sistemi metallici lungo la linea di vista tra l’oggetto e l’osservatore. Questi sistemi non associati sembrano essere dovuti ad una grande varietà di strutture, da componenti di gas diffuso fino alle galassie di campo.

Il capitolo 1 della presente tesi descrive nel dettaglio come si possa comprendere la fisica dell’IGM studiando i sistemi in assorbimento degli spettri di QSO, presentando i principali risultati ottenuti tradizionalmente in questo campo. Sucessivamente la tesi si concentra...
sul lavoro svolto durante il periodo di dottorato, ed in particolare sulle tre principali aree di ricerca affrontate:

1. Studio di linee di vista (LOSs) verso QSO singoli (Capitoli 2, 3):
   - è stato analizzato un campione di 22 spettri di QSO ad alta risoluzione, al quale sono state applicate le statistiche tradizionali: tutti gli assorbimenti Lyα negli spettri sono stati fittati con profili di Voigt, al fine di ottenere parametri fisici associati ai sistemi di assorbimento; in particolare sono state studiate, e confrontate con risultati ottenuti in passato, quantità come l’evoluzione in redshift della densità in numero di righe e la funzione di correlazione a due punti;
   - è stato sviluppato un nuovo metodo per analizzare la foresta Lyα. A differenza dell’approccio tradizionale, dove da uno spettro di QSO si ottiene semplicemente un insieme discreto di assorbitori, questo metodo ricostruisce il campo di densità di idrogeno responsabile degli assorbimenti partendo da assunzioni sulla fisica dell’IGM. Il risultato è un campo di densità continuo la cui analisi permette di risolvere alcuni problemi introdotti invece dalla classica analisi di profili di Voigt.
   - il nuovo algoritmo è stato testato con simulazioni idrodinamiche “N-body” dell’IGM;
   - il nuovo algoritmo è stato applicato ai dati osservativi citati in precedenza per studiare il campo di densità di idrogeno (la sua evoluzione con il redshift e la funzione di correlazione a due punti) ed il cosiddetto “effetto di prossimità”, portando ad una stima della sovradensità del campo nella regione dove si trova un QSO.

2. Studio di linee di vista verso sistemi di QSOs (Capitoli 4, 5):
   la tesi presenta anche uno studio di linee di vista multiple; il confronto di diverse linee di vista, vicine in separazione angolare e in redshift di emissione, permette di avere informazioni non solo sulle regioni di universo tracciate dalle linee di vista, ma anche sulle regioni tra le linee di vista stesse. A questo riguardo è stata studiata la “funzione di correlazione trasversale” ottenuta dall’analisi del flusso trasmesso da un campione di 15 QSO che formano 21 coppie. Per lo studio di linee di vista multiple è stata osservata una coppia di QSO per due notti (7-8 agosto 2007); le osservazioni sono state fatte con lo spettrografo ad alta risoluzione UVES, al “Very Large Telescope” (VLT), situato al Cerro Paranal, in Cile. Questi oggetti sono stati osservati per studiare le corrispondenze tra assorbimenti metallici e galassie di campo. Il lavoro fatto ed i risultati preliminari ottenuti fino ad ora sono presentati nel capitolo 5.

3. Nuova strumentazione (Capitolo 6):
   Un ruolo importante per il progresso della scienza dell’IGM verrà svolto dal nuovo spettrografo a media risoluzione ed altà sensibilità, X-shooter. X-shooter riceverà la prima luce all’osservatorio ESO VLT in luglio 2008 per poter essere in pieno funzionamento agli inizi del 2009. L’ampio intervallo spettrale coperto e le caratteristiche di questo strumento sono particolarmente vantaggiose per lo studio di spettri di QSOs. Nella tesi viene descritto il caso scientifico “Tomography of the IGM” per lo strumento, ed il lavoro svolto all’ESO nel contesto del progetto X-shooter. Questo lavoro include le misure di laboratorio di sorgenti di calibrazione per il braccio nel
vicino infrarosso dello strumento ($1 - 2.5\mu m$) e la partecipazione al progetto osservativo per la costruzione di un catalogo spettro-fotometrico di stelle standard per la calibrazione assoluta in flusso degli oggetti che si osserveranno con X-shooter.
The study of the Inter Galactic Medium (IGM) has greatly evolved during recent years. The discovery of QSOs in the sixties opened this field of astronomy (Schmidt 1965). Looking at the absorption spectra of these objects people realized the presence of an extraordinary amount of absorption features, in particular a “forest” of absorption lines blue-ward of the Lyα emission, much more numerous than in the region red-ward of the Lyα emission. Most of the lines in the forest were proven to be due to the Lyα resonant transition in neutral hydrogen (Gunn & Peterson 1965; Bahcall & Salpeter 1965; Burbidge et al. 1966). With the acquisition of more and more data, the number of Lyα lines strongly supported the idea that galactic and intergalactic gas, and not only material intrinsic to the QSO, is the source of most quasar absorption (Lynds 1971; Oemler & Lynds 1975; Young et al. 1979). The first models for the Lyα forest tried to describe absorbers as pressure confined clouds, each of them responsible for a Lyα absorption line, confined by a hotter and more tenuous IGM (Sargent et al. 1980; Ikeuchi & Ostriker 1986; Baron et al. 1989). These models turned out to have significant problems in explaining the large variety of structures that the Lyα forest should trace, the evolution of the line number density with redshift, the mechanisms of formation of these clouds. Finally the advent of semi-analytic and hydro-dynamical simulations suggested a new framework for the forest of hydrogen absorption in QSO spectra. Most studies on the large-scale properties of the IGM have concentrated on the crucial problem of the physical scale of Lyα forest clouds. The large sizes found (e.g. Weymann & Foltz 1983; Foltz et al. 1984; Smette et al. 1992, 1995; Bechtold 1994; Dinshaw et al. 1994, 1995; Fang et al. 1996; Crotts & Fang 1998; Petitjean 1998; Monier et al. 1999; Lopez et al. 2000; D’Odorico et al. 1998, 2002; Williger et al. 2000; Young et al. 2001; Aracil et al. 2002; Becker et al. 2004) have led to the realization that these clouds are really part of the general large scale structure. Ionization arguments (Rauch & Haehnelt 1995), analytical and Monte Carlo modelling of absorption in double lines of sight (Smette et al. 1992; Charlton et al. 1995; Fang et al. 1996; Crotts & Fang 1998; Viel et al. 2002) and cosmological hydro-simulations (Cen et al. 1994; Petitjean et al. 1995; Zhang et al. 1995; Hernquist et al. 1996; Miralda-Escudé et al.
1.1. Cosmological Context

Nowadays there is a general consensus in the scientific community on the so called Standard Cosmological Model. The Cosmological Principle postulates the homogeneity and isotropy of the universe; in other words, there is not a privileged observer in the universe and its statistic properties at a certain time are independent of the position (see Milne 1935; Bondi & Gold 1948; Hoyle 1948). Both isotropy and homogeneity are confirmed by observations of the mass distribution on large scale (the Large Scale Structure, LSS Colless et al. 2001; Adelman-McCarthy et al. 2006) and of the Cosmic Microwave Background (CMB) (Spergel et al. 2007). On the other hand it is evident that the universe is neither isotropic nor homogeneous on galactic scales. This discrepancy is due to small primordial fluctuations superimposed on the otherwise homogeneous density field, which grow forming the currently observed galaxies. In this section we will focus our attention on some relevant concepts we will refer to in the next chapters. We also give some equations that we consider useful for a better comprehension of the text. A more rigorous treatment of cosmological equations can be found for instance in Padmanabhan (1995) and Peacock (1999).

1.1.1 Basic Equations

The evidence that the distance between galaxies is increasing with time leads to the idea that at a certain point in the past the observed universe should have been compressed in a small volume. This model is called Big Bang (BB). It is worth to notice that the idea of the BB as a “big explosion” is not correct: it is more realistic to figure out the BB as the starting instant of the expansion of the universe. The space-time geometry of a homogeneous, isotropic and expanding universe is described by the so called Robertson-Walker (RW) Metric:

\[ ds^2 = dt^2 - a^2(t) \left( \frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right). \]  

(1.1.1)
1.1. Cosmological Context

The $k$ parameter can assume the values -1,0,1 for a negatively curved, flat or positively curved universe, respectively; $ds^2$ represents the proper time describing the distance of two events; $t$ is the time coordinate, $r$ the radial space coordinate, $\theta$ and $\phi$ the angular coordinates. The quantity $a(t)$ is the Scale Factor: it describes how distances in the universe scale with time. We can write the following relation:

$$1 + z(t) = \frac{a(t_0)}{a(t)} ,$$

(1.1.2)

where we introduce the concept of cosmological redshift. The value $z = 0$ corresponds to the present time, $t = t_0$. The classical interpretation of redshift is related to the Doppler shift due to the velocity of the source with respect to the observer. In a very similar way we can associate a recession velocity $v$ to the cosmological redshift. Following our similitude we can write a relation between this “velocity” and the distance $d$ of the source:

$$d = \frac{\dot{a}(t)}{a(t)} v .$$

(1.1.3)

We define the Hubble parameter:

$$H(t) = \frac{\dot{a}(t)}{a(t)} ,$$

(1.1.4)

the Hubble constant:

$$H_0 = H(t_0) ,$$

(1.1.5)

$l$ as the system of physical coordinates and $x$ as the system of comoving coordinates:

$$l = ax ,$$

(1.1.6)

where $a \equiv (1 + z)^{-1}$ is the scale factor which describes the expansion of the universe. A comoving observer is the observer that will perceive the universe to be isotropic, while non-comoving observers will see region of the sky systematically blue-shifted or red-shifted.

To find the evolution of the scale factor the Einstein’s General Relativity (GR) is assumed to describe gravity. For the RW metric the Einstein’s field equations take the form of the Friedmann equations:

$$\left(\frac{\ddot{a}}{a}\right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2} + \frac{\Lambda}{3} ,$$

(1.1.7)

$$\frac{\dot{a}}{a} = -\frac{4\pi G}{3} (\rho + 3p) + \frac{\Lambda}{3} .$$

The parameter $k$ represents the curvature of the universe (eq. 1.1.1); $\rho$ is the mean density of the universe an $p$ is the pressure. When wecompute density and pressure we have to take into account both the radiation and the matter content of the universe. In the cosmological framework matter and raditation evolve in a different way, in particular matter density scales as $a(t)^{-3}$, while the energy density of radiation as $a(t)^{-4}$. For this reason we consider three different epochs in the history of the universe dominated respectively by matter, radiation and vacuum. The epoch during which the densities of matter and radiation where almost equal is called the equivalence epoch. The quantity $\Lambda$ is called cosmological constant refers to the vacuum component of the universe. The nature of this term represents one of the most interesting and open questions in cosmology, it will be
1.1. Cosmological Context

discussed later in the thesis (sec. 1.1.2.1). The curvature of the universe is an important parameter to understand the topology of the space-time, the evolution of the density of different species, and the final fate of the universe. A flat universe is described by an Euclidean space; in absence of a vacuum component, a flat universe expands forever but at a continually decelerating rate, with expansion asymptotically approaching a fixed rate. With a vacuum component, the expansion rate of the universe initially slows down, due to the effect of gravity, but eventually increases. A negatively curved universe is described by spherical geometry; in a negatively curved universe without a vacuum component, gravity eventually stops the expansion of the universe, after which it starts to contract until all matter in the universe collapses to a point, a final singularity termed the “Big Crunch”, by analogy with Big Bang. However, if the universe has a large amount of dark energy, then the expansion of the universe can continue forever. An positively curved universe is described by hyperbolic geometry, and will continue to expand forever. In this thesis the universe is assumed to be flat, based in the recent results by Spergel et al. (2007). The density of the universe is directly related to its geometry through the parameter \( \Omega \):

\[
\Omega = \frac{\rho}{\bar{\rho}_{cr}},
\]

where \( \bar{\rho}_{cr} \) is the Critical Density of the universe:

\[
\bar{\rho}_{cr}(t) = \frac{3H^2(t)}{8\pi G}.
\]

The value of the parameter \( \bar{\rho}_{cr} \) is related to the geometry of the universe: if \( \Omega \) is greater than 1 we live in a close universe; if \( \Omega \) is lower than 1 we live in an open universe; if \( \Omega \) is equal to 1 our universe is flat. \( \Omega \) can be expressed as the sum of the densities of all the species contributing to the energy budget of the universe:

\[
\Omega = \Omega_b + \Omega_{DM} + \Omega_R + \Omega_\nu + \Omega_\Lambda
\]

where \( \Omega_b \) refers to the baryonic density, \( \Omega_{DM} \) to the dark matter (DM) density, \( \Omega_r \) to the radiation density, \( \Omega_\nu \) to the neutrinos density, and \( \Omega_\Lambda \) to the vacuum density. The Hubble parameter can be written as a function of \( a \) and \( \Omega \):

\[
H(a) = H_0 E(a) = H_0 \sqrt{\Omega_m \left( \frac{a_0}{a} \right)^3 + \Omega_R \left( \frac{a_0}{a} \right)^4 + \Omega_\Lambda},
\]

\( \Omega_m \) being the matter component of the universe, \( \Omega_m = \Omega_b + \Omega_{DM} \).

1.1.1.1 Power Spectrum

This paragraph describes the properties of the matter density field in comoving coordinates. First of all we define the density contrast as:

\[
\delta(x) = \frac{\rho(x) - \bar{\rho}}{\bar{\rho}}.
\]

This quantity represents the deviation of the density field from its mean value \( \bar{\rho} \). We can represent the spatial distribution of the density contrast as a superposition of plane waves.
or a Fourier series. We consider the comoving wavenumber $k = (n_x \frac{2\pi}{L}, n_y \frac{2\pi}{L}, n_z \frac{2\pi}{L})$ and impose periodic boundary conditions on a cubic volume $V = L^3$:

$$\delta(x) = \sum_{n_x, n_y, n_z = -\infty}^{\infty} \delta_k e^{i k \cdot x}, \quad (1.1.13)$$

where $\delta_k$ is the Fourier transform of $\delta(x)$. If we take an arbitrarily large $L$ we can recover the integral form of the Fourier transform:

$$\delta(x) = \frac{1}{(2\pi)^3} \int \tilde{\delta}_k e^{i k \cdot x} d^3 k \quad (1.1.14)$$

Assuming $\delta(x)$ is a real quantity implies $\delta_k^* = \delta_{-k}$. The variance of the field is defined as:

$$\sigma^2 = \langle (\delta(x) - \langle \delta(x) \rangle)^2 \rangle = \langle \delta(x)^2 \rangle; \quad (1.1.16)$$

if we consider two vectors $x_1$ and $x_2$, the modulus of their difference being $x = |x_1 - x_2|$, we define the correlation function $\xi(x)$ as:

$$\xi(x) = \langle \delta(x_1) \delta(x_2) \rangle. \quad (1.1.17)$$

If we compare eq. 1.1.17 and eq. 1.1.18 we obtain:

$$\xi(x) = \langle \int \tilde{\delta}_{k_1} e^{i k_1 \cdot x_1} d^3 k_1 \int \tilde{\delta}_{k_2} e^{i k_2 \cdot x_2} d^3 k_2 \rangle$$

$$= \int d^3 k_1 \int d^3 k_2 e^{i k_1 \cdot x_1 + i k_2 \cdot x_2} \langle \delta_{k_1}^* \delta_{k_2} \rangle. \quad (1.1.18)$$

We define the power spectrum:

$$\langle \delta_{k_1}^* \delta_{k_2} \rangle = P(|k_1 - k_2|) \delta^D(k_1 - k_2), \quad (1.1.19)$$

where $\delta^D$ represents a Dirac delta function. Comparing eq. 1.1.17, 1.1.18 and 1.1.19 we get:

$$\xi(x) = \frac{1}{(2\pi)^2} \int d^3 k P(k) e^{i x \cdot k}; \quad (1.1.20)$$

in practice $\xi(x)$ and $P(k)$ are a Fourier pair. The variance is equivalent to $\xi(0)$, and we can express it as:

$$\sigma^2 = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) dk. \quad (1.1.21)$$

### 1.1.2 The Standard Cosmological Model

In the last decades astronomers found growing evidence of the presence of a mass component that interacts only gravitationally. This mass component does not emit electromagnetic radiation and for this reason it has been called Dark Matter (DM). A classic test in
favour of the existence of DM consists in the analysis of the rotation curves of spiral galaxies. The prediction of the shape of these rotation curves, based on the gravitational effect of the luminous mass, does not match with the observed rotation curves. This evidence stimulated astronomers to postulate the existence of a dark mass component sustaining the rotation of spirals up to the observed radii. Another strong indirect evidence of the existence of the DM is the measurement of the temperature anisotropies of the Cosmic Microwave Background (CMB), that match very well with the prediction of DM models. The nature of DM has strong effects on the cosmology, in particular on the growth of perturbations that give rise to the galaxy population. The proposed DM models differ in the relativistic or non-relativistic nature of DM particles. In the case of relativistic DM particles (hot dark matter models, top-down scenario), larger structures form first, then they fragment in smaller structures. On the contrary, in the case of non-relativistic DM particles (cold dark matter models, bottom-up scenario) smaller structures form first, then merge to form larger structures. The Λ-Cold Dark Matter (ΛCDM) model gained increasing importance in recent times; it assumes non-relativistic DM particles and takes into account the cosmological constant Λ. In this model matter density Ω_m = Ω_b + Ω_{DM} is smaller than 1, and a vacuum component Ω_Λ, is also taken into account. The ΛCDM scenario, together with the Big Bang model, implies that the universe was very dense and very hot in the early times. In this conditions matter was completely ionized and all particle species (photons included) were in thermodynamical equilibrium. At these stages the universe was opaque to electromagnetic radiation. During the following expansion the temperature decreased, until light nuclei and electrons started to combine in atoms, thus decoupling from photons: from this moment on, photons can propagate freely in the universe. This is called the recombination epoch, occurred at z ∼ 1100. We can observe at z = 0 this radiation as an almost uniform cosmic background in the microwaves (CMB). The existence of the CMB is a very stringent confirmation of the Big Bang model. Moreover, due to the coupling of matter and radiation before recombination, small CMB temperature anisotropies trace the fluctuations in the density field at this epoch.

1.1.2.1 The Cosmological Constant

In the development of his theory of gravity, Einstein was interested in finding static solutions for his equations, both due to his hope that general relativity would embody Mach’s principle that matter determines inertia, and simply to account for the astronomical data as they were understood at the time. The way to do this was to add a new term involving a free parameter Λ, the cosmological constant, in the side of his equation which describes the geometry of the space-time. Assuming Λ > 0 this modification allowed a static solution for a universe filled with dust of zero pressure and mass density

\[ \rho = \frac{\Lambda}{8\pi G}. \quad (1.1.22) \]

When the expansion of the universe was discovered by Hubble (1929), the cosmological constant solution was abandoned by the scientific community. Meanwhile, particle theorists (see Weinberg 1989, for a detailed discussion) realized that anything that contributes to the energy density of the vacuum, the state of lowest energy of a system, acts just like
1.1. Cosmological Context

A cosmological constant. In this picture the cosmological constant would be part of the energy-momentum tensor, being a source for the gravitational field. This interpretation introduces the so-called “cosmological constant problem,” that is currently one of the most significant unsolved problems in fundamental physics: applying the standard quantum field theory the expected contribution at the vacuum energy density is of order:

$$\rho_\Lambda \sim (10^{18} \text{Gev})^4 \sim 2 \times 10^{110} \text{erg/cm}^3,$$

(1.1.23)

while cosmological observations imply

$$\rho_\Lambda \sim (10^{-12} \text{Gev})^4 \sim 2 \times 10^{-10} \text{erg/cm}^3.$$

(1.1.24)

The ratio between 1.1.23 and 1.1.24 is the origin of the famous discrepancy of 120 orders of magnitudes between the theoretical and observational values of the cosmological constant. This and other problems connected with the cosmological constant (see Padmanabhan 2003) have led the community to consider the possibility that this “Dark Energy” in the universe is not just a constant but is of more complicated nature, evolving with time. Assuming a dark energy term independent of time, recent results (Spergel et al. 2007; Santos et al. 2008) are consistent with the presence of a non-zero cosmological constant; at the present even in the case of more exotic dark energy models the cosmological solution seem to be the best (Kurek et al. 2008).

1.1.2.2 Cosmological parameters

The ΛCDM scenario is completely described by the values of a set of cosmological parameters: a basic set consists of the Hubble constant $H_0$, the total matter density $\Omega_m$, the baryon $\Omega_b$, radiation $\Omega_r$ and neutrino $\Omega_\nu$ densities, the cosmological constant $\Lambda$, the density perturbation spectral index $n$, the density perturbation amplitude $\sigma_8$. These parameters define the overall evolution of the universe and the growth of structures. The most powerful technique to measure the value of cosmological parameters relies on the measurement of the anisotropies in the CMB. The angular power spectrum $C_l$ of anisotropies, depends on all the cosmological parameters and it was observed by several groups (see i.e. de Bernardis et al. 2000). The Wilkinson Microwave Anisotropy Probe (WMAP) satellite provides the most accurate measurement of the CMB fluctuation, and the most accurate determination of cosmological parameters (Spergel et al. 2007). However, to obtain tight constraints other data sets need to be considered in addition to WMAP. Suitable data sets come from galaxies (Sánchez et al. 2006), galaxy clusters (Borgani et al. 2001; Allen et al. 2003), distant supernovae (Perlmutter et al. 1999; Knop et al. 2003) and the Lyα forest (Viel et al. 2006). In spite of the different nature of these datasets their results are in good agreement. In this thesis, I will adopt the following set of parameters: $\Omega_m = 0.26$, $\Omega_\Lambda = 0.74$, $\Omega_b = 0.0463$, $n_s = 0.95$, $\sigma_8 = 0.85$ and $H_0 = 72 \text{ km/s}^{-1} \text{ Mpc}^{-1}$, in agreement with the previous cited observational results.
1.1.3 Structure Formation

1.1.3.1 Linear regime

The commonly adopted theory for structure formation is the gravitational instability scenario, in which primordial density perturbations grow through gravitational Jeans instability to form all the structures we observe today. To derive the evolution of these primordial density fluctuations into bound objects, we can proceed as follows. Let us describe the universe in terms of a fluid made of collisionless dark matter and baryons, with a mean density $\bar{\rho}$. At any time and location, the mass density can be written as $\rho(x, t) = \bar{\rho}(t)[1 + \delta(x, t)]$, where $x$ indicates the comoving spatial coordinates and $\delta(x, t)$ is the mass density contrast. The time evolution equation for $\delta$ during the linear regime ($\delta \ll 1$) reads (Peebles 1993):

$$\ddot{\delta}(x, t) + 2H(t)\dot{\delta}(x, t) = 4\pi G\bar{\rho}(t)\delta(x, t) + \frac{c_s^2}{a(t)^2}\nabla^2\delta(x, t). \quad (1.1.25)$$

Here, $c_s$ is the sound speed and $H(t) = H_0[\Omega_m(1 + z)^3 + \Omega_\Lambda]^{1/2}$. The second term on the left hand side of the above equation represents the effect of cosmological expansion. This, toghether with the pressure support (second term on the right hand side) acts against the growth of the perturbation due to the gravitational collapse (first term on the right hand side). Pressure in the baryonc gas is essentially provided by collisions, while in the collisionless dark matter component the pressure support arises from the readjustement of the particle orbits. The above equation can be used also to follow separately the evolution of the different components of a multi-component medium. In this case, $c_s$ would be the velocity of the perturbed component which is gravitationally dominant (as it drives the collapse of the perturbation). The equation has two independent solutions, one of which grows in time and governs the formation of structures. The evolution of the single Fourier components of the density constrast is then given by:

$$\ddot{\delta}_k + 2H\dot{\delta}_k = \left(4\pi G\bar{\rho} - \frac{k^2c_s^2}{a^2}\right)\delta_k. \quad (1.1.26)$$

This sets a critical wavelength, the Jeans lengt ($L_J$), at which the competing pressure and gravitational forces cancel (Jeans 1928):

$$L_J = \frac{2\pi a}{k_J} = \left(\frac{\pi c_s^2}{G\bar{\rho}}\right)^{1/2}. \quad (1.1.27)$$

For $l \gg L_J$ the pressure term is negligible because the response time for the pressure wave is long compared to the growth time for the density contrast, and the zero pressure solutions apply. On the contrary, at $l < L_J$ the pressure force is able to counteract gravity and the density contrast oscillates as a sound wave. It is conventional to introduce also the Jeans mass as the mass within a sphere of radius $L_J/2$:

$$M_J = \frac{4\pi}{3}\bar{\rho}\left(\frac{L_J}{2}\right)^3. \quad (1.1.28)$$
In a perturbation with mass greater than $M_J$ the pressure force is not counteracted by gravity and the structures collapse. This sets a limit on the scales that are able to collapse at each epoch and has a different value according to the component under consideration, reflecting the differences in the velocity of the perturbed component. Given the initial power spectrum of the perturbations, the evolution of each mode can be followed through eq. 1.1.26 and then integrated to recover the global spectrum at any time. The CDM power spectrum in three dimensions is given by:

$$P_{DM}^{(3)}(k) = A_{DM} k^n T_{DM}^2(q),$$

(1.1.29)

where $n$ is the spectral index and $T_{DM}(q)$ is the CDM transfer function (Bardeen et al. 1986):

$$T_{DM}(q) = \frac{\ln(1 + 2.34q)}{2.34q} [1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4]^{-1/4}$$

(1.1.30)

with $q \equiv k/(h\text{Mpc}^{-1})\Gamma^{-1}$. The shape parameter $\Gamma$ depends on the Hubble parameter, $\Omega_m$ and $\Omega_b$ (Sugiyama 1995):

$$\Gamma = \Omega_m h \exp \left[ -\Omega_b \left( 1 + \frac{\sqrt{2h}}{\Omega_m} \right) \right].$$

(1.1.31)

The normalization parameter $A_{DM}$ is fixed through the value of $\sigma_8$ (the rms density fluctuations in spheres of radius 8 $h^{-1}$ Mpc). Although the initial power spectrum is a pure power law, perturbation growth results in a modified final power spectrum. In fact, while on large scales the power spectrum follows a simple linear evolution, on small scales it changes shape due to the additional non-linear gravitational growth of perturbations and it results in a bended spectrum, $P(k) \propto k^{n-4}$. The amplitude of the power spectrum, however, is not specified by current models of inflation and must be determined observationally. Note that most of the power of the fluctuation spectrum of the standard CDM model is on small scales; therefore these are the first to become non-linear.

### 1.1.3.2 Non-linear regime: Dark Matter

The simplest model to follow the evolution of the perturbations in Gaussian density fields is the one of a spherically symmetric, constant density region, for which the collapse can be followed analytically. At a certain point the region reaches the maximum radius of expansion, then it turns around and starts to contract ("turnaround point"). In the absence of any symmetry violation, the mass would collapse into a point. However, long before this happens, the dark matter experiences a violent relaxation process and quickly reaches virial equilibrium. If we indicate with $z$ the redshift at which such a condition is reached, the halo can be described in terms of its virial radius, $r_{vir}$, circular velocity, $v_c = \sqrt{GM/r_{vir}}$, and virial temperature, $T_{vir} = \mu m_p v_c^2/2k_B$, whose expressions
are (Barkana & Loeb 2001):

\[ r_{\text{vir}} = 0.784 \left( \frac{M}{10^8 h^{-1} M_\odot} \right)^{1/3} \left[ \frac{\Omega_m}{\Omega_m^{\ast}} \frac{\Delta_{\text{vir}}(z)}{18\pi^2} \right]^{-1/3} \left( \frac{1 + z}{10} \right)^{-1} h^{-1} \text{kpc}, \]  

(1.1.32)

\[ v_c = 23.4 \left( \frac{M}{10^8 h^{-1} M_\odot} \right)^{1/3} \left[ \frac{\Omega_m}{\Omega_m^{\ast}} \frac{\Delta_{\text{vir}}(z)}{18\pi^2} \right]^{1/6} \left( \frac{1 + z}{10} \right)^{1/2} \text{km/s}^{-1}, \]  

(1.1.33)

\[ T_{\text{vir}} = 2 \times 10^4 \left( \frac{\mu}{0.6} \right) \left( \frac{M}{10^8 h^{-1} M_\odot} \right)^{2/3} \left[ \frac{\Omega_m}{\Omega_m^{\ast}} \frac{\Delta_{\text{vir}}(z)}{18\pi^2} \right]^{1/3} \left( \frac{1 + z}{10} \right) \text{K}. \]  

(1.1.34)

Here, \( \mu \) is the mean molecular weight, \( m_p \) is the proton mass, and (Bryan & Norman 1998):

\[ \Delta_{\text{vir}}(z) = 18 \pi^2 + 82(\Omega_m^{\ast} - 1) - 39(\Omega_m^{\ast} - 1)^2, \]  

(1.1.35)

\[ \Omega_m^{\ast} = \frac{\Omega_m(1 + z)^3}{\Omega_m(1 + z)^3 + \Omega_\Lambda}. \]  

(1.1.36)

Although spherical collapse captures some of the physics governing the formation of halos, their inner structure should be investigated through numerical simulations. Navarro et al. (1996, 1997) have simulated the formation of dark matter halos of masses ranging from dwarf galaxies to rich clusters, finding that their density profile has a universal shape, independent of the halo mass, the initial density fluctuation spectrum and the cosmological parameters:

\[ \rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}, \]  

(1.1.37)

where \( \rho_s \) and \( r_s \) are characteristic density and radius. Usually, the quantity \( c \equiv r_{\text{vir}}/r_s \), the concentration parameter, is introduced. Thus, it results that DM halos produced by numerical simulations are characterized by cuspy density profiles. On the other hand, kinematical observations in galaxies are at odds with the predicted DM distribution, favouring a cored profile (Salucci & Burkert 2000). This discrepancy between simulations and observations is argument of a very hot debate.

### 1.1.3.3 Non-linear regime: Baryonic gas

In contrast to dark matter, as long as the gas is fully ionized, the radiation drag on free electrons prevents the formation of gravitationally bound systems. The residual ionization of the cosmic gas keeps its temperature locked to the CMB temperature through different physical processes, down to a redshift \( 1 + z_t \simeq 1000(\Omega_b h^2)^{2/5} \) (Peebles 1993). When the radiation decouples from matter, perturbations in this component are finally able to grow in the pre-existing dark matter halo potential wells and eventually form the first bound objects. The process leading to the virialization of the gaseous component of matter is similar to the dark matter one. In this case, during the contraction following the turnaround point, the gas develops shocks and gets reheated to a temperature at which pressure support can prevent further collapse. The mass of these bound objects can be derived from eqs. 1.1.27-1.1.28, where \( c_s \) is the sound velocity of the baryonic gas. In
1.2 Gunn-Peterson test and the Ionization of the IGM

QSO spectra have been used to study the ionization history of the IGM through the so-called Gunn-Peterson test (Gunn & Peterson 1965). This test makes it possible to determine the value of the density of neutral hydrogen in the IGM measuring the transmission blue-ward of the Lyα emission in QSO spectra. As the emission redshift of the object goes to higher values, i.e., approaching the cosmic reionization epoch, the density of neutral hydrogen grows and therefore the transmission blue-ward of the Lyα emission line drops. Fig. (1.1) shows a clear example of this behaviour. With simple calculations it is possible to estimate the neutral fraction from QSO spectra; defining the cross section for the Lyα transition as

\[ \sigma(\nu) = \sigma \left( \nu \frac{a_0}{a} \right) = \pi e^2 \frac{f}{m_e c} \phi \left( \nu \frac{a_0}{a} \right), \]  

\[ a \] being the scale factor, \( e \), \( m_e \) the charge and mass of the electron, \( f = 0.4162 \) the oscillator strength for the Lyα transition, \( \phi \) a shape function that as a first approximation

\[ \text{At redshift of 100, the fractional abundance of molecular hydrogen was } 10^{-6} \text{ (Galli & Palla 1998). Other molecules that were important coolants are HD and LiH, which had abundances } 10^{-3} \text{ and } 10^{-4} \text{ times lower than that of } H_2. \]
Figure 1.1: Optical spectra of $z \gtrsim 5.8$ QSOs observed with Keck/ESI, in the observed frame. The spectrum of SDSS1044-0125 has been taken from Fan et al. (2000). In each spectrum, the expected wavelengths of prominent emissions lines, as well as the Lyman limit, are indicated by the dashed lines. Reproduced from Becker et al. (2001).
we can define as a Dirac function $\phi(\nu) \simeq \delta(\nu - \nu_{\text{Ly}\alpha})$. Then, $\tau_{\text{GP}}$ will be the optical depth for an absorption, defined as the integral over the light path of the neutral hydrogen density times the cross section:

$$\tau_{\text{GP}} = \int \sigma(\nu)n_{\text{HI}} \, dl,$$

$$dl = cdt = \frac{da}{a},$$

$$\tau_{\text{GP}} = \int \sigma(\nu)n_{\text{HI}} \frac{da}{a} = \int \frac{\sigma(\nu)n_{\text{HI}} c}{H_0 a E(a)} da,$$

$H_0$ being the Hubble constant and $E(a) = \sqrt{\Omega_m(\frac{a_0}{a})^3 + \Omega_r(\frac{a_0}{a})^4 + \Omega_\Lambda}$ its evolution with the scale factor. This integral can be solved, the equation inverted to obtain $n_{\text{HI}}$ for a given observed optical depth at a given redshift:

$$\tau_{\text{GP}} = \frac{\pi e^2}{m_e \nu_{\text{Ly}\alpha}} \frac{\sigma(\nu)}{H_0 E(z)}$$

$$n_{\text{HI}} = 2.4131 \times 10^{-11} h \tau(z) E(z)$$

where $h = H_0/(100$ km s$^{-1}$ Mpc$^{-1}$). To obtain the ionization fraction, the mean value of hydrogen of the universe as a function of redshift has to be estimated. Under the assumption that almost all the hydrogen in the universe is in diffuse form, the following relation holds:

$$\frac{n_{\text{HI}}}{\bar{n}_{\text{H}}} \approx \Omega_b \rho_0 = 1.124 \times 10^{-5} \Omega_b h^2 \text{prot cm}^{-3},$$

with $Y$ the helium abundance (estimated from the Big Bang nucleosynthesis for instance), $\Omega_b$ the baryonic density fraction in the universe. The combination of Eq. 1.2.4 and 1.2.5 gives the relation between the ionization fraction of hydrogen in the universe and the observed optical depth, at fixed cosmology and redshift. The highest redshift quasars that can be observed are at $z \sim 5 - 6$, when the Einstein-De Sitter approximation (i.e. $\Omega_\Lambda = \Omega_r = 0$, $\Omega_m = 1$) is relatively accurate; in this condition we can quantitatively estimate the ionizing fraction, obtaining:

$$\tau_{\text{GP}} < 0.1 \text{ @ } z \sim 5 \Rightarrow \frac{n_{\text{HI}}}{\bar{n}_{\text{H}}} < 10^{-6} h$$

$$\tau_{\text{GP}} \sim 5 \text{ @ } z \sim 6 \Rightarrow \frac{n_{\text{HI}}}{\bar{n}_{\text{H}}} < 4 \times 10^{-5} h$$

These relations are valid until $n_{\text{HI}}/\bar{n}_{\text{H}} \ll 1$. This means that even assuming a huge absorption, with almost no transmission blue-ward of the Ly$\alpha$ emission in a $z = 6$ QSO, the ionized fraction would be $\sim 99.995\%$. Therefore, the absorption in the Ly$\alpha$ forest is very sensitive to the fraction of neutral hydrogen in the universe and tell us that the IGM at least below $z \sim 6$ is an highly ionized medium.

### 1.2.1 Constraining Reionization with QSO absorption spectra

The GP effect has been also used to constrain the end of the reionization epoch: fig. 1.2 shows the results by Becker et al. (2001); they analized a sample of QSOs at low resolution,
from the Sloan Digital Sky Survey (SDSS) looking at their GP optical depth, claiming that the universe is approaching the reionization epoch at \( z \sim 6 \), and finding the first evidence of a complete Gunn-Peterson trough in the spectrum of SDSS J1030+0524 \( (z=6.28) \), where no transmitted flux is detected over a large region \( (300 \ \text{Å}) \) immediately blue-ward of the Ly\( \alpha \) emission line. Fan et al. (2002) detected the appearance of very large portions of completely absorbed flux in the observed spectra, interpreting it as an evidence for the complete reionization of the universe to occur at \( z \simeq 6 \). Due to the extremely high sensitivity of \( \tau_{GP} \) to tiny neutral hydrogen amounts, the detection of a Gunn-Peterson trough only translates into a lower limit for the hydrogen neutral fraction. Recently, many studies have tried to clarify whether the SDSS data actually require that the IGM was reionized as late as \( z \simeq 6 \) (Gallerani et al. 2006; Becker et al. 2007): in particular, Gallerani et al. (2006) have shown that currently available QSO spectra are compatible with a highly ionized universe at that redshift. The spectroscopy of the Ly\( \alpha \) forest for QSOs at \( z > 6 \) discovered by the SDSS seems to indicate that the ionization state of the IGM might be very different along different lines of sight. For example, the analysis of the spectrum of one of the most known distant QSO (SDSS J1148+5251) show some residual flux both in the Ly\( \alpha \) and Ly\( \beta \) troughs which, when combined with the Lyman\( \gamma \) region (Furlanetto & Oh 2005) implies this flux is not consistent with pure transmission, despite previous claims (White et al. 2003). The presence of unabsorbed regions in the spectrum corresponds to a highly ionized IGM along that particular line of sight. The fact that transmission is detected along some lines of sight, while the medium seems quite neutral along others, possibly implies that the IGM ionization properties are different along different lines of sight at \( z \gtrsim 6 \), thus suggesting that we might be observing the end of a patchy reionization process (Wyithe & Loeb 2006).

### 1.3 The UV background from the Proximity Effect

The main source of ionization in the low density IGM is the UV cosmic background (UVB). Quasar spectra can be used to estimate the strength of this flux: QSOs are strong ionization sources, and the signature of their ionizing power is the deficiency of Ly\( \alpha \) absorptions in the vicinity of their emission redshift. This statistical deficiency of absorption can be used to infer the size of the QSO influence sphere and, knowing its luminosity, to infer the strength of the UV background. To explain that, we follow here the approach by Giallongo et al. (1996). The line distribution far away from the QSO as a function of density and redshift can be parametrized as:

\[
\frac{\partial^2 n}{\partial z \partial N(\text{H}1)} = A_0 (1 + z)^\gamma \begin{cases} N(\text{H}1)^{-\beta} & N(\text{H}1) < N_{\text{break}} \\ N(\text{H}1)^{-\beta_1} N(\text{H}1)^{-\beta_2} & N(\text{H}1) \geq N_{\text{break}} \end{cases}
\]

where

\[
N_{\text{break}}(z) = \frac{N_{\infty,b}}{1 + \omega(z)}
\]

Near the QSO, highly ionized absorbers are observed with a column density

\[
N(\text{H}1) = \frac{N_{\infty}}{1 + \omega}
\]
1.3. The UV background from the Proximity Effect

Figure 1.2: Evolution of transmitted flux ratio and effective Gunn-Peterson optical depth as functions of redshift. The solid line is the expected evolution if the number density of Lyα clouds increases as $N(z) \propto (1+z)^{2.5}$. No flux is detected in the spectrum of SDSS 1030+0524 at $z_{\text{abs}} \sim 6$, indicating $\tau_{\text{eff}} > 5.0$. The non-detection of flux in the Lyα trough gives a substantially stronger 1σ upper limit of $\tau_{\text{eff}} > 20$. Reproduced from Becker et al. (2001).
1.3. The UV background from the Proximity Effect

i.e., the ratio between the intrinsic column density \( N_\infty \), which the same absorber would have at an infinite distance from the QSO, and the factor \( 1 + \omega \), where

\[
\omega(z) = \frac{F}{4\pi J}
\]

(1.3.4)
is the ratio between the flux \( F \) that the absorber receives from the QSO and the flux \( J \) that the absorber receives from the general UVB (see Bajtlik et al. (1988) and Bechtold (1994) for details of the model). The following conservation law near the QSO emission redshift can be obtained

\[
f(N) = g(N_\infty) dN_\infty/dN = g(N_\infty)(1 + \omega),
\]

(1.3.5)

where \( f(N) \) and \( g(N_\infty) \) are the column density distributions near the QSO and at infinite distance. Then the following double power law distribution can be used to get a simultaneous estimate of the Ly\( \alpha \) parameters of the \( N(\text{H}^1) \), \( z \) distributions and of the UVB \( J_{LL}(z) \) evaluated at the Lyman limit:

\[
\frac{\partial^2 n}{\partial z \partial N(\text{H}^1)} = A_0 (1 + z)^\gamma (1 + \omega)^{1-\beta_f}
\]

(1.3.6)

\[
\times \begin{cases} N(\text{H}^1)^{-\beta_f} & N(\text{H}^1) < N_{\text{break}} \\ N(\text{H}^1)^{-\beta_s} N(\text{H}^1)^{\beta_s - \beta_f} & N(\text{H}^1) \geq N_{\text{break}} \end{cases}
\]

(1.3.7)
is the observed break, which is shifted to lower and lower \( N(\text{H}^1) \) as the QSO emission redshift is approached. Giallongo et al. (1996) used six high redshift QSOs to estimate \( J \); considering the evolution of the UVB with redshift:

\[
J = J_{z=3} \left( \frac{1 + z}{4} \right)^j,
\]

(1.3.8)

and applying a maximum likelihood analysis for the lines with \( \log N(\text{H}^1) \geq 13.3 \) the authors obtained a value \( J_{-22} = J \times 10^{-22} = 5 \pm 1 \) and \( j = -0.28 \pm 1.41 \), negative but consistent with no evolution in the redshift interval \( z = 2 - 4 \). The UVB is a crucial quantity to interpret the evolution of the number density of Ly\( \alpha \) absorption with redshift (see Sec. 1.5.1), and the QSOs contribution to it is not enough to explain the observed relations: a contribution from galaxies is then required. Bianchi et al. (2001) computed the contribution to UVB both from QSOs and galaxies. Their conclusion, summarized in fig. 1.3, is that a galaxy dominated background with a fraction of ionizing photon escaping from a galaxy, \( f_{\text{esc}} \lesssim 0.1 \) is consistent with the estimates of \( J \), and \( f_{\text{esc}} \gtrsim 0.05 - 0.1 \) is needed in order to explain the observed evolution of Ly\( \alpha \) absorbers with redshift. Similar results have been obtained by Bolton & Haehnelt (2007); they used a large set of hydrodynamical simulations, combined with measurements of the Ly\( \alpha \) opacity of the IGM taken from the literature, to obtain robust estimates of the photoionization rate per hydrogen atom at \( z = 5 \) and 6. The authors suggest that the combined ionizing emission from star forming galaxies and QSOs is capable of maintaining the IGM in its highly ionized state if \( f_{\text{esc}} \gtrsim 0.2 \) and that the ionizing emission from star forming galaxies is likely to dominate the total ionizing photon budget at redshift \( z > 5 \).
Figure 1.3: UV background at $\lambda = 912\,\text{Å}$ for the models with $f_{\text{esc}}=0.05$, 0.1 and 0.4 (solid lines), in a flat $\Omega_m = 1$ universe. Also shown are the separate contribution of the QSOs (dotted line) and of the galaxies (dashed lines, each corresponding to a value of $f_{\text{esc}}$). The shaded area refer to the Lyman limit UV background estimated from the proximity effect (Giallongo et al. 1996; Cooke et al. 1997; Scott et al. 2000). The arrow shows an upper limit for the local ionising background (Vogel et al. 1995). The datapoint at $z = 3$ is derived from a composite spectrum of Lyman-break galaxies (Steidel et al. 2000). Reproduced from Bianchi et al. (2001).
1.4 Absorption Line Definition

To gain physical information from the lines of a spectrum, we have to model them with some analytical profile whose parameters are related to properties of the gas from which lines originated. The energy intensity of radiation, $I_\nu$, will vary when passing through a cloud of gas due to absorption and emission; this process is described by the equation of radiative transfer,

$$dI_\nu = -k_\nu \rho I_\nu dx + j_\nu \rho dx$$  \hspace{1cm} (1.4.1)

where $k_\nu$ is the absorption coefficient per unit mass, and the emission coefficient $j_\nu$ is defined as the emission per unit mass, per unit frequency interval, into a unit solid angle. Equation (1.4.1) becomes,

$$\frac{dI_\nu}{k_\nu \rho dx} = -I_\nu + \frac{j_\nu \rho}{k_\nu \rho}$$  \hspace{1cm} (1.4.2)

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu(\tau),$$  \hspace{1cm} (1.4.3)

where $d\tau_\nu = k_\nu \rho dx$ defines the optical depth through the cloud and $S_\nu(\tau)$ is the source function. Since we are concerned only with absorption lines, we can put $S_\nu(\tau) = 0$ and thus easily solve Eq. (1.4.3) as $I_\nu = I_{00} \exp(-\tau_\nu)$. The absorption coefficient, $k_\nu$, is related to the absorption cross-section per atom, $\sigma_\nu$, by $k_\nu \rho = \sigma_\nu N_{abs}$, where $N_{abs}$ is the number of atoms per unit volume capable of absorbing radiation at the appropriate frequency. The cross-section in turn can be expressed as the product of the line profile and a constant for a particular line:

$$\sigma_\nu = \sigma_{\nu_0} \phi(\nu)$$  \hspace{1cm} (1.4.4)

where the line profile is normalized so $\int \phi(\nu) d\nu = 1$ and hence $\int \sigma_\nu = \sigma_{\nu_0}$. The oscillator strength $f$ is defined by

$$\sigma_{\nu_0} = \frac{\pi e^2}{4\pi \epsilon_0 mc} f$$  \hspace{1cm} (1.4.5)

where $m$ and $e$ are the mass and the charge of the electron. The line profile is determined by the combined effects of different broadening.

**The doppler broadening** due to the motion of individual absorbing atoms. The frequency $\nu$ of a line absorbed by an atom moving with velocity $v_r$ in the line of sight is $\nu = \nu_0 (1 - v_r/c)$, where $\nu_0$ is the rest frequency of the line. The sign of $\Delta_\nu = \nu - \nu_0$ is negative (i.e., a redshift) if $v_r$ is positive, i.e. away from the observer. The normalized profile of the line $\phi(\nu)$ will be given by the distribution of line of sight velocities $v_r$ since the radiation that we observe is normally the result of the radiation of many atoms. If we assume that the absorbing atoms have a gaussian velocity dispersion (a Maxwell-Boltzmann thermal distribution), we obtain for the thermal Doppler broadening,

$$\phi(\nu) = \frac{1}{\sqrt{\pi} \Delta \nu_D} e^{-(\Delta_\nu/\Delta \nu_D)^2}$$  \hspace{1cm} (1.4.6)
1.4. Absorption Line Definition

with the Doppler width

$$\Delta \nu_D = \frac{\nu_0}{c} \sqrt{\frac{2kT}{m}} = \frac{\nu_0}{c} b_{th}$$

where $\nu_0$ is the frequency of the line centre, and $b_{th} = \sqrt{2kT/m}$ is the thermal Doppler parameter. There can be also larger scale collective motions of group of atoms, called turbulence, that contribute to the Doppler line broadening. In case the turbulence can be modelled as a Gaussian velocity distribution with quadratic mean velocity $\sqrt{\frac{3}{2}}V$ and mean velocity $2V/\sqrt{\pi}$, the resulting profile is still given by Eq. (1.4.6), but with Doppler parameter $b_{tot}^2 = b_{th}^2 + V^2$.

The natural and pressure line broadening. Every spectral line has an intrinsic width due to the fact that excited states have in general very short lifetimes and by Heisenberg’s uncertainty principle, $\Delta E \Delta t \sim h/2\pi$, so there will be a spread in energy level. The pressure or collisional broadening is instead the effect of de-excitation of the upper level caused by interactions with other particles. The two effects give a similar line shape, the so called Lorentzian profile

$$\phi(\nu) = \frac{\Gamma}{4\pi^2 (\Delta \nu)^2 + \left(\frac{\Gamma}{4\pi}\right)^2}$$

where the Lorentzian width $\Gamma$ is the sum of the collisional and natural line widths and microscopically is the de-excitation rate of the upper level, or total damping constant. The overall line profile always contains both components. The line centre (core) is dominated by the Doppler profile, while the collisional/natural component is dominating the wings of the line. The convolved profile is given by

$$\phi(\nu) = \frac{a}{\pi^{3/2} \Delta \nu_D} \int_{-\infty}^{\infty} \frac{e^{-x^2}}{(u-x)^2 + a^2} dx$$

where $u = \Delta \nu/\Delta \nu_D$ is the frequency distance from the line centre in Doppler widths and $a = \Gamma/(4\pi \Delta \nu_D)$ is the ratio of the Lorentz width to the Doppler width. The convolved profile is called the Voigt profile. The integral has to be evaluated numerically, and $a/\pi$ times the integral in Eq. (1.4.9) is called the Voigt function

$$H(a, u) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-x^2}}{(u-x)^2 + a^2} dx.$$  

(1.4.10)

Now, going back to the optical depth definition and switching from frequency to wavelength, we obtain the expression

$$\tau(\lambda) = \frac{\pi e^2}{4\pi \epsilon_0 mc} f N \frac{\lambda_0}{\sqrt{\pi b}} H(a, u)$$

$$= 1.498 \times 10^{-2} f N \frac{\lambda_0}{b} H(a, u)$$

(1.4.11)

where, $\lambda_0$ is the wavelength of the centre of the line and $N$ is the column density, i.e., the number of atoms along the line of sight in a cylindrical volume of unitary base. The optical
1.4. Absorption Line Definition

The depth at the centre of the line is:

\[ \tau_0 = 1.489 \times 10^{-15} \frac{f N(\text{cm}^{-2}\lambda(\text{Å}))}{b(\text{km s}^{-1})} \]  

(1.4.13)

Absorption lines are observed against a continuous background. For each point in the profile, observations give the depth \( r \) of the line,

\[ r = \frac{(I_c - I)}{I_c} \]  

(1.4.14)

where \( I \) is the observed spectral intensity and \( I_c \) the interpolation of the absorption-free continuum over the absorption feature. The overall line strength is given by the equivalent width, \( w_{\text{obs}} \), defined as:

\[ w_{\text{obs}} = \int \frac{I_c - I}{I_c} d\lambda = \int (1 - e^{-\tau(\lambda)}) d\lambda. \]  

(1.4.15)

If \( r_0 \) is the normalized depth at the line centre and \( \Delta \lambda_{1/2} \) the full width at half minimum (FWHM), for a rectangular profile \( w_{\text{obs}} = r_0 \Delta \lambda_{1/2} \). Hence the equivalent width can also be defined as the width in wavelength of a rectangular profile line 100\% deep which has the same area in a flux-wavelength plot as the actual line. For redshifted absorption lines, \( w_{\text{obs}} = w_r \times (1 + z_{\text{abs}}) \).

1.4.1 The curve of growth

It is apparent from Eq. (1.4.15) that the observed equivalent width does not depend on the spectral resolution. Thus, this is the parameter that can be measured for lines in spectra of low to intermediate resolution, when the instrumental width of lines is larger than the typical Doppler parameter, \( b \), of QSO absorption lines. From Eq. (1.4.15), (1.4.12), (1.4.10) it is possible to relate the equivalent width to the column density for different values of the Doppler parameter. The function that gives this relation is called the curve of growth and an example corresponding to the Hi Lyα transition is plotted in Fig. (1.4) for \( b = 5, 10, 20 \) and 30 km s\(^{-1}\) (Petitjean 1998). Three distinct regimes can be identified:

- When the column density is small (\( \tau_0 < 0.1 \)), the absorption line is optically thin, the Voigt function reduces to a Gaussian profile and the equivalent width does not depend on \( b \). This is the linear part of the curve of growth, where the determination of \( N \) from \( w \) is easy and reliable. For any transition,

\[ N(\text{cm}^{-2}) = 1.13 \times 10^{20} \frac{w_r(\text{Å})}{\lambda_0^2(\text{Å}) f} \]  

(1.4.16)

- The logarithmic or flat part of the curve of growth is characterized by the large dependence of \( N \) on \( b \) at a given \( w \). In this regime, the determination of \( b \) and \( N \) are very uncertain except when several lines of the same ion are used. Equivalent width and optical depth at the centre of the lines (see Eq. (1.4.13)) are related by:

\[ \frac{w}{\lambda_0} = 2\frac{b}{c} \sqrt{\ln \tau_0} \]  

(1.4.17)
1.4. Absorption Line Definition

Figure 1.4: Curve of growth: logarithm of the equivalent width ($w$ in Å) versus logarithm of the column density ($N$ in cm$^{-2}$) for different values of the Doppler parameter ($b$ in km s$^{-1}$). The curves are calculated for H~I Ly$\alpha$ $\lambda = 1215$Å. The three characteristic regimes are illustrated (see the text). Reproduced from Petitjean (1998)
• We have so far taken the line as having a Gaussian profile, since the ratio $\Gamma/(4\pi\Delta\nu_D)$ is usually much less than one and the Gaussian element dominates except in the wings. However, as the line becomes stronger, the core will saturate while the wings can still grow. Eventually they will supply the major part of the equivalent width. In this case the equivalent width can be found approximately by assuming that the line has a pure Lorentzian profile. This gives rise to the damping portion of the curve of growth in which the equivalent width no longer depends on $b$ and the column density determination is again accurate. In that case (see Eqs. (1.4.10) and (1.4.13)),

$$\frac{w}{\lambda_0} = 2.64 \frac{b\sqrt{a}}{c} \sqrt{\tau_0}$$

(1.4.18)

### 1.5 Lyα forest evolution as a function of redshift

---

**Figure 1.5:** Illustration of structure evolution of intergalactic gas from high to low redshift. The upper spectrum of a $z = 3.63$ quasar is a Keck/HIRES observation, while the lower spectrum is a FOS/HST observations of a $z = 1.33$ quasar. Higher redshift quasars show a much thicker forest of Lyα lines. Slices through N-body/hydrodynamic simulation results at the two epochs $z = 3$ and $z = 1$ are shown in the right-hand panels. Three contour levels in column density are shown: $10^{11}$ cm$^{-2}$ (dotted lines), $10^{12}$ cm$^{-2}$ (solid lines) and $10^{13}$ cm$^{-2}$ (thick solid lines). Evolution proceeds so that the voids become more empty so that even the low column density material is found in filamentary structures at low redshifts. Reproduced from Charlton et al. (2000).

The appearance of the Lyα forest region in QSO spectra is strongly dependent on the emission redshift of the QSO itself. Fig. (1.5) is a clear example of the observed evolution of the Lyα forest and of the environment that hosts the QSO. Qualitatively, the evolution

---
can be understood in terms of two main factors. The UV radiation field is expected to decrease from \( z \sim 2 \) to \( z = 0 \), whether it is comprised of photons from quasars or star-forming galaxies. Thus as the UV field decreases the neutral fraction increases and the number of detected lines should increase. This tendency is balanced and overcome by a second effect: as cosmic expansion proceeds, the density of the gas decreases and even if the photoionizing UV field was constant, the neutral fraction would decrease and the number of Ly\( \alpha \) lines detected would decrease. There are two main ways to look at this evolution, the more classical line number density evolution, based on the counts of line found in the spectra, and the new-fashioned evolution of the H\( \text{I} \) opacity, based on the flux statistics of the spectra. In the following, we describe both of them.

1.5.1 Line number density evolution

![Figure 1.6: The number density evolution of the Ly\( \alpha \) forest over the column density range \( N(\text{H}\text{I}) = 10^{13.64-17} \text{cm}^{-2} \), which is comparable to the HST data (open triangles) of Weymann et al. (1998). The data are shown for the binned sample for display. The filled symbols are derived from the UVES observations of Kim et al. (2002). Open squares, the star, open circles, and the diamond are taken from the HIRES data by Hu et al. (1995), Lu et al. (1996), Kim et al. (1997), and Kirkman & Tytler (1997), respectively. The horizontal error bars represent the \( z \) interval over which the number density was estimated. The vertical error bars represent the Poisson \( \sigma \) error. Reproduced from Kim et al. (2002).](image)

In an individual QSO line of sight, observations of the Ly\( \alpha \) forest can extend over a redshift range \( \Delta z \) greater than unity. Then we are sampling a significant fraction of a Hubble time, and it is natural to look for changes in the rate of incidence of absorption lines with redshift. Peterson (1978) first found evidence for an increase in the number of Ly\( \alpha \) clouds with redshift beyond what was expected for a population of objects with invariant cross-section and uniform comoving number density. Sargent et al. (1980) drew attention to the need of defining a consistent sample by setting a common lower limit to the intrinsic equivalent width \( w_r \) of the Ly\( \alpha \) lines, and they found lack of evolution
for their sample. The presence of evolution was at first subject to some debate (see the summary by Murdoch et al. (1986)), but it is now clear that the Lyα forest as a whole evolves quite strongly with $z$. The observationally determined evolution in the number density of absorbers above a certain column density threshold is usually expressed in the form:

$$\frac{dN}{dz} = \left(\frac{dN}{dz}\right)_0 (1 + z)^\gamma$$

(Sargent et al. 1980; Young et al. 1982). Much observational effort has been devoted to studying the redshift number density evolution, but unfortunately the resulting conclusions are far from robust. This is because Lyα column densities $N(H\text{I})$ are distributed according to a power law $N(H\text{I})^{-\beta}$ with index $\beta \sim 1.5$ (see next section) so the majority of lines in any column density limited sample are always close to the threshold, and small variations in imposing the threshold can cause large changes in the estimated number of lines, and in $\gamma$. Moreover, the normalization $(dN/dz)_0$ is very sensitive to line blending: since it depends on spectral resolution, data quality, and redshift it is possible that individual studies differ by a factor of two or more. Fig. (1.6) (Kim et al. 2002) shows a recent computation of $dN/dz$, giving as best fit the value $\gamma = 2.47 \pm 0.18$. As can be seen in fig. (1.6) there is a clear change in evolution at low redshifts, due to the drop in the UVB, resulting from the decline in the QSO population. In the absence of structure evolution, it is possible to derive an analytical approximation for the evolution of $dN/dz$ with $J_\nu$ and the Hubble expansion. For absorbers in photoionization equilibrium with the background, it is easy to show that the evolution of lines above a given threshold in

Figure 1.7: Number density evolution of the Lyα forest with $N_{HI} = 10^{13.64-16}$ cm$^{-2}$, for two different cosmologies. Dotted lines refer to the evolution compatible with an ionising UV background due only to QSOs. Solid lines show the evolution when both QSOs and galaxies contribute to the background, for the models with $f_{esc}=0.05$ (upper line), 0.1 and 0.4 (lower line). Data points come from several observations in the literature for the column density range $N_{HI} = 10^{13.64-16}$ cm$^{-2}$, as given by Kim et al. (2001). The modelled evolution has been normalized to the observed evolution in the redshift range $2 < z < 3$. Reproduced from Bianchi et al. (2001).
The differential density distribution function, \( f(N(HI)) \), is defined as the number of absorption lines per unit absorption distance path and per unit column density as a column density can be written as (Davé et al. 1999):

\[
\left( \frac{dN}{dz} \right)_{N(HI)>N(HI)_{\text{min}}} \propto \frac{1}{H(z)} \left[ \frac{(1+z)^5}{\Gamma_{HI}(z)} \right]^{\beta-1},
\]

where \( H(z) \) is the Hubble parameter and \( \beta \) the coefficient of the power-law distribution of clouds with column density. \( \Gamma_{HI}(z) \) is the photo-ionisation rate

\[
\Gamma_{HI}(z) = \int_{v_0}^{\infty} \frac{4\pi J(\nu, z)}{h\nu} \sigma_{HI}(\nu) d\nu,
\]

with \( \nu_0 \) the frequency of the Lyman limit and \( \sigma_{HI}(\nu) = \sigma_{HI}(\nu_0/\nu)^3 \) the HI photo-ionisation cross-section. Bianchi et al. (2001) showed that the UVB due only to QSOs is not enough to explain the change in the \( dN/dz \) evolution, and the contribution from galaxies has to be taken into account, as shown in fig. 1.7.

### 1.5.2 Evolution of the HI mean opacity

With the new view of the Ly\( \alpha \) forest the emphasis of the analysis has shifted to statistical measures of the transmitted flux, more suitable for absorption arising from a continuous density field. The tool to study the evolution of the Ly\( \alpha \) forest with redshift in this case is the mean HI opacity, directly related with the flux of the spectra, and therefore with the absorption strength of the underlying Ly\( \alpha \) forest, by means of the following relation:

\[
F_\lambda = F_c e^{-\tau_{HI}(\lambda)},
\]

where \( F_\lambda \) is the observed flux at a certain wavelength and \( F_c \) the continuum of the spectrum. The mean HI opacity provides a more straightforward comparison between different data sets than the line counts. Moreover, the mean HI opacity does not rely on the subjective line counting method, although it is more subject to the uncertainties in the determination of the continuum. Fig. (1.8), taken from Kim et al. (2002), shows results from many authors. Again from this figure it can be seen that the UVB estimated only from QSOs cannot explain the evolution of the Ly\( \alpha \) forest. The shaded area is indeed the result of simulations by Davé et al. (1999) where only the contribution from QSOs is taken into account. In Kim et al. (2007) the authors did a careful analysis of a large sample of data, studying the flux probability distribution function (PDF) but also the HI effective optical depth and its evolution with redshift. They took carefully into account the metal contamination in the spectra; they get to the conclusion that the effect of the removal of metal absorption, on the measured effective optical depth due to HI absorption, is small and their measurement of the evolution of \( \tau_{HI} \) is in agreement with previous results.

### 1.6 The Differential Density Distribution Function

The differential density distribution function, \( f(N(HI)) \), is defined as the number of absorption lines per unit absorption distance path and per unit column density as a
1.6. The Differential Density Distribution Function

Figure 1.8: The H\textsubscript{i} opacity as a function of z. Filled symbols represent the mean H\textsubscript{i} opacity. The open circle at $z \geq 2.0$ represents of J2233-606 when two high column density systems are excluded. Other symbols at $z > 1.5$ are open circles (Hu et al. 1995), large square (Lu et al. 1996), the open triangle (Kirkman & Tytler 1997), and diamonds (Rauch et al. 1997). Symbols at $z < 1.5$ with arrows are from: open diamonds (Impey et al. 1996), open triangles (Bahcall et al. 1993), open stars (Impey et al. 1999) and open squares (Penton et al. 2000). Owing to the low resolution and low S/N spectra of the HST observations, $\langle \tau \rangle$ could be highly underestimated if the Ly\textalpha forest at $N(\text{H}\textsubscript{i}) \leq 10^{14}$ cm\textsuperscript{-2} contains the bulk of the neutral hydrogen. $\langle \tau_{\text{H}\textsubscript{i}} \rangle$ estimates, while the y-axis error bars were estimated from simply changing the adopted continuum by $\pm 5$ per cent. The dotted line represents the commonly used formula by (Press et al. 1993), $\langle \tau_{\text{H}\textsubscript{i}}(z) \rangle = 0.0037(1 + z)^{3.46}$. The shaded area enclosed with dot-dashed lines indicates the ranges of $\langle \tau \rangle$ expected from different cosmological simulations by Davé et al. (1999). Reproduced from Kim et al. (2002).
function of \( N(\text{H} \text{I} ) \). This quantity plays a central role in absorption line studies, providing similar information as the luminosity function in the study of galaxies. The absorption distance path \( X(z) \) is defined by

\[
X(z) = \int_0^z dz (1+z)^2 \frac{H_0}{H(z)},
\]

(1.6.1)

\( H(z) \) being the Hubble parameter dependent on the chosen cosmology. Empirically \( f(N(\text{H} \text{I} )) \) is fitted by a power law: \( f(N(\text{H} \text{I} )) = f_0 N^{-\beta}(\text{H} \text{I} ) \). We already saw in section 1.4.1 that there are two regimes, the linear and the damping portion of the curve of growth, in which the determination of the column density is quite unambiguous. In the central part where the line starts to saturate but the damping wings are not yet visible, the column densities are relatively difficult to measure exactly. Simultaneous fits to the higher order Lyman lines (which are less saturated owing to their lower oscillator strengths) can help to some degree. Fig. 1.6 shows the observed \( f(N(\text{H} \text{I} )) \) at different

![Figure 1.9: The differential density distribution functions at \( z \geq 2.1, 3.3 \) and 3.8, without the incompleteness correction. The data are shown in the binned sample for display. Open triangles are the differential density distribution functions for the damped Lyα systems at \( 1.5 < z < 2.5 \) (Storrie-Lombardi & Wolfe 2000). The dotted line represents the incompleteness-corrected \( f(N(\text{H} \text{I} )) \) at \( z \geq 2.85 \) from Hu et al. (1995), \( f(N(\text{H} \text{I} )) = 4.9 \times 10^7 N(\text{H} \text{I} )^{-1.46} \). Reproduced from Kim et al. (2002).](image)

redshift ranges without the incompleteness correction due to line blending. The dotted line represents the incompleteness-corrected \( f(N(\text{H} \text{I} )) \) at \( \langle z \rangle = 2.85 \) from Hu et al. (1995), \( f(N(\text{H} \text{I} )) = 4.9 \times 10^7 N(\text{H} \text{I} )^{-1.46} \). Triangles for the damped Lyα systems at \( 1.5 < z < 2.5 \) are taken from Storrie-Lombardi & Wolfe (2000). At \( z \sim 2.1 \), \( f(N(\text{H} \text{I} )) \)
at $N(\text{H}i) = 10^{12.5-14.5}$ cm$^{-2}$ is in good agreement with the incompleteness-corrected $f(N(\text{H}i))$ at $z \sim 2.8$. It is also true for $f(N(\text{H}i))$ at $z \sim 3.3$, although the goodness-of-the-fit is lower at $N(\text{H}i) = 10^{12.5-13}$ cm$^{-2}$. In general, $f(N(\text{H}i))$ is well approximated by a single power-law $f(N(\text{H}i)) \propto N(\text{H}i)^{-1.5}$ at $N(\text{H}i) = 10^{12.5-22}$ cm$^{-2}$ (Petitjean et al. 1993). As noted by Petitjean et al. (1993) and Kim et al. (1997), $f(N(\text{H}i))$ starts to deviate from the empirical power-law at $N(\text{H}i) > 10^{14}$ cm$^{-2}$. The amount of this deviation increases as $z$ decreases since the higher $N(\text{H}i)$ forest disappears more rapidly as $z$ decreases. In addition, the deviation $N(\text{H}i)$ at which $f(N(\text{H}i))$ starts to deviate decreases as $z$ decreases. At $z = 3.8$, 3.3 and 2.1, $f(N(\text{H}i))$ deviates from the power-law at $N(\text{H}i) \sim 10^{16}$ cm$^{-2}$, $N(\text{H}i) \sim 10^{14.5}$ cm$^{-2}$ and $N(\text{H}i) \sim 10^{14.2}$ cm$^{-2}$, respectively. At $N(\text{H}i) > 10^{15.6}$ cm$^{-2}$, the deviation from the single power-law increases more than $3\sigma$ at $z \sim 2.1$ and more than $2\sigma$ at $z \sim 3.3$.

1.7 The Width of the Lines and the Temperature of the IGM

The width of the absorption lines, described by the Doppler parameter $b$ of the Voigt fitting in the Ly$\alpha$ forest, is directly related to the temperature of the absorbing gas in the IGM, by means of eq.(1.4.7). For a photoionized gas, a temperature-density relation holds: $T = T_0 (1 + \delta)_{\gamma_T-1}$, where $T$ is the gas temperature, $T_0$ is the gas temperature at the mean gas density, $\delta$ is the baryon density contrast, and $\gamma_T$ is a constant which depends on the reionization history (Hui & Gnedin 1997). For an abrupt reionization at $z \gg 5$, the temperature of the mean gas density decreases as $z$ decreases after the reionization, eventually approaching an asymptotic $T_0$. For a generally assumed QSO-dominated UV background with a sudden turn-on of QSOs at $5 < z < 10$, $T_0$ decreases as $z$ decreases at $2 < z < 4$ (Hui & Gnedin 1997). Under the assumption that there are some lines which are broadened primarily by the thermal motion at any given column density, this equation of state translates into a lower cutoff $b(N(\text{H}i))$ envelope in the $N(\text{H}i)-b$ distribution: $T$ and $\delta$ can be derived from $b$ and $N(\text{H}i)$. For the equation of state $T = T_0 (1 + \delta)_{\gamma_T-1}$, $b_c(N(\text{H}i))$ becomes

$$\log(b_c) = \log(b_0) + (\Gamma_T - 1) \log(N(\text{H}i)),$$

where $\log(b_0)$ is the intercept of the cutoff in the $\log(N(\text{H}i))$-$\log b$ diagram and $(\Gamma_T - 1)$ is the slope of the cutoff (Schaye et al. 1999). The cutoff slope $(\Gamma_T - 1)$ is proportional to $(\gamma_T - 1)$. This cutoff envelope provides a probe of the gas temperature of the IGM at a given $z$, thus giving a powerful constraint on the thermal history of the IGM (Hu et al. 1995; Lu et al. 1996; Kim et al. 1997; Kirkman & Tytler 1997; Zhang et al. 1997; Schaye et al. 1999; Bryan & Machacek 2000; McDonald et al. 2000; Ricotti et al. 2000; Schaye et al. 2000). In practice, defining $b_c(N(\text{H}i))$ in an objective manner is not trivial due to the finite number of available absorption lines, sightline-to-sightline cosmic variances, limited S/N, and unidentified metal lines. There are several methods proposed to derive $b_c(N(\text{H}i))$, we refer to Kim et al. (2001), Hu et al. (1995), McDonald et al. (2000), Theuns & Zaroubi (2000) for a review of them. All these different methods agree on a typical Doppler parameter in the Ly$\alpha$ forest of $b \sim 20 - 30$ km s$^{-1}$, corresponding to an IGM temperature $T_{IGM} \sim 10^4$K. Bolton et al. (2007) recently found a possible evidence...
for an inverted temperature-density relation ($\gamma_T < 1$) from the flux distribution of the Ly$\alpha$ forest at $z \simeq 3$. This result suggests that the voids in the IGM maybe significantly hotter and the thermal state of the low density IGM may be significantly more complex than is usually assumed at $z \simeq 3$.

1.8 Two point correlation function of the Ly$\alpha$ lines

The forest of Ly$\alpha$ absorption lines can be used to study the spatial distribution of matter in the universe in a redshift and scale range not reachable with other observations. Other classes of objects that can provide this type of information are QSOs themselves or galaxies; however, Ly$\alpha$ absorbers are preferred to QSOs for this kind of analysis because the absorbers are much more numerous and are expected to trace more ‘normal’ environments. The current understanding of the Ly$\alpha$ forest as tracer of the matter density field, highlighted its cosmological nature and therefore its fundamental role in understanding the formation of cosmic structures. Traditionally the clustering properties of the Ly$\alpha$ for-

![Figure 1.10: Two-point correlation function in the velocity space](image)

Figure 1.10: Two-point correlation function in the velocity space: a) (upper panel) for the complete sample of Lyman-$\alpha$ lines, b) (lower panel) for lines with column densities $< 10^{13.6} \, \text{cm}^{-2}$. The short-dashed and long-dashed lines represent the 1$\sigma$ and 2$\sigma$ confidence limits for a Poissonian process. Reproduced from Cristiani et al. (1997).
Two point correlation function of the Ly$_\alpha$ lines are studied with the Two Point Correlation Function (TPCF) of the lines in analogy with studies of the clustering of galaxies. This function is defined as the excess, due to clustering, of the probability $dP$ of finding a Ly$_\alpha$ absorber in a linear portion $dl$ at a distance $r$ from another absorber: $dP = \phi_{\text{Ly}_\alpha}(z) dl [1 + \xi(r)]$, where $\phi(z)$ is the average linear density of the absorbers as a function of $z$. Operatively this quantity is estimated with the formula (Peebles 1980):

$$\xi(v) = \frac{N_{\text{obs}}(v)}{N_{\text{exp}}(v)} - 1,$$  

(1.8.1)

where $N_{\text{obs}}$ is the observed number of line pairs with velocity separations between $v$ and $v + dv$, and $N_{\text{exp}}$ is the number of pairs expected in the same range of separations from a random distribution in redshift. The TPCF is known to be a satisfactory estimator when used to investigate weak clustering on scales considerably smaller that the total interval covered by the data, which is the domain where Ly$_\alpha$ clustering lies. In the middle of the '90s, it became clear that Ly$_\alpha$ lines cluster at small separations, i.e. $\Delta v \lesssim 300$ km s$^{-1}$; the clustering signal was seen to increase with increasing column density of the lines, but no significant signal was observed for lines with $\log N(\text{H}1) \lesssim 13.6$. Fig. (1.10) and (1.11) show the results from a paper by Cristiani et al. (1997). The TPCF of the lines can constrain cosmological parameters, once compared with numerical simulations characterized by different parameters.

![Figure 1.11: Two-point correlation function in the velocity space for lines with column densities $> 10^{13.8}$ cm$^{-2}$. Confidence limits as in Fig. (1.10). The continuous line shows a model described in the paper. Reproduced from Cristiani et al. (1997).](image-url)
1.9 Flux Statistics of the Lyα forest

The advent of the new picture of the Lyα forest opened up the possibility to probe the density fluctuations of matter with the Power Spectrum of the transmitted flux in QSO absorption spectra (Croft & Gaztanaga 1998; Croft et al. 1999; Hui 1999; McDonald et al. 2000; Hui et al. 2001; Croft et al. 2002; McDonald 2003; Viel et al. 2003, 2004b). Important results from hydro-simulations are: a tight correlation between the HI and the dark matter distribution (on scales larger than the Jeans length ($L_J$) of the IGM), and a simple temperature density relation for the underdense to moderately overdense gas, see section 1.7 (Hui & Gnedin 1997). For most of the gas in this density regime, the Lyα optical depth is proportional to the neutral hydrogen density (Gunn & Peterson 1965) which, since the gas is in photoionization equilibrium, is proportional to the density times the recombination rate:

\[
\tau \propto \rho^2 T^{-0.7} = A(\rho/\bar{\rho})^{\gamma_T},
\]

\[
A = 0.433 \left(\frac{1+z}{3.5}\right)^6 \left(\frac{\Omega_b h^2}{0.02}\right)^2 \left(\frac{T_0}{6000 K}\right)^{-0.7} \times \left(\frac{h}{0.65}\right)^{-1} \left(\frac{H(z)/H_0}{3.68}\right)^{-1} \left(\frac{\Gamma}{1.5 \times 10^{-12} s^{-1}}\right)^{-1},
\]

with $\gamma_T$ varying in the range 1.6 – 1.8 and $\Gamma$ the HI photoionization rate. Because Eq. (1.9.1) describes the analog of Gunn-Peterson absorption for a non-uniform, photoionized medium (ignoring the effect of peculiar velocities), it has been called the Fluctuating Gunn-Peterson Approximation (Rauch et al. 1997; Weinberg et al. 1998). The transmitted flux in a QSO spectrum, $F = e^{-\tau}$, is monotonically related to $\rho$ in this approximation. Because the relation between $\tau$ and $\rho$ is fairly simple, one can extract information about the underlying mass density field from the observed flux distribution. The standard procedure to derive the matter power spectrum from the QSO spectra follows three simple steps (Croft et al. 2002; Viel et al. 2004b):

- estimate of the 1D flux power spectrum from the spectrum $P_F(k)$, using as flux estimator $F/\langle F \rangle - 1$, where $F$ is the flux of the continuum-fitted spectrum.
- obtaining the 3D flux power spectrum by means of

\[
P_F^{3D} = -\frac{2\pi}{k} \frac{dP_F^{1D}}{dk}
\]

- then obtain the matter power spectrum through a bias function $b(k)$:

\[
P_F(k) = b^2(k) P_M(k)
\]

The bias function is estimated from hydro-simulations and depends on cosmological parameters, mean flux level, and temperature. The measurement of the amplitude of the matter power spectrum on scales of a few Mpc with the Lyα forest have been used to constrain the spectral index of primordial density fluctuations $n$ and the $\sigma_8$ parameter, the rms fluctuation amplitude of the density for a 8 h$^{-1}$ Mpc sphere. Viel et al. (2006)
show that the Lyα forest data are consistent to within $1\sigma$ with the three-year WMAP results (Spergel et al. 2007). The errors analyzing the Lyα data are large, and dominated by systematic uncertainties.
CHAPTER 2

New Sample, Old Tools

We have collected a large sample of high-resolution QSO spectra in order to better understand the physics of the IGM, to reconstruct and study the hydrogen density field at redshifts $z \sim 3$ by means of a new approach that will be described in Chapter 3. This Chapter describes the data set used and how it has been analyzed to study the classical statistics of the Ly$\alpha$ forest described in the previous chapter. The line number density evolution with redshift and the two point correlation function have been studied to show the consistence with other results in the literature, the improvements given by this sample and the careful analysis that have been carried.

All the studies described in this chapter and the next one, chapter 3, refer to Saitta et al. (2007) and Bruscoli et al. (2008 in prep.). My own contribution to this project was to

- develop and test with numerical simulations the method to reconstruct the H-density field,
- fit all the features in the observed spectra with Voigt profiles and make the classical analysis (line number density evolution, two point correlation function) of the Ly$\alpha$ forest,
- apply the method to observed data to study the properties of the field (the evolution in redshift, the correlation function and the power spectrum).

2.1 Simulated Data Sample

We used simulations run with the parallel hydro-dynamical (TreeSPH) code GADGET-2 based on the conservative ‘entropy-formulation’ of SPH (Springel 2005). They consist of a cosmological volume with periodic boundary conditions filled with an equal number of dark matter and gas particles. Radiative cooling and heating processes are followed for a
primordial mix of hydrogen and helium. A mean UVB produced by QSOs and galaxies is assumed as given by Haardt & Madau (1996) with helium heating rates multiplied by a factor 3.3 in order to better fit observational constraints on the temperature evolution of the IGM. This background gives naturally a $\Gamma \sim 10^{-12}$ (H ionisation rate) at the redshifts of interest here (Bolton et al. 2005). The star formation criterion is a very simple one that converts in collision-less stars all the gas particles whose temperature falls below $10^5$ K and whose density contrast is larger than 1000 (it has been shown that the star formation criterion has a negligible impact on flux statistics). More details can be found in Viel et al. (2004b). 2 × $400^3$ dark matter and gas particles in a 120 $h^{-1}$ comoving Mpc box (although for some cross-checks have been analyzed some smaller boxes of 60 $h^{-1}$ comoving Mpc) are used. The gravitational softening is set to 5 $h^{-1}$ kpc in comoving units for all particles. We stress that the parameters chosen here, including the thermal history of the IGM, are in reasonably good agreement with observational constraints including recent results on the CMB and other results obtained by the Ly$\alpha$ forest community (e.g. Viel et al. 2006).

Mock spectra are obtained from the simulations, in two different ways depending on the application:

- Studying the gas distribution at redshifts $\sim 1.7$-3.5: the 120 Mpc simulation box at $z = 2$ is pierced to create a set of 364 mock lines of sight covering a redshift range $\Delta z \simeq 0.11$. For each of these lines of sight, we know the density contrast, the temperature, and the peculiar velocity pixel by pixel. Peculiar velocities are small, typically less than 100 km s$^{-1}$, and randomly oriented, so their contribution, e.g. to the correlation function, is in general negligible. However, we choose to modify the redshifts of the density field in order to take into account the peculiar velocity field using the formula $v_{pec}(z_{old}) = c (z_{new} - z_{old})/(1+(z_{new}+z_{old})/2)$, where $v_{pec}$ is the velocity used to shift a line from redshift $z_{old}$ to redshift $z_{new}$, to highlight the effects of our transformation alone.

- Studying the proximity effect and the overdensity field around QSOs: we analysed the output at $z = 2.2$ and run a Friend-of-Friend (FoF) algorithm to identify the most massive collapsed haloes that should host the QSO. We found about 200 (54) haloes whose total mass is larger than $10^{12} M_\odot/h$ ($10^{13} M_\odot/h$). We then pierced the simulated box along the 200 LOSs intersecting the center of the haloes with $M \geq 10^{12} M_\odot/h$ and along 200 random LOSs. This latter sample constitutes our 'control' sample. We explicitly checked that the correlation function of the haloes with masses larger than $10^{12} M_\odot/h$ is well fitted by a power-law function with $r_0 = 6$ Mpc/h and slope $-1.8$ in agreement with other theoretical and observational results.

The simulated lines of sight have been fitted with Voigt profiles using an automated version of VPFIT.
### Table 2.1: Summary of the main characteristics of our QSO sample.

<table>
<thead>
<tr>
<th>QSO</th>
<th>$z_{em}$</th>
<th>$J$ mag</th>
<th>$\Delta z_{Ly\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE1341-1020</td>
<td>2.139</td>
<td>18.68</td>
<td>1.658-2.087</td>
</tr>
<tr>
<td>Q0122-380</td>
<td>2.203</td>
<td>17.34</td>
<td>1.711-2.150</td>
</tr>
<tr>
<td>PKS1448-232</td>
<td>2.220</td>
<td>17.09</td>
<td>1.725-2.166</td>
</tr>
<tr>
<td>PKS0237-23</td>
<td>2.233</td>
<td>16.61</td>
<td>1.737-2.179</td>
</tr>
<tr>
<td>J2233-606</td>
<td>2.250</td>
<td>16.97</td>
<td>1.753-2.196</td>
</tr>
<tr>
<td>HE0001-2340</td>
<td>2.267</td>
<td>16.74</td>
<td>1.765-2.213</td>
</tr>
<tr>
<td>HE1122-1648</td>
<td>2.400</td>
<td>16.61</td>
<td>1.878-2.344</td>
</tr>
<tr>
<td>Q0109-3518</td>
<td>2.407</td>
<td>16.72</td>
<td>1.884-2.350</td>
</tr>
<tr>
<td>HE2217-2818</td>
<td>2.415</td>
<td>16.47</td>
<td>1.891-2.358</td>
</tr>
<tr>
<td>Q0329-385</td>
<td>2.440</td>
<td>17.20</td>
<td>1.912-2.383</td>
</tr>
<tr>
<td>HE1158-1843</td>
<td>2.451</td>
<td>17.09</td>
<td>1.921-2.394</td>
</tr>
<tr>
<td>HE1347-2457</td>
<td>2.560</td>
<td>17.35</td>
<td>2.016-2.502</td>
</tr>
<tr>
<td>Q0453-423</td>
<td>2.662</td>
<td>17.69</td>
<td>2.100-2.602</td>
</tr>
<tr>
<td>PKS0329-255</td>
<td>2.698</td>
<td>17.88</td>
<td>2.125-2.636</td>
</tr>
<tr>
<td>HE0151-4326</td>
<td>2.761</td>
<td>17.48</td>
<td>2.186-2.699</td>
</tr>
<tr>
<td>Q0002-422</td>
<td>2.768</td>
<td>17.50</td>
<td>2.189-2.705</td>
</tr>
<tr>
<td>HE2347-4342</td>
<td>2.878</td>
<td>17.12</td>
<td>2.281-2.814</td>
</tr>
<tr>
<td>HS1946+7658</td>
<td>3.051</td>
<td>16.64</td>
<td>2.429-2.984</td>
</tr>
<tr>
<td>HE0940-1050</td>
<td>3.088</td>
<td>17.08</td>
<td>2.463-3.020</td>
</tr>
<tr>
<td>Q0420-388</td>
<td>3.122</td>
<td>17.44</td>
<td>2.489-3.053</td>
</tr>
<tr>
<td>PKS2126-158</td>
<td>3.275</td>
<td>17.54</td>
<td>2.620-3.204</td>
</tr>
<tr>
<td>HE1341-1020</td>
<td>3.620</td>
<td>16.22</td>
<td>2.911-3.543</td>
</tr>
</tbody>
</table>

Figure 2.1: $Ly\alpha$ forest redshift coverage of the considered QSO sample.
2.2 Observed Data Sample

Most of the observational data used in this work were obtained with the UVES spectrograph (Dekker et al. 2000) at the Kueyen unit of the ESO VLT (Cerro Paranal, Chile) in the framework of the ESO Large Programme (LP): “The Cosmic Evolution of the IGM” (Bergeron et al. 2004). Spectra of 18 QSOs were obtained in service mode with the aim of studying the physics of the IGM in the redshift range 1.7-3.5. The spectra have a resolution $R \sim 45000$ and a typical signal to noise ratio (SNR) of $\sim 35$ and 70 per pixel at 3500 and 6000 Å, respectively. Details of the data reduction can be found in Chand et al. (2004); Aracil et al. (2004).

We added to the main sample 4 more QSOs spectra with comparable resolution and SNR:

- J2233-606 (Cristiani & D’Odorico 2000). Data for this QSO were acquired during the commissioning of UVES in October 1999.
- HE1122-1648 (Kim et al. 2002). Data for this QSO were acquired during the science verification of UVES in February 2000. The reduced and fitted spectrum was kindly provided to us by Tae-Sun Kim.
- HS1946+7658 (Kirkman & Tytler 1997). Data for this QSO were acquired with Keck/HIRES in July 1994.
- B1422+231 (Rauch et al. 1996). Data for this QSO were acquired with Keck/HIRES in 1996. The reduced and fitted spectrum was kindly provided to us by Tae-Sun Kim.

Table 2.1 summarises the main properties of our QSO sample. None of our QSOs is a Broad Absorption Line (BAL) object. Magnitudes are taken from the GSC-II catalogue (McLean et al. 2000). Figure 2.1 shows the distribution in redshift of the Ly$\alpha$ forests for all the QSOs of the sample. For most of the application described in the following we considered for each QSO the redshift range between 1000 km s$^{-1}$ red-ward of the Ly$\beta$ emission, in order to avoid contamination by associated Ly$\beta$ lines, and 5000 km s$^{-1}$ blue-ward of the Ly$\alpha$ emission to exclude the region affected by the proximity effect due to the ionising flux of the QSO. Studying the proximity effect itself instead the redshift range goes to the Ly$\alpha$ emission. The coverage is good over the whole redshift range $z \simeq 1.7 - 3.5$, with most of the signal concentrated between $z \sim 2$ and 2.5. In Fig. 2.2, we show a portion of the Ly$\alpha$ forest of the QSO HE0001-2340 compared with the same wavelength region in a mock spectrum extracted from the considered simulation box at $z = 2$. 
2.2. Observed Data Sample

![Portion of the Ly$\alpha$ forest for an observed (lower panel) and a simulated (upper panel) line of sight in our sample.](image)

Figure 2.2: Portion of the Ly$\alpha$ forest for an observed (lower panel) and a simulated (upper panel) line of sight in our sample.

### 2.2.1 Creation of the line lists

All the lines in the Ly$\alpha$ regions were fitted with the FITLYMAN tool (Fontana & Ballester 1995) of the ESO MIDAS data reduction package$^1$. In the case of complex saturated lines we used the minimum number of components to reach $\chi^2 \leq 1.5$. Whenever possible, the other lines in the Lyman series were used to constrain the fit. The spectra of HE1122-1648, HS1946+7658, B1422+231 and all the simulated lines of sight were fitted with the VPFIT$^2$ package. Both software tools model absorption features with a Voigt profile convolved with the instrument line spread function. The minimum H$\text{I}$ column density detectable at 3$\sigma$, at the lower SNR of the spectra in our sample, is $\log N(\text{H} \text{I}) \simeq 12$. Metals in the forest were identified and the corresponding spectral regions were masked to avoid effects of line blanketing. We eliminated Ly$\alpha$ lines with Doppler parameters $b \leq 10$ km s$^{-1}$, that are likely unidentified metal absorptions. In a total amount of 8435 fitted Ly$\alpha$ lines, 368 (4.4%) fall in the masked intervals, 1150 (13.6%) are at less than 1000 km s$^{-1}$ red-ward the Ly$\beta$ emission or at less than 5000 km s$^{-1}$ blue-ward the Ly$\alpha$ emission, while 599 were eliminated because they have $b \leq 10$ (7.1%). The output of this analysis is a list of Ly$\alpha$ lines for each QSO with central redshift, column density, Doppler parameter and the corresponding errors obtained with FITLYMAN. In the line fitting approach to the study of the Ly$\alpha$ forest, each line is considered as the signature of an absorber. As a consequence statistical measures are computed with the population of absorption lines, representative of the population of absorbers. Our sample of fitted Ly$\alpha$ lines is the largest, homogeneous sample ever gathered up to now.

---

$^1$http://www.eso.org/midas

$^2$http://www.ast.cam.ac.uk/~rfc/vpfit.html
2.3 Line Number density evolution

In Fig. 2.3 we plot the result for the QSOs in our sample for the standard column density interval $13.64 < \log N(\text{H}i) < 17 \text{ cm}^{-2}$ in order to compare our statistics with the HST low redshift measurement\textsuperscript{3} (Weymann et al. 1998). The best fit to our data gives: $dn/dz = (166 \pm 4) \times ((1 + z)/3.5)^{2.8\pm0.2}$ (1 $\sigma$ errors). There is no substantial change in the trend with respect to previous results by Kim et al. (2001, 2002) who used smaller samples of UVES QSO spectra of the same quality. However, our points are systematically higher on the plot, with an increase in $\log dn/dz$ amounting to $\sim 0.03$ at $z \sim 2$ up to $\sim 0.1$ at $z \sim 3$. The discrepancy arises from the fact that we have taken into account the decrease in the available redshift interval due to the presence of metal lines ‘masking’ the Ly$\alpha$ features. High resolution spectra allow to identify a larger number of metal lines: in our sample these metal masks correspond to about 9 percent of the total redshift interval covered by the observed Ly$\alpha$ forests. Fig. 2.4 shows the number density evolution for two different $\text{H}i$ column density ranges: $13 \leq \log N(\text{H}i) \leq 14$ and $14.5 \leq \log N(\text{H}i) < 17 \text{ cm}^{-2}$. The linear fit in these intervals gives slopes of $1.9 \pm 0.2$ and $3.8 \pm 0.4$ for the weak and the strong lines selection, respectively. This trend was already noticed by Kim et al. (2002): stronger lines have a steeper number density evolution than the weaker ones.

\textit{Figure 2.3: Number density evolution of the Ly$\alpha$ forest lines over the column density range $13.64 < \log N(\text{H}i) < 17 \text{ cm}^{-2}$ for the 22 QSOs in our sample (open triangles). The solid line traces the best linear fit obtained for those data (see text). For comparison, we report also previous measurements at high redshift and the result of the low redshift HST campaign.}

\textsuperscript{3}The lower limit in column density is due to the fact that HST measurements have been transformed from equivalent width into $\text{H}i$ column densities assuming a typical Doppler parameter of $30 \text{ km s}^{-1}$.
2.4 TPCF of the Lyα lines

As we have seen in the previous chapter, the TPCF was usually computed in redshift space, to avoid the effect of peculiar velocities. Since in this context peculiar velocities are negligible (see e.g. Rauch et al. 2005), we compute the correlation function in real space, measuring separations in comoving Mpc. At the characteristic redshift of our sample, $z = 2.5$, a velocity separation $\Delta v = 100 \text{ km s}^{-1}$ corresponds to $\Delta r \approx 0.9$ comoving Mpc, in our fiducial cosmology. $N_{\text{exp}}$ is obtained by averaging the results of 1000 numerical simulations of the number of lines observed in each QSO spectrum. In particular, the set of line redshifts is randomly generated in the same redshift interval as the data according to the observed distribution $\propto (1 + z)^\beta$, where we adopt the value $\beta = 2.8$ found in the previous section. The same mock line lists are used to estimate the error on the observed correlation function by determining the $1 \sigma$ standard deviation of the correlation functions of the randomly distributed lines. Lines closer than 0.3 comoving Mpc, are merged into a single line with redshift equal to the mean redshift, weighted with the column densities, and column density equal to the sum of the column densities. The minimal separation is set by the intrinsic blending due to the tipical width of the lines (see Giallongo et al. 1996). We compute the correlation function for the whole data set (Fig. 2.5) and for two column density cuts (Fig. 2.6) to investigate the clustering properties of strong and weak lines. Previous results (Cristiani et al. 1995; Lu et al. 1996; Cristiani et al. 1997) already showed a significant clustering signal for strong absorptions, which is confirmed and strengthened by our data. Furthermore, we also see a significant clustering for the weak lines, consistent with previous results by Misawa et al. (2004). The amplitude is
2.4. TPCF of the Lyα lines

Figure 2.5: Two point correlation function for the observed Lyα lines in the column density range $12 < \log N(\text{H} \text{I}) < 17 \text{ cm}^{-2}$. In the bottom panel lines closer than one Jeans length have been merged into one line, see text. The dashed lines represent the $1\sigma$ confidence levels from a random distribution of lines.

Figure 2.6: Two point correlation function for the observed Lyα lines in two column density ranges as reported in the panels.
2.4. TPCF of the Lyα lines

Figure 2.7: Two point correlation function for the observed Lyα lines in the column density range $13.8 < \log N(\text{H}_i) < 17$ and in two redshift ranges reported in the panels.

about one order of magnitude lower than for the stronger lines but the clustering signal in the first bin is significant at the 7σ level. As already said in chapter 1, the Jeans length ($L_J$) likely represents the typical size of IGM structures detected as Lyα absorptions. This length (varying from $\sim 1.2$ to $1.6$ comoving Mpc for the maximum and minimum redshift of our sample, respectively) is also comparable with the clustering scale of the Lyα lines as shown in Figs. 2.5 (upper panel) and 2.6. In order to verify that the clustering signal we are detecting is not only due to structures internal to the absorbers, we perform the following test. Lines with separation less than the local $L_J$ are merged into a single line with column density equal to the sum of the column densities of the component features and redshift equal to the $N(\text{H}_i)$-weighted mean of the component redshifts and the TPCF is re-computed. The result, reported in the lower panel of Fig. 2.5, shows that the clustering signal is preserved substantially at the same level of the one computed with all the lines, with a slightly decreased significance due to the smaller statistics. This is an indication that Lyα absorbers cluster among themselves and not only inside themselves. The present data set is large enough to allow studying the evolution of the correlation function with redshift. We consider the column density range for which the signal is stronger, $13.8 < \log N(\text{H}_i) < 17$, and we divide our sample in two parts. The first sub-sample is formed by objects with emission redshift $z_{\text{em}} \leq 2.5$, for which the average Lyα forest redshift is $\langle z_{\text{Lyα}} \rangle = 2.07$, and the second sub-sample has objects with $z_{\text{em}} > 2.5$ and $\langle z_{\text{Lyα}} \rangle = 3.02$. Results are shown in Fig. 2.7: the high redshift lines are less clustered than the low redshift lines. This apparent evolution with redshift is biased by the fact that the relation $\delta - \log N(\text{H}_i)$ is also $z$-dependent. Indeed, the same column density range selects objects with a lower density contrast at higher redshift explaining the lower
2.4. TPCF of the Lyα lines

Figure 2.8: Two point correlation function for the observed Lyα lines in the high-redshift range. Here the cut in column density is defined to correspond to a constant cut in density contrast, \( \delta \gtrsim 3 \), corresponding to \( \log N(H\text{I}) > 14.3 \) at \( z = 3.02 \), see text.

clustering signal. To verify this effect, we selected lines on the ground of a constant density contrast, \( \delta \gtrsim 3 \), which corresponds to \( \log N(H\text{I}) > 13.8 \) at the average redshift of the low redshift sub-sample, and to \( \log N(H\text{I}) > 14.3 \) at the higher average redshift. The correlation function for the latter sub-sample is shown in Fig. 2.8. Selecting the same kind of structures, there is no longer evidence of a significant evolution with redshift. Tab.2.2 shows a detailed budget of the number of lines used to compute the TPCF in all the different selections described before.
2.4. TPCF of the Lyα lines

Table 2.2: Detailed budget of the number of lines used to compute TPCFs. The first column refers to the selection done, in terms of emission redshift of the considered objects.

<table>
<thead>
<tr>
<th>Selection</th>
<th>N_{QSO}</th>
<th>N_{a}^{tot}</th>
<th>N_{b}^{mask}</th>
<th>N_{c}</th>
<th>N_{d}^{prox}</th>
<th>N_{e}^{comb}</th>
<th>N_{f}^{merg}</th>
<th>N_{g}^{col}</th>
<th>N_{h}^{fin}</th>
</tr>
</thead>
<tbody>
<tr>
<td>all QSOs</td>
<td>22</td>
<td>8435</td>
<td>368</td>
<td>644</td>
<td>1150</td>
<td>1955</td>
<td>380</td>
<td>4953</td>
<td>1147</td>
</tr>
</tbody>
</table>
| 13.8 < log N(H\text{I}) < 17  
\Delta r > 0.3 com. Mpc |         |             |              |      |              |              |              |             |             |
| all QSOs           | 22      | 8435        | 368          | 644  | 1150         | 1955         | 380          | 1319        | 4781        |
| 12 < log N(H\text{I}) < 13.8  
\Delta r > 0.3 com. Mpc |         |             |              |      |              |              |              |             |             |
| all QSOs           | 22      | 8435        | 368          | 644  | 1150         | 1955         | 380          | 170         | 5930        |
| 12 < log N(H\text{I}) < 17  
\Delta r > 0.3 com. Mpc |         |             |              |      |              |              |              |             |             |
| $z_{em} \leq 2.5$   
13.8 < log N(H\text{I}) < 17  
\Delta r > 0.3 com. Mpc | 11      | 3188        | 103          | 169  | 445          | 665          | 100          | 2042        | 381         |
| $z_{em} > 2.5$     
13.8 < log N(H\text{I}) < 17  
\Delta r > 0.3 com. Mpc | 11      | 5247        | 265          | 475  | 705          | 1290         | 280          | 2911        | 766         |
| $z_{em} > 2.5$     
14.3 < log N(H\text{I}) < 17  
\Delta r > 0.3 com. Mpc | 11      | 5247        | 265          | 475  | 705          | 1290         | 280          | 3345        | 332         |
| all QSOs           | 22      | 8435        | 368          | 644  | 1150         | 1955         | 2944         | 64          | 3472        |
| 12 < log N(H\text{I}) < 17  
\Delta r > 1L_{J} |         |             |              |      |              |              |              |             |             |

^a total number of fitted Lyα lines; ^b number of Lyα lines falling in the metal masks; ^c number of Lyα lines with $b < 10$ or $b > 100$; ^d number of lines falling closer than 1000 km s$^{-1}$ red-wards the Lyβemission or closer than 5000 km s$^{-1}$ blue-wards the Lyαemission; ^e number of eliminated lines because one of the three previous conditions occurs; ^f number of merged lines because their separation is less than the $\Delta r$ threshold indicated in the selection; ^g number of merged lines not fulfilling the column density selection; ^h number of lines used to compute the TPCF.
Traditionally, absorption spectra were decomposed into Voigt profiles which were then identified with individual discrete absorption systems. Information on the physical state of the gas responsible for the absorptions comes directly from the fit parameters: redshift, column density and Doppler parameter (linked to the temperature). In the new “cosmic web” view of the IGM the emphasis of the analysis has shifted to statistical measures of the transmitted flux (e.g. the flux power spectrum) more suitable for absorption arising from a continuous density field. However the interpretation of statistical quantities of the continuous flux field and their relation with the physical properties of the gas requires a non-trivial comparison with full hydro-dynamical high-resolution simulations that are computationally expensive. We present a new technique to reconstruct the hydrogen density field from the Lyα forests lines observed in high resolution spectra. The aim of this method is to extend the line fitting approach by identifying a new statistical estimator linked to the physical properties of the underlying IGM. This new estimator will also overcome the two main drawbacks of the Voigt fitting method:

(i) the subjectivity of the decomposition into components: the same absorption can be resolved by different scientists (or software tools) in different ways, both in the number of components, and in the values of the output parameters for a single component;

(ii) the blanketing effect of weak lines: they can be hidden by the stronger lines, so that their exact number density is unknown and has to be inferred from statistical arguments. Unfortunately, since the weak lines are also the most numerous, the uncertainty in their exact number is transformed into a systematic error of the computed statistical quantities.

The procedure to recover the H-density from the list of Lyα line column densities in a QSO line of sight, has been dubbed FLO (From Lines to Overdensities).
3.1 Introducing FLO

The starting point of FLO is the physical point made by Schaye (2001) that, in general, Ly\(\alpha\) absorbers will not be far from local hydrostatic equilibrium, i.e. their characteristic size will be typically of the order of the local \(L_J\). Making this hypothesis, assuming photo-ionisation equilibrium and introducing the equation of state of the IGM (Hui & Gnedin 1997), \(T = T_0 (\rho/\rho_0)^{\gamma-1}\) (where \(T_0\), the temperature at the mean density, and \(\gamma\) depend on the ionisation history of the Universe), the following relation linking the H\(\text{i}\) column density and the density contrast, \(\delta = n_\text{H}/\langle n_\text{H} \rangle - 1\), is obtained:

\[
N(\text{H}\text{i}) \simeq 3.7 \times 10^{13} \text{ cm}^{-2} (1 + \delta)^{1.5-0.26\alpha} T_{0.4}^{-0.26} \Gamma_{12}^{-1}
\times \left(\frac{1 + z}{4}\right)^{9/2} \left(\frac{\Omega_b h^2}{0.024}\right)^{3/2} \left(\frac{f_\gamma}{0.178}\right)^{1/2},
\]

where, \(\alpha \equiv \gamma - 1\), \(T_{0.4} \equiv T_0/10^4\), \(\Gamma_{12} \equiv \Gamma/10^{-12}\) is the H photo-ionisation rate and \(f_\gamma \approx \Omega_b/\Omega_m\) is the fraction of the mass in gas. The relation is valid for optically thin gas (\(\log N(\text{H}\text{i}) < 17\) cm\(^{-2}\)). In order to apply eq. 3.1.1 we have, first of all, to go through the Voigt fitting process of the Ly\(\alpha\) forest absorptions in a QSO spectrum. Then, to transform the list of H\(\text{i}\) column densities of Ly\(\alpha\) lines into the matter density field which generated them, we have to perform the following steps:

1. group Ly\(\alpha\) lines into absorbers of size of \(1\) \(L_J\) with column density equal to the sum of column densities and redshift equal to the weighted average of redshifts, using column densities as weights. We adopt the approximated formula for \(L_J\) (Zaroubi et al. 2006):

\[
L_J \simeq 1.33 \left(\frac{\Omega_m h^2}{0.135}\right)^{-1/2} \left(\frac{T_0}{1.8 \times 10^4}\right)^{1/2} \left(\frac{1.6}{\gamma}\right)^{1/2} \left(\frac{1 + z}{4}\right)^{-1/2} \text{ Mpc}
\]

in comoving units, where \(h \equiv H_0/100\) km s\(^{-1}\) Mpc\(^{-1}\), and the other parameters have already been defined. The absorbers are created with a friend-of-friend algorithm:

1. the spatial separation between all the possible line pairs is computed and the minimum separation is compared with \(L_J\), computed at the \(N(\text{H}\text{i})\)-weighted redshift mean of the pair;
2. if the two lines of the pair are more distant than the local \(L_J\), they are classified as two different absorbers, stored and deleted from the line list;
3. if the two lines are closer than the local \(L_J\), they are replaced in the line list by one line with a redshift equal to the \(N(\text{H}\text{i})\)-weighted mean of the two redshifts and a column density equal to the sum of the two column densities;
4. the procedure is iterated until all the lines are converted into absorbers.

2. transform the list of column densities of absorbers into a list of \(\delta\) with eq. 3.1.1;
(3) bin the redshift range covered by the studied Ly$\alpha$ forest; We binned the redshift range of the Ly$\alpha$ forest into steps of 1 $L_J$ and distributed the absorbers onto this grid, proportionally to the superposition between absorber size (which is again 1 $L_J$) and bin. Empty bins are filled with 1 absorber with hydrogen density contrast corresponding to the redshift of the bin and the minimum detectable column density in our data, log $N$(H$\text{i}$) = 12 cm$^{-2}$. The obtained field has been normalised in order to have have $\langle \delta+1 \rangle = 1.0$ for the whole considered sample. This operation is necessary to recover the correct asymptotic behaviour of the two point statistics of the field (see below).

With the introduction of this new statistical estimator we drastically reduce the drawbacks of the standard Voigt fitting approach. On the one hand, the weight of weak lines on statistical measures is cut down, since their contribution to the $\delta$ field is low. On the other hand, we verify that, in the process of Voigt fitting complex absorption features, the total H$\text{i}$ column density is a much more robust quantity than the number of components. To this purpose, we compare the results of line fitting for a sub-sample of 12 QSO spectra fit also with the software tool VPFIT and that were kindly provided to us by Tae-Sun Kim. The total number of lines in each line of sight is not conserved, in particular, significant differences are observed for the complex absorption systems where, in general, VPFIT fits more lines than FITLYMAN. Most of these discrepancies are due to the identification of low column density lines. However, the total column density in these complex absorbers appears to be much more stable between the two fitting methods. In Fig. 3.1, we plot the

![Figure 3.1: Comparison between the fitting results by FITLYMAN (solid line) and VPFIT (dotted line) for the QSO Q0109-3518. The lower panel shows the total number of lines, while the upper one shows the sum of the column densities of all the lines in redshift bins of width $\Delta z = 0.01$](image)

result of the comparison between FITLYMAN and VPFIT for one QSO of the sample, Q0109-
3518 ($z_m = 2.407$). We divide the line of sight into redshift bins of width $\Delta z = 0.01$; we sum both the number and the column densities of the lines in each bin, and plot them against redshift. It is evident that while the number of lines is different in every bin, the two column density distributions trace each other more faithfully. The two redshift intervals where the column density measured with FITLYMAN goes to zero are due to masked metal lines falling at those redshifts. VPFIT has been used to fit the lines of 3 QSOs in our sample and also to analyse the output spectra from the simulations (see next section). We verify the stability of FLO by applying it to the line lists obtained by VPFIT and FITLYMAN for the 12 common QSOs. In Fig 3.2, we show the comparison between the two recovered fields (1 L$_J$ binning) by means of a contour scatter plot. There is a tight correlation between the two values in the overdense regions, while the scatter increases for $\delta \lesssim 0$.

![Figure 3.2: Contour scatter plot of the FITLYMAN versus VPFIT reconstructed density fields. The contours show the number density of pixels which increases by a factor of 10 at each level.](image)

### 3.2 Testing FLO with Cosmological Simulations

To assess the ability of FLO to recover the H-density field, we used the simulated data sample described in 2.1. We applied FLO to the mock spectra obtained from the simulations and compared the resulting H-density field with the original one; results are described in the following sections.

#### 3.2.1 Reconstruction of the $\delta$ field

In this and in the next sections we refer to the $\delta$ reconstructed field as the one binned at 1 L$_J$ since we study the properties of the IGM at this length scale. The Ly$\alpha$ lines in
each simulated line of sight are selected to have, as in the case of observations, \( b \geq 10 \) km s\(^{-1}\). We introduce a further constraint, \( b \leq 100 \) km s\(^{-1}\), which is required by the fact that simulated spectra are not continuum fitted. Shallow and broad oscillations in the simulated spectra are fitted as absorption lines with Doppler parameters of the order of thousands of km s\(^{-1}\). In the real spectra these kind of oscillations are instead fitted with the continuum. The selected lines are grouped into absorbers and transformed into the corresponding density field following the procedure described in Section 3.1. The

![Figure 3.3: Contour scatter plot of the true versus reconstructed \( \delta \) field from simulations. The contours show the number density of pixels which increases by a factor of 10 at each level.](image)

reconstructed field is compared with the original density field (i.e. the output of the simulation), which is also binned into 1 \( L_1 \) steps. An upper threshold is adopted both for the true and and the recovered \( \delta \) field, \( \delta_{\text{thr}} = 50 \), since 99.95 percent of pixels in the simulated lines of sight have values \( \delta \leq \delta_{\text{thr}} \) and the algorithm to recover the \( \delta \) field (eq. 3.1.1) is valid for values \( \delta \leq \text{few} \times 10 \). The upper cut is applied before the normalisation process. The average values of the \( \delta \) field considering all the 364 simulated spectra are: \( \langle \delta + 1 \rangle \simeq 0.9 \) and 1.3 for the true and recovered field, respectively. The fields are normalised using these values. Figure 3.3 shows the contour scatter plot of the original versus reconstructed density field, the number of points within the different curves increases by a factor of 10 at each contour level. As can be seen from the figure, our method reconstructs fairly well the original field above the mean density, while in under-dense regions the recovered density is lower than the true density, if they are below our lower threshold. Indeed, the lower horizontal tail observed in the scatter plot is due to the treatment of the empty bins during the absorber-field transformation. The upper horizontal tail is instead due to the cut applied to over-densities larger than the maximum original value. Figure 3.4 shows the distribution of \( \delta \) densities in the true and recovered field. The peak at \( \log(\delta + 1) \simeq -0.83 \) contains \( \simeq 53 \) percent of the points and is due to the procedure that assigns to empty bins the value of \( \delta \) corresponding to the redshift of the bin and to the minimum observed column density \( \log N(\text{H}_1) = 12 \). On the other hand, the small bump at \( \log(\delta + 1) \simeq 1.57 \) is due to the upper cut applied to the recovered density field and it includes \( \sim 0.4 \) percent of the total number of points. The transformation
3.2. Testing FLO with Cosmological Simulations

Figure 3.4: Distribution of $\delta$ values in the true (solid line) and recovered (dashed line) field normalized to the total number of points.

recovers more than half of the points at $\log(\delta + 1) \simeq -0.15$ and recovers all the points within 30 percent in the range $-0.08 \lesssim \log(\delta + 1) \lesssim 1.45$. As it is clear, we are not dealing correctly with the under-dense regions, even if they are above our observational detection limit. This is likely due to the fact that our primary hypothesis, the local hydrostatic equilibrium, is not valid for these regions. This was also discussed by Schaye (2001) and here we have the evidence that under the mean density the gas is still expanding. In the next section, it will be shown that the fact of not reproducing correctly more than half of the points in the under-density regime does not significantly affect statistical measures like the correlation function.

3.2.2 Two point statistics of the $\delta$ field

We computed the correlation function for the original and recovered $\delta$ field, with the formula:

$$\xi_\delta(r) = \langle \delta(r + dr)\delta(r) \rangle,$$

(3.2.1)

where $r$ is the physical separation of two points in comoving Mpc. $\xi_\delta(r)$ quantifies the clustering properties of the considered field, showing a signal significantly different from zero at separations where the field presents structures (over or under-densities). We adopt as minimal separation for the computation of $\xi_\delta(r)$ the lowest value of $L_J$ for our sample, $\simeq 1.504$ comoving Mpc, corresponding to the maximum redshift, in order to account for all the bins. The bin size is instead the largest value of $L_J$ for our sample, $\simeq 1.532$.
Figure 3.5: Correlation function of $\delta$ from simulations. Crosses represent the correlation function obtained from the original data, triangles the one obtained from the reconstructed field. 

comoving Mpc, corresponding to the minimum redshift. Figure 3.5 plots the results of the correlation function for the true and recovered $\delta$ field. The value in each bin is the median value of 50 sample of 88 lines of sight obtained with a bootstrap technique from the 364 lines of sight of the total sample. This procedure is required in order to compare this result with the analogous one for the observed data (see Section 3.3.1.1). We have 22 observed spectra but each one covers a redshift range corresponding to about 4 simulated spectra. Error-bars are $1\sigma$ computed with the percentiles of the distribution of values in each bin. The recovered correlation function is in very good agreement with the true one at every separation. We also estimate the one-dimensional power spectrum of the hydrogen density field, that is defined by the Fourier transform of the correlation function:

$$\xi(r) = \frac{V}{2\pi} \int |\delta_k|^2 e^{-ikr} dk$$

$$P_{\delta}^{1D}(k) \equiv \langle |\delta_k|^2 \rangle$$

(3.2.2)

The power spectrum computation follows six simple steps:

1) a grid of wave-numbers is built in the Fourier space, starting from $k_{\text{min}} = 2\pi/\Delta r$, where $\Delta r$ is the length of a line of sight in comoving Mpc, and formed by $n_{\text{pix}}/2$ evenly spaced elements, where $n_{\text{pix}}$ is the number of pixels of the original $\delta$ field;

2) the Fourier transform of the $\delta$ field is computed;

3) the products $\delta_k\delta_{k'}$ are averaged in each bin;
3.3. FLO Applied to the observed data sample

3.3.1 Tracing the gas at redshifts $\sim 1.7-3.5$

The procedure explained in Section 3.1 is then applied to obtain the corresponding density contrast field for each line of sight. In the case of observations, we have to take into account the presence of the masked intervals covering regions occupied by metal absorption systems. We eliminate the pixels that fall by more than 30 percent into a masked interval. Before the normalisation step, we apply an upper threshold as for simulations ($\delta_{\text{thr}} = 50$) since we want to compare our result with the one obtained in Section 4.2. The pixels above the threshold correspond to $\sim 0.9$ percent of the total number of pixels. Fig. 3.2 and 3.3 show the errors associated with the change in fitting tool (we are using VPFIT...
Figure 3.7: Distribution with redshift of the average normalised, reconstructed $\delta$ values for 4 sub-samples of QSOs selected by their emission redshift as displayed in the plot. Horizontal lines represent the redshift coverage of each QSO sample, while vertical lines are the spread in average $\delta$ values for the single QSO in the samples.
for the simulated data sample and FITLYMAN for the observed one) and the ones intrinsic to FLO, respectively. Since the intrinsic errors turn out to be larger than those induced by the fitting technique, we can safely compare the results obtained from the simulations and those obtained from the data sample, presented in this section. Figure 3.7 shows the average values of the recovered $\delta$ fields for 4 sub-samples built from the 22 observed lines of sight selected on the ground of the QSO emission redshifts. The spread of average $\delta$ values for the single QSO forming the samples is also shown. There is no significant trend with redshift, as it is expected if the density field follows on average the evolution of the cosmic mean value.

### 3.3.1.1 Two point statistics of the $\delta$ field

The correlation function for the observed $\delta$ field is computed with the same formula given for simulations. The result is shown in Fig. 3.8. Here the value in each bin is obtained averaging all the sample, while the error bars are computed creating 50 samples of 22 lines of sight drawn out of our sample with a bootstrap technique and taking the percentiles of the distribution corresponding to $1\sigma$ errors. The smaller separation is $\simeq 1.245$ comoving Mpc, corresponding to the $L_J$ at the largest redshift of the sample, while the bin size is $\simeq 1.628$ comoving Mpc which is the $L_J$ at the lowest redshift of the sample. The clustering signal in the first two bins is significant at the $3\sigma$ level. We have superimposed to the data points the $\delta$-TPCF obtained from simulations. We notice that the $\delta$-TPCF of the

![Figure 3.8: Correlation function of the $\delta$ field reconstructed from our 22 observed QSO spectra. Points refer to the data, the line instead represent the prediction from the simulation.](image-url)
3.3. FLO Applied to the observed data sample

Observational data as a whole is in very good agreement with the simulated one. Since

\[ P_{\delta}^{1D}(k) \] is very sensitive to cosmological parameters, it is very important to check if the prediction of such a function are in agreement with the observed values. In the case of the observed spectra, the masked metal lines make the starting grid of pixels unevenly spaced, thus not fitted for the application of the FFT. To overcome this problem, as a 1st order approximation, the masked bins have been put to the average density. This procedure is based on the observation that the Ly\( \alpha \) forest gas traces on average the average density and on the analogous method adopted in the computation of the power spectrum of the transmitted flux (Viel et al. 2004b). Fig. 3.9 shows the result of this computation: the power spectrum obtained from our data is in excellent agreement with the one obtained from the density fields reconstructed with FLO from the simulated spectra based on a concordance cosmological model. Error bars are computed using a jackknife estimator on the whole sample of observed QSOs.

3.3.2 The distribution of matter close to QSO from the proximity effect

In the standard analysis of the proximity effect it is assumed the matter distribution is not altered by the presence of the QSO. The only difference between the gas close and far away from the QSO is the increased photoionization rate due to the QSO emission. A consequence of this hypothesis is that there should be correlation between the strenght of the proximity effect and the luminosity of the QSO. However, observational results are not conclusive on this subject (Lu et al. 1991; Bechtold 1994; Srianand & Khare 1996; Liske & Williger 2001). It is infact likely, that QSOs occupy overdense regions. Hierarchical models of structure formation predict that super-massive black holes, that are thought to power QSOs, are in massive halos which are strongly biased to high-density regions. In this paragraph we investigate the density distribution of matter close to QSOs from the proximity effect using FLO. Similar analysis have been carried out in previous works using

\[ \Omega_m=0.74 \]
\[ \Omega_{\Lambda}=0.26 \]
\[ \sigma_8=0.85 \]
3.3. FLO Applied to the observed data sample

A different approach, the determination of the cumulative probability distribution function of pixel optical depth. Rollinde et al. (2005) have studied the density structure around 12 QSOs belonging also to our sample. These authors marginally detect the presence of an overdensity for $2 \lesssim r \lesssim 10 \ h^{-1} \ Mpc$ proper, which is consistent with the density profile around the most massive halo at $z = 2$ in the Millennium simulation (Springel 2005), assuming an hydrogen ionization rate $\Gamma = 10^{-12}s^{-1}$. Guimarães et al. (2007) have investigated the distribution of matter density close to 45 high-redshift ($z_{em} \sim 4$) QSOs observed at medium spectral resolution. Their study reveals gaseous overdensities for brighter QSOs.

3.3.3 QSO emission redshift determination

Emission redshift of QSOs at $z_{em} > 1.5$ are generally computed by the positions of the most prominent UV emission lines, in particular H I Ly$\alpha$ and C IV. However, it was assessed by several studies (e.g. Gaskell 1982; Wilkes 1986; Espey et al. 1989; Corbin 1990; Tytler & Fan 1992; Laor et al. 1995; McIntosh et al. 1999; Vanden Berk et al. 2001) that high ionization emission lines C IV, N V, C III], and H I Ly$\alpha$ are on average blue-shifted by several hundreds of km s$^{-1}$ with respect to low ionization lines O I and Mg II and Balmer lines. On the other hand, Balmer lines and narrow forbidden lines usually give redshifts within 100 km s$^{-1}$, and possibly much less, of the systemic stellar and interstellar medium redshifts (Tytler & Fan 1992, and references therein).

The knowledge of the correct systemic redshift of the QSO is of fundamental importance when using the proximity effect to estimate both the intensity of the UV ionising background and the density structure close to the QSO itself. To this purpose, we carried out a detailed analysis of the emission line redshifts of the QSOs in our sample both using data from the literature and directly fitting the lines in the UVES spectra. In decreasing order of precision, we adopted the redshifts measured by: i) narrow forbidden lines, mainly [OIII] $\lambda 5007$ A, ii) H$\beta$, iii) Mg II, and iv) O I. There are 5 QSOs in the sample for which none of these lines was measured. We excluded those objects from the proximity study, cutting the portion of the spectrum within 5000 km s$^{-1}$ of the best estimate of the emission redshift (generally obtained from Ly$\alpha$ and/or C IV emission lines).

Table 3.1 gives the QSO redshifts and the details of the estimate. The determinations from the UVES spectra were obtained by rebinning the region of the emission, normalising it to the local continuum and fitting a Gaussian profile to the line.

3.3.3.1 QSO luminosity determination

Getting closer and closer to the QSO, the UV ionising field becomes dominated by the intrinsic QSO emission flux. In order to derive the matter density distribution around the QSO using the observed variation of the absorption features in the QSO spectrum, a reliable determination of the intrinsic luminosity of the QSO has to be obtained. Magnitudes
of the objects in our sample have been taken from the GSC-II catalogue (McLean et al. 2000) and reported in Tab. 3.1. In order to estimate the corresponding intrinsic bolometric luminosity, we have adopted the following procedure. The QSO template library defined in Fontanot et al. (2007) has been considered, which is based on high quality SDSS QSO spectra in the redshift interval $2.2 < z < 2.25$. This redshift range was chosen in order to maximize the level of completeness of the sample and the wavelength interval longwards of the Ly$\alpha$ emission. Moreover, in this redshift range, the dynamical response of the SDSS spectrograph is such that the Ly$\alpha$ lines is completely sampled in all spectra. In the original paper, the authors considered the rest frame spectra of the 215 QSOs forming the final sample and they used a continuum fitting technique in order to extend the information blueward of the Ly$\alpha$. A mean continuum slope $\gamma = 0.7 \pm 0.3$ was obtained for the objects in the library, where $f_\nu \propto \nu^{-\gamma}$ and $f_\nu$ is the quasar flux in units of ergs cm$^{-1}$ s$^{-1}$ Hz$^{-1}$. Fontanot et al. (2007) demonstrated that this library is suitable for predicting QSOs colors up to $z \lesssim 5.2$. For the purposes of this work we considered the continuum fitted slopes for $\lambda > 500 \angstrom$. While at shorter wavelengths, a fixed slope $\gamma = -1.8$ (Madau et al. 1999) has been imposed. We used the template spectra in the library to compute a synthetic $b_J$ and $r_F$ magnitude at each emission redshift listed in tab. 3.1. For reproducing the $b_J$ and $r_F$ photographic magnitudes, the spectral response of the PEGASE code (Fioc & Rocca-Volmerange 1997) was assumed. Then, the templates have been renormalized in each band separately, by requiring the synthetic magnitude to match the observed one, and renormalized spectra have been used to give a prediction for the AB magnitudes at 912 $\angstrom$ ($M_{bJ}^{912}$ and $M_{rF}^{912}$ respectively). The quantity $\Delta M_{912} = M_{bJ}^{912} - M_{rF}^{912}$ has been adopted as an estimator of the agreement between the slope of the template and the intrinsic slope of the considered QSO: we then associate to each observed QSO the template with the smaller $\Delta M_{912}$. For 18 out of 22 QSOs in our sample this procedure gives $\Delta M_{912}$ values lower than 0.001 mag. For the remaining 4 objects the values are respectively 0.1 (PKS0329-255), 0.2 (HE1347-2457 and B1422+231) and 0.4 (HE1341-1020). In these cases there is no template in the library, which reproduces the correct intrinsic slope of the observed QSOs, and $\Delta M_{912}$ could be taken as a measure of the systematic error for the 4 objects, when we consider the luminosity corresponding to the template with the smaller $\Delta M_{912}$. We check that our conclusions do not change if we exclude these QSOs from the main sample. We then use the selected and renormalized template spectra to estimate the monochromatic luminosities $L_\nu$ (ergs s$^{-1}$ Hz$^{-1}$)(Bechtold 1994). Integrating $L_\nu$ between $300 < \lambda < 912\angstrom$ we compute the quantity $\Gamma_\nu$ (s$^{-1}$ cm$^2$):

$$\Gamma_\nu = \int \frac{L_\nu}{h\nu} \sigma_\nu \, d\nu \quad (3.3.1)$$

where

$$\sigma_\nu = 6.30 \times 10^{-18} \text{cm}^{-2} \left(1.34 \left(\frac{\nu}{\nu_{\text{HI}}}\right)^{-2.99} - 0.34 \left(\frac{\nu}{\nu_{\text{HI}}}\right)^{-3.99}\right) \quad (3.3.2)$$

is the absorbing cross-section for neutral hydrogen (Osterbrock 1989; Bolton & Haehnelt 2007). We relate $\Gamma_\nu$ to the hydrogen photoionization rate of the quasar through the formula:

$$\Gamma_{\text{QSO}} [\text{s}^{-1}] = \frac{1}{4\pi r_{\text{abs}}^2} \Gamma_\nu \quad (3.3.3)$$

where $r_{\text{abs}}$ is the distance of the absorber from the quasar expressed in cm.
### Table 3.1: Relevant properties of the QSOs forming our sample. See text for further details.

<table>
<thead>
<tr>
<th>QSO</th>
<th>$z_{em}$</th>
<th>line</th>
<th>Ref.</th>
<th>$\Delta z_{\text{Ly}a}$</th>
<th>$b_J$</th>
<th>$r_F$</th>
<th>$L_{LL} \times 10^{41}$</th>
<th>$\Gamma \times 10^{40}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE1341-1020a</td>
<td>2.142</td>
<td>Mg II</td>
<td>1</td>
<td>1.6599-2.142</td>
<td>18.68</td>
<td>17.52</td>
<td>0.51312</td>
<td>0.11030</td>
</tr>
<tr>
<td>Q0122-380</td>
<td>2.2004</td>
<td>Hβ</td>
<td>2</td>
<td>1.709-2.2004</td>
<td>17.34</td>
<td>16.70</td>
<td>3.0814</td>
<td>0.77208</td>
</tr>
<tr>
<td>PKS1448-232</td>
<td>2.224</td>
<td>Mg II</td>
<td>1</td>
<td>1.729-2.224</td>
<td>17.09</td>
<td>16.87</td>
<td>3.6234</td>
<td>0.97592</td>
</tr>
<tr>
<td>PKS0237-23</td>
<td>2.233</td>
<td>Mg II</td>
<td>1</td>
<td>1.737-2.233</td>
<td>16.61</td>
<td>16.21</td>
<td>5.1119</td>
<td>1.2824</td>
</tr>
<tr>
<td>J2233-606</td>
<td>2.248</td>
<td>O I</td>
<td>1</td>
<td>1.7496-2.248</td>
<td>16.97</td>
<td>17.01</td>
<td>4.6694</td>
<td>1.3706</td>
</tr>
<tr>
<td>HE0001-2340</td>
<td>2.265</td>
<td>Mg II</td>
<td>1</td>
<td>1.764-2.265</td>
<td>16.74</td>
<td>16.46</td>
<td>4.5967</td>
<td>1.1794</td>
</tr>
<tr>
<td>HE1122-1648b</td>
<td>2.40</td>
<td></td>
<td>3</td>
<td>1.878-2.344</td>
<td>16.61</td>
<td>16.32</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HE2217-2818b</td>
<td>2.412</td>
<td></td>
<td>1</td>
<td>1.888-2.355</td>
<td>16.47</td>
<td>16.16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Q0329-385</td>
<td>2.437</td>
<td>Mg II</td>
<td>5</td>
<td>1.9096-2.437</td>
<td>17.20</td>
<td>16.91</td>
<td>5.2736</td>
<td>1.5572</td>
</tr>
<tr>
<td>HE1158-1843a,b</td>
<td>2.448</td>
<td></td>
<td>1</td>
<td>1.919-2.391</td>
<td>17.09</td>
<td>16.84</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HE1347-2457a</td>
<td>2.5986</td>
<td>Hβ</td>
<td>4</td>
<td>2.046-2.5986</td>
<td>17.35</td>
<td>16.14</td>
<td>3.4681</td>
<td>0.74552</td>
</tr>
<tr>
<td>Q0453-423</td>
<td>2.669</td>
<td>O I</td>
<td>1</td>
<td>2.106-2.669</td>
<td>17.69</td>
<td>16.74</td>
<td>2.9047</td>
<td>0.65574</td>
</tr>
<tr>
<td>PKS0329-255b</td>
<td>2.696</td>
<td></td>
<td>1</td>
<td>2.129-2.635</td>
<td>17.88</td>
<td>17.71</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HE0151-4326b</td>
<td>2.763</td>
<td></td>
<td>1</td>
<td>2.186-2.701</td>
<td>17.48</td>
<td>16.93</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Q0002-422</td>
<td>2.769</td>
<td>O I</td>
<td>1</td>
<td>2.191-2.769</td>
<td>17.50</td>
<td>16.89</td>
<td>5.5006</td>
<td>1.5077</td>
</tr>
<tr>
<td>HE2347-4342a</td>
<td>2.880</td>
<td>O I</td>
<td>1</td>
<td>2.285-2.880</td>
<td>17.12</td>
<td>16.30</td>
<td>7.9118</td>
<td>1.9612</td>
</tr>
<tr>
<td>HS1946+7658</td>
<td>3.058</td>
<td>O I</td>
<td>6</td>
<td>2.435-3.058</td>
<td>16.64</td>
<td>15.76</td>
<td>17.283</td>
<td>4.5913</td>
</tr>
<tr>
<td>HE0940-1050</td>
<td>3.0932</td>
<td>Hβ</td>
<td>4</td>
<td>2.465-3.0932</td>
<td>17.08</td>
<td>16.08</td>
<td>12.845</td>
<td>3.4360</td>
</tr>
<tr>
<td>Q0420-388a</td>
<td>3.1257</td>
<td>O I</td>
<td>1</td>
<td>2.493-3.1257</td>
<td>17.44</td>
<td>16.70</td>
<td>8.7491</td>
<td>2.5681</td>
</tr>
</tbody>
</table>

---

* QSOs with associated absorption systems; b QSO not considered in the proximity study.

References: (1) this paper; (2) Sulentic et al. 2004; (3) Kim et al. 2002; (4) P. Marziani, private comm.; (5) Espey et al. 1989; (6) Fan & Tytler 1994; (7) Scott et al. 2000.
3.3. FLO Applied to the observed data sample

3.3.3.2 The proximity effect

The proximity effect was initially called the inverse effect based on the observation that the number of Ly\(\alpha\) absorption lines was decreasing with respect to the average number density approaching the QSO emission redshift.

As a first test, we wanted to highlight this effect also with the FLO approach. Equation 3.1.1 has been applied to the Ly\(\alpha\) absorbers built from the 22 line lists in our sample, deliberately neglecting the contribution of the QSO to the UV ionising flux. The resulting 22 \(\delta\) fields with pixels of \(1 \, L_\odot\) in redshift space have been re-binned into bins of 3 proper Mpc in length in the region within 10 proper Mpc from the QSO emission redshift and into bins of 5 proper Mpc in length in the outer region. Those bins that are covered by more than 33 percent by masked intervals are eliminated from the final count for the single object. Then in each bin, we have taken the mean of all the \(\delta\) values contributing to that bin (one per QSO at maximum). The resulting \(\delta\) fields are shown in Fig. 3.10 together with the number of QSOs contributing to each bin. The dashed lines in the upper panel represent the average of all the points at separations larger than 30 Mpc, that are likely no longer affected by the ionising flux from the QSO. The dotted lines are the \(1\sigma\) standard deviations.

![Figure 3.10: Proximity effect observed in the mean \(\delta\) fields for the 22 QSOs in our sample neglecting the UV flux from the QSO. The lower panel shows the number of objects contributing to each bin.](image-url)
3.3.3.3 The density structure around QSOs

As a first approximation we have not taken into account the effect of the QSO ionizing flux on their environments. In order to study the characteristics of the over density hosting the QSOs, we used in eq. 3.1.1 a value of $\Gamma_{12}$ corrected for the QSO contribution:

$$\Gamma_{12} = \frac{\Gamma_{UVB} + \Gamma_{QSO}}{10^{-12}} = \frac{\Gamma_{UVB}}{10^{-12}}(1 + \frac{\Gamma_{QSO}}{\Gamma_{UVB}})$$ (3.3.4)

where $\Gamma_{QSO}$ is defined in eq. 3.3.3. In fig. 3.11 we present, for the 22 QSO in the sample the comparison between the median recovered density field with and without the contribution of the QSO ionizing flux.

Figure 3.11: Mean $\delta$ field obtained from the 22 QSOs spectra in our sample as a function of the proper separation from the QSO when the QSO ionizing flux is taken into account. Red stars show the result without the ionizing contribution from QSO.

3.4 Future Perspectives

FLO is a powerful tool to study the characteristics of the IGM and to compare observations with numerical simulations. In principle it could be a powerful tool to place constraints on cosmological parameters. To constrain cosmological parameters with FLO we have to deeply understand the inaccuracies of the method, in particular the problems with underdensities and big overdensities, that forced us to introduce the two cuts in $\delta$ (see
fig. 3.3); solving this could help us to understand the bias in the one dimensional power spectrum between the original simulated field and the reconstructed one (see fig. 3.6). A good comprehension of the one dimensional power spectrum is crucial to go to the 3D power spectrum, that would allow us to put constrains on cosmological parameters through FLO.
Multiple LOSs offer an invaluable alternative to address the spatial distribution of the absorbers, enabling a more direct interpretation of the observed correlations. Besides, the cross-correlation of QSO spectra is much less affected by errors uncorrelated in different LOSs, like those introduced by continuum fitting. The interesting range of separations lies at about the Jeans scale of the photoionised IGM ($\sim 1$ arcmin or $\sim 1.4 h_{70}^{-1}$ comoving Mpc at $z \sim 2$). At this scale there should be a transition from a smooth gas distribution, which produces nearly identical absorption features in neighbouring LOSs, to a correlated density distribution, where the correlation strength decreases with increasing separation of the LOSs (Viel et al. 2002). Common absorption features in the spectra of multiply lensed quasars (Smette et al. 1992, 1995) indicate that the absorptions are produced in structures exceeding a few tens of kpc. Studies of close quasar pairs provide evidence for dimensions of a few hundred kpc (D’Odorico et al. 1998; Aracil et al. 2002). Recently, lensed and more widely separated QSO pairs have been used to recover the kinematics of the gaseous cosmic web (Rauch et al. 2005), confirming that the Hubble expansion and gravitational instability are the main processes influencing the Ly$\alpha$ forest gas.

We have used a unique sample of 15 UVES spectra forming 21 QSO pairs with angular separations evenly distributed between of $\sim 1$ and 14 arcmin (3 pairs at $\sim 1$ arcmin), to compute the cross-correlation function of the transmitted flux in the Ly$\alpha$ forest at $z \sim 2$. The following section will describe in detail this analysis published in D’Odorico et al. (2006). My own contribution to this project was to collaborate in reducing and analyzing the data.

4.1 Data Sample

The exploitation of the potential offered by multiple QSO lines of sight has been limited by the dearth of suitable groups of QSOs close and bright enough to permit high resolution
Figure 4.1: Relative positions, U magnitudes and redshifts of the QSOs composing the sestet (left plot) and the triplet (right plot) in our sample
4.1. Data Sample

<table>
<thead>
<tr>
<th>Object</th>
<th>z</th>
<th>$M_B$</th>
<th>$\text{Ly}\alpha$ range</th>
<th>S/N per pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair A</td>
<td>PA1</td>
<td>2.645</td>
<td>19.11</td>
<td>2.094–2.585</td>
</tr>
<tr>
<td></td>
<td>PA2</td>
<td>2.610</td>
<td>19.84</td>
<td>2.094–2.550</td>
</tr>
<tr>
<td>Triplet</td>
<td>T1</td>
<td>2.041</td>
<td>18.20</td>
<td>1.633–1.991</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>2.05</td>
<td>18.30</td>
<td>1.592–1.999</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>2.053</td>
<td>18.10</td>
<td>1.665–2.002</td>
</tr>
<tr>
<td>Sestet</td>
<td>S1</td>
<td>1.907</td>
<td>19.66</td>
<td>1.665–1.859</td>
</tr>
<tr>
<td></td>
<td>S2</td>
<td>2.387</td>
<td>19.53</td>
<td>1.858-2.331</td>
</tr>
<tr>
<td></td>
<td>S3</td>
<td>2.102</td>
<td>19.31</td>
<td>1.633–2.051</td>
</tr>
<tr>
<td></td>
<td>S4</td>
<td>1.849</td>
<td>19.59</td>
<td>1.575–1.802</td>
</tr>
<tr>
<td></td>
<td>S5</td>
<td>2.121</td>
<td>18.85</td>
<td>1.633–2.069</td>
</tr>
<tr>
<td></td>
<td>S6</td>
<td>2.068</td>
<td>20.19</td>
<td>1.592–2.017</td>
</tr>
<tr>
<td>Pair U</td>
<td>UM680</td>
<td>2.144</td>
<td>18.60</td>
<td>1.653–2.092</td>
</tr>
<tr>
<td></td>
<td>UM681</td>
<td>2.122</td>
<td>19.10</td>
<td>1.634–2.070</td>
</tr>
<tr>
<td>Pair Q</td>
<td>Q2343+12</td>
<td>2.549</td>
<td>17.00</td>
<td>1.994–2.490</td>
</tr>
<tr>
<td></td>
<td>Q2344+12</td>
<td>2.773</td>
<td>17.50</td>
<td>2.183–2.711</td>
</tr>
</tbody>
</table>

Table 4.1: Characteristics of the observed QSO spectra

spectroscopy. Two major breakthroughs have dramatically improved this situation:

- the 2dF QSO Redshift Survey (Croom et al. 2004, www.2dfquasar.org), whose complete spectroscopic catalogue contains more than $\sim 23000$ QSOs in a single homogeneous data base, which is approximately 50 times more than the previous largest QSO survey to a similar depth ($B < 21$);

- the SDSS QSO catalogue (Schneider et al. 2005); this catalogue consists of 46420 objects with $0.078 < z < 5.414$, $15.1 < i < 21.78$ and $-30.2 < M_i < -22.0$.

- the UVES spectrograph (Dekker et al. 2000) at the Kueyen unit of the ESO VLT (Cerro Paranal, Chile) which has a remarkable efficiency especially in the extreme UV (close to the atmospheric cut-off). In this way, relatively low-redshift ($z \lesssim 2.5$) QSOs become accessible to high resolution observations of the Ly\alpha forest down to faint apparent magnitudes and their surface density becomes high enough to provide several lines of sight for a tomography of the IGM.

We searched the 2dF QSO database for the best groups with apparent magnitude $B \leq 20$ and $z > 1.8$. A triplet found in the Calán-Tololo QSO survey (Maza et al. 1993, 1995) was added to the sample. A great observational effort was carried out to collect UVES spectra of the selected QSO, which all have magnitudes fainter than $B \sim 19$, with the exception of the triplet for which $B \sim 18$. ESO allocated 51 hours of observation to our project up to now, which allowed us to obtain acceptable S/N spectra of one pair (from now on called Pair A, with angular separation of $\sim 1.3\,''$), one sestet and the triplet of QSOs (see Fig. 4.1). Reduction of the QSO spectra was conducted with the pipeline of the instrument (version 2.1, Ballester et al. 2000) provided by ESO in the context of the data
reduction package MIDAS. In most cases, we could not apply the standard procedure due to the faintness of the spectra in the blue region and had to pre-filter the cosmic rays to make the optimal extraction work properly. Single extracted spectra were summed and rebinned and wavelengths were corrected to the vacuum-heliocentric reference frame. The final spectra have resolution $R \sim 40000$ in the Ly$\alpha$ and C$\text{IV}$ forest, while the S/N per pixel varies on average between $S/N \sim 3$ and 10 in the Ly$\alpha$ forest and between 4 and 15 in the C$\text{IV}$ forest (see Table 4.1 for details). Continuum determination, in particular in the Ly$\alpha$ forest region, is a very delicate step in the process of spectra reduction. Tentative procedures realised up to now to objectively determine the continuum position through automatic algorithms do not give satisfactory results. We adopted a manual subjective method based on the selection of the regions free from clear absorption that are successively fitted with a spline polynomial of 3rd degree. The limitations introduced by the uncertainty in the true continuum level and shape should be less important in the computation of the cross-correlation function than in the case of single line of sight analysis. This because the undulations are uncorrelated between adjacent lines of sight, in particular, if the two QSOs do not have too similar redshifts (see Viel et al. 2002, for a discussion in the case of power spectra). We added to the sample two more pairs from our archive: UM680/UM681 (Pair U, separated by $\sim 1$ arcmin) and Q2344+1228/Q2343+1232 (Pair Q, separated by $\sim 5.57$ arcmin). They were observed with UVES at the same resolution as the new data, and reduced following the same procedure. More details on their properties are given in the paper by D’Odorico et al. (2002). The list of the QSOs and the main characteristics of the spectra are reported in Table 4.1. The Ly$\alpha$ forests of two of the closest pairs in our sample are shown in Fig. 4.2 The total sample is formed by 21 pairs uniformly distributed between angular separations of $\sim 1$ and 14 arcmin, corresponding to comoving spatial separations between $\sim 1.4$ and $21.6 \, h^{-1}$ Mpc. The median redshift of the Ly$\alpha$ forest is $z \sim 1.8$. This is the largest sample of high resolution spectra of QSO pairs ever collected, unique both for the number density – we have 6 QSOs in a region of $\sim 0.04$ square degrees – and the variety of line of sight separations investigated. As a

<table>
<thead>
<tr>
<th>Object</th>
<th>Redshift</th>
<th>$M_B$</th>
<th>Ly$\alpha$ range</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE 1341-1020</td>
<td>2.134</td>
<td>17.1</td>
<td>1.64431–2.08216</td>
</tr>
<tr>
<td>Q 0122-380</td>
<td>2.189</td>
<td>16.7</td>
<td>1.69072–2.13625</td>
</tr>
<tr>
<td>PKS 1448-232</td>
<td>2.220</td>
<td>16.96</td>
<td>1.71688–2.16674</td>
</tr>
<tr>
<td>HE 0001-2340</td>
<td>2.280</td>
<td>16.7</td>
<td>1.76750–2.22575</td>
</tr>
<tr>
<td>Q 0109-3518</td>
<td>2.406</td>
<td>16.44</td>
<td>1.87381–2.34966</td>
</tr>
<tr>
<td>HE 2217-2818</td>
<td>2.406</td>
<td>16.0</td>
<td>1.87381–2.34966</td>
</tr>
<tr>
<td>Q 0329-385</td>
<td>2.423</td>
<td>16.92</td>
<td>1.88816–2.36638</td>
</tr>
<tr>
<td>HE 1158-1843</td>
<td>2.453</td>
<td>16.93</td>
<td>1.91347–2.39589</td>
</tr>
</tbody>
</table>

Table 4.2: List of the considered QSOs from the ESO Large Program (see text)
Figure 4.2: Lyα-forests of two of the closest pairs in our sample. Pair A (upper two panels) separated by 1.3′ and the pair at 1.06′ formed by the QSOs S5 and S6 of the sestet (lower two panels)
4.1.1 The simulated spectra

In order to both assess the nature of the Ly\(\alpha\) forest inferred from simulations and to constrain the cosmological scenario of the same simulations, we compared the results obtained for our sample of observed QSO spectra with analogous results for a sample of mock Ly\(\alpha\) forests. The simulations adopted are the same of Chap. 3. The \(z = 1.8\) output of the simulated box was pierced to obtain 50 triplets of lines of sight carefully reproducing the observed triplet mutual separations. The same was done for 50 sestets of lines of sight reproducing the observed sestet and 50 pairs of lines of sight at the same angular separation as Pair U. Fifty different realisations of Pair A spectra and 50 of Pair Q were obtained from the output box at redshift \(z = 2.4\).

4.2 The flux auto-correlation function

The unnormalised auto-correlation function of the flux along the line of sight is defined as:

\[
\xi_{f \parallel}^f(\Delta v_{\parallel}) = \langle \delta_f(v_{\parallel}, \theta)\delta_f(v_{\parallel} + \Delta v_{\parallel}, \theta) \rangle,
\]

following previous studies on the same subject (e.g. McDonald et al. 2000; Rollinde et al. 2003; Becker et al. 2004). The auto-correlation function for our sample of QSO spectra was obtained by averaging over all the pixels of all the QSOs. The results were binned in 50 \(\text{km s}^{-1}\) velocity bins. The 1 \(\sigma\) error on this measure is estimated extracting 50 samples of 15 objects from our sample with a bootstrap method, computing the auto-correlation function for each sample and determining the standard deviation of the distribution. The obtained correlation function is in very good agreement with the analogous determination by Rollinde et al. (2003) and in qualitative agreement with the Ly\(\alpha\) lines two point correlation function at similar redshifts (Kim et al. 2001). The same procedure was adopted to compute \(\xi_{f \parallel}^f\) for the QSO spectra of the control sample taken from the UVES LP. These spectra have a much larger S/N (\(\sim 50\) in the Ly\(\alpha\) forest region) than our spectra, they have been selected to be free from damped Ly\(\alpha\) absorptions and they do not belong to known close QSO groups. The auto-correlation function for the simulated spectra was computed as the arithmetic mean of the correlation functions obtained for 50 realisations of the observed sample and the error is the corresponding standard deviation. It is important to recognise that the computed error bars both for the observed and simulated \(\xi_{f \parallel}^f\) are strongly correlated. This is due to the fact that every pixel contributes to the correlation function in several velocity bins. In Fig. 4.3 we show the comparison of the three auto-correlation functions in velocity space, error bars are included just to give a qualitative idea of their size. The agreement at the 3 \(\sigma\) level is good. However, we would like to point out two kinds of discrepancies at the 1 \(\sigma\) level which, in our opinion, are the signatures of two interesting effects:

1. Both observed correlation functions have an amplitude in the first bin (centred at 50 \(\text{km s}^{-1}\)) which is larger than the simulated one. We ascribe this to a scantiness of strong
4.2. The flux auto-correlation function

Figure 4.3: Comparison of the auto-correlation function for our sample of observed QSO spectra (solid triangles), for the LP sample of observed QSO spectra (crosses, shifted in velocity for clarity) and for the simulated sample of spectra (empty triangles) as a function of velocity separation along the line of sight. Error bars on $\xi_f$ are highly correlated and they are plotted only to give an idea of their extent.

lines ($\log N(\text{HI}) \gtrsim 15.5$) in the simulated spectra (see e.g. Theuns et al. 2002) due to the limited extent of simulations in comoving space, which implies that large over densities (giving rise to strong Ly$\alpha$ lines) are not appropriately probed.

2. The correlation function computed with our sample of QSOs tends to show a stronger clustering than the simulated one up to large separations, while the LP correlation function is compatible with the mock correlation function already at the second bin. The main difference between the two samples is the fact that our QSOs were selected to belong to close pairs and associations, thus they could trace over-clustered regions. Although we excluded the spectral portion within 5000 km s$^{-1}$ of the QSO emission, an enhanced clustering signal could still be present. We notice that cosmic variance could also be at the origin of these small differences. Indeed auto-correlation functions of single observed QSO spectra show large variations among them and we are working with relatively small samples. On the other hand, all the simulated spectra have been drawn from the same box of 120 $h^{-1}$ comoving Mpc, which although notably large for a simulation, possibly underestimates the cosmic variance (see McDonald et al. 2000, for possible ways of correcting the simulated correlation function for errors induced by the finite box size).
4.2.1 Contribution of the stronger absorptions

Lyα lines of increasing H1 column density correspond to structures of increasing overdensity and smaller characteristic size. Going from log \(N(\text{H}1) \sim 12\) to log \(N(\text{H}1) \gtrsim 15\) we probe from under-dense, modestly over-dense regions, to highly over-dense filaments and their intersections. We expect that the major contributors to the clustering signal are the stronger absorption lines, which trace the higher density peaks. The aim of this section is twofold. On the one hand, by excluding the stronger lines we want to verify the clustering properties of the gas in the ‘true’ IGM, far from the densest regions. On the other hand, we want to test the relative behaviour of the observed and simulated auto-correlation functions without those lines. Note that a careful study of the impact of strong absorption systems on the flux power spectrum was performed by Viel et al. (2004a) and McDonald et al. (2005). To select the strong absorption lines without carrying on the time consuming Voigt fitting procedure of the observed and simulated spectra, we can relate the transmitted flux values to a corresponding column density assuming a typical Doppler parameter. We eliminated the pixels with \(f < f_{\text{thres}} = 0.1\) or \(\tau_{\text{H}1} > 2.3\). This approximately corresponds to excluding the absorption lines with log \(N(\text{H}1) > 13.96\), if we adopt the formula for the optical depth at the centre of the line:

\[
\tau_{0,\text{Ly} - \alpha} = 7.58295 \times 10^{-13} \frac{N \text{ (cm}^{-2}\text{)}}{b \text{ (km s}^{-1}\text{)}} ,
\]

and a Doppler parameter \(b = 30\) km s\(^{-1}\) characteristic of the Lyα forest absorptions at redshift \(z \sim 2\) (Kim et al. 2001). This technique is efficient for data with \(S/N \gtrsim 10\), so we applied the selection only to the LP QSO sample and to the sample of simulated spectra for which, in this case, we adopted a Gaussian noise giving \(S/N = 50\). Fig. 4.4 shows the resulting correlation functions. The amplitude of the observed \(\xi_{\parallel}^f\) has reduced to about one third of the initial value but still shows a significant clustering signal. This is an indication that matter is still clustered also far from the most over-dense regions. The threshold on \(\tau_{\text{H}1}\) corresponds to \(\rho/\bar{\rho} \lesssim 6.5\). What is worth noting is that the amplitude of the simulated correlation function has decreased of about one half and is now fully compatible with the observed correlation function. This suggests that strong lines are deficient in the simulated spectra so that their contribution to the correlation function is smaller than for the observed spectra. When they are excluded from the computation the two correlation functions become consistent.

4.2.2 Selection effects in the observed spectra

There is the possibility that our observed sample is affected by a selection bias since the QSOs were required to be at close angular separations and, in most cases, they have similar emission redshifts. As a consequence, the observed lines of sight could be biased toward more clustered regions than average. We tried to estimate this effect on the auto-correlation function by excluding larger and larger portions of the observed spectra from the QSO emissions (up to 20000 km s\(^{-1}\)). No significant decrease of the amplitude of the correlation function has been detected. This test could prove inadequate as far as
the sestet is concerned since the redshifts of some member QSOs are separated by up to $\sim 50000$ km s$^{-1}$ (see Fig. 4.1). As a consequence, our measure could be affected by the transverse proximity effect (e.g. Jakobsen et al. 2003; Worseck & Wisotzki 2006) or by possible clustering associated with the presence of a foreground QSO. Indeed, in the case of Pair Q, we found a damped Ly$\alpha$ system in the higher redshift QSO spectrum at the redshift of the other QSO (D’Odorico et al. 2002). From a preliminary analysis of the sestet, strong Ly$\alpha$ absorptions ($\log N$(H$_i$) $\sim$ 15) are observed in the spectra of S3, S5 and S6 at $\sim 600 - 750$ km s$^{-1}$ from the emission redshift of the lower redshift QSO, S1, and a strong Ly$\alpha$ with associated C$iv$ absorption is present in the spectrum of S2, the furthest away QSO, at $\sim 1000$ km s$^{-1}$ from the redshift of S1. Considering the other foreground object, S4, Ly$\alpha$ absorptions within 200 km s$^{-1}$ of its redshift are observed along the line of sight to S2 ($\log N$(H$_i$) $\sim$ 14) and S3 ($\log N$(H$_i$) $\sim$ 15) without associated metals detected. We will carry out a detailed study of the coincidences among multiple lines of sight in the future, with data of improved S/N. In the following, we will use the auto-correlation function determined from the LP sample to avoid systematic uncertainties introduced by the above mentioned possible selection effect.

### 4.3 Clustering of the flux across the line of sight

In this section we exploit the potentialities of our sample of QSO pairs by determining the clustering properties of the IGM across the lines of sight. The great advantage with respect to the correlation function along the line of sight, in particular for a sample like
ours showing a large variety of pair separations, is that we have the guarantee of sampling true spatial separations between the pixels, the effect of peculiar velocities being negligible or absent.

4.3.1 The cross-correlation function in redshift space

As a first approach, we computed the cross-correlation function extending in a natural way the procedure adopted for the auto-correlation function. Every pixel along the line of sight is considered as an element of the density field at the QSO angular position in the sky and at a distance from the observer (comoving along the line of sight) corresponding to the redshift of the pixel:

\[ r_{\parallel}(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} , \]  

(4.3.1)

In eq. 4.3.1, there is the implicit hypothesis that peculiar velocities contribute negligibly to the measured redshift in the Ly\(\alpha\) forest. This statement is supported by recent measurements using QSO pairs (Rauch et al. 2005), and we will see a posteriori that the comparison of the auto and cross-correlation functions makes this procedure allowable.

The cross-correlation function of the transmitted flux between two lines of sight at angular separation \(\Delta \theta\) is defined as:

\[ \xi^f(\Delta r) = \langle \delta_f(\theta, r_{\parallel,1}) \delta_f(\theta + \Delta \theta, r_{\parallel,2}) \rangle , \]  

(4.3.2)

where, \(\Delta r = \sqrt{(r_{\parallel,1}^2 + r_{\parallel,2}^2 - 2 r_{\parallel,1} r_{\parallel,2} \cos \Delta \theta)}\) is the spatial separation between pixel 1 at \(r_{\parallel,1}\) along one line of sight and pixel 2 at \(r_{\parallel,2}\) along the paired line of sight. The error on the correlation function cannot be estimated with a bootstrap technique, as done before, due to the limited number of pairs contributing to each separation (in particular at the smaller separations). In order to give an evaluation of the significance of the observed signal, the same cross-correlation function was computed replacing one QSO in each pair with a control QSO, then repeating the operation replacing the other QSO of every pair with another control QSO. The control sample is formed by the LP QSO spectra. We chose LP QSOs at redshifts close to those of the original objects and shifted the spectra in order to match it exactly. We derived an indicative error as the r.m.s. deviation from zero of the two control functions in two regions: \(\Delta r < 5 \, h^{-1} \, \text{Mpc}\), for which the uncertainty is larger, and \(\Delta r > 5 \, h^{-1} \, \text{Mpc}\). We computed the cross-correlation function also for the sample of mock spectra. The simulated spectra are characterised by the redshift of the output box and a velocity extent. In order to assign a redshift value to every pixel, we gave the central pixel of every spectrum the redshift of the corresponding output box, then we numbered the pixels one by one transforming the velocity size of the pixel into a redshift size. Once the redshifts were determined pixel by pixel, we followed the same procedure adopted for the observed spectra for the 50 simulated samples and computed the average cross-correlation function and its 1 \(\sigma\) standard deviation. The result of our computation is shown in Fig. 4.5 compared with the observed cross-correlation function of the pairs and of the control sample. There is a very good agreement between the two functions. Only in the first bin there is an indication of a lower value for the simulated cross-correlation with respect to the observed one. This is consistent with the result
4.3. Clustering of the flux across the line of sight

Figure 4.5: Cross-correlation function of the transmitted flux for our sample of pairs and groups of QSOs (solid squares) compared with the analogous correlation function computed with the sample of mock spectra (empty squares). The error bars on the simulated data are $1\sigma$ standard deviations for 50 simulated samples while the error bars on the observed data have been obtained from the r.m.s. deviation of the two control samples plotted as dashed lines (see text). Simulated data are slightly shifted in $\Delta r$ for clarity.
obtained for the auto-correlation function if we take into account the fact that the first bin of $\xi^f_\perp$ corresponds to the second bin of $\xi^f_\parallel$.

### 4.3.2 The cross-correlation coefficient

A measure of the transverse clustering properties of the IGM which is less affected by peculiar velocities is the flux cross-correlation coefficient,

$$\chi^f_\perp(\Delta \theta) = \langle \delta_f(\theta, v_\parallel) \delta_f(\theta + \Delta \theta, v_\parallel) \rangle,$$

where, every pixel along one line of sight is correlated with the one face-to-face in redshift space along the paired line of sight and the result is averaged over all the pixels in the common redshift interval. Every pair of QSOs at angular separation $\Delta \theta$ gives one value of $\chi^f_\perp(\Delta \theta)$, and a sample with several pairs at different separations, as is our sample, gives an estimate of the correlation function. At a given redshift, the angular separation $\Delta \theta$ corresponds to a velocity separation $\Delta v_\perp = c F(z) \Delta \theta$, where $c$ denotes the speed of light, and $F(z)$ is a dimensionless function of redshift that includes all the dependence on the global cosmological metric. In the cosmological model that we have adopted:

$$F(z) = \frac{E(z) \int_0^z [dz/E(z)]}{(1 + z)}.$$

We computed the range of velocity separations covered by each of our pair of spectra then we grouped the pairs in velocity bins of variable width and computed the average cross-correlation coefficient for every group. Given the small number of pairs in every group (a maximum of 3 QSO pairs) the uncertainties reported on these determinations are computed by applying a simple error propagation to eq. 4.3.3, thus they account only for the pixel statistics and the noise associated with the transmitted flux but they are not representative of the true error due to the cosmic variance. In the case of the simulated spectra, we had 50 realisations of each of our QSO pairs so we could obtain in every velocity interval defined for the observed pairs an average cross-correlation coefficient with its error, that in this case is the standard deviation of the distribution of values. Results are shown in Fig. 4.6. The two points at the smallest separations are given by Pair U + S5S6 and Pair A, respectively. The two samples of values are in very good agreement except for the bin at $\Delta v_\perp \sim 500$ km s$^{-1}$ ($\simeq 5$ arcmin at $z \sim 1.8$) for which the observed value is $\sim 7 \sigma$ above the simulated value. The observed $\chi^f_\perp(\Delta \theta)$ in this velocity bin is obtained by averaging two pairs: S1S3 and T1T3 for which $\chi^f_\perp \simeq 0.040 \pm 0.003$ and $0.0177 \pm 0.0015$, respectively. The two pairs of spectra were searched for peculiar features that could boost the signal by selecting those pixels for which $\delta_f(\theta, v_\parallel) \delta_f(\theta + \Delta \theta, v_\parallel) > 0.4$. In Fig. 4.7 we show the results of this research. In the case of S1 (up left plot) about 44 per cent of selected pixels have values $f > 2$ or $f < 0$, that is they are strongly affected by noise (in the case of S3 only $\sim 14$ per cent of pixels are in those range of values). As a consequence, the large value of the cross-correlation coefficient for pair S1S3 could be due to the low S/N in the spectrum of S1 and cannot be determined reliably. On the other hand, in the case of pair T1T3, Fig. 4.7 (centre plots) shows a feature giving a strong signal at $\lambda \simeq 3300$ Å. We identified in the spectra of T1 and T3 two coincident strong H I
Figure 4.6: Comparison of the cross-correlation coefficients for our sample of observed spectra (solid dots) and of simulated ones (empty dots) as a function of the velocity separation, $\Delta v = c F(z) \Delta \theta$, corresponding to the angular separation, $\Delta \theta$, of the QSO pairs. The angular separation computed at $z = 1.8$ is reported in the top axes. Observed values are shifted by 10 km s$^{-1}$ for clarity. The lower observed value at $\Delta v \simeq 500$ km s$^{-1}$ was obtained excluding pair S1S3 and removing a strong coincident Ly$\alpha$ line in the T1T3 pair (see text). Error bars on the observed values along the x axis represent the velocity range covered by the considered pairs. Errors on the $\chi^2_x$ represent the uncertainty on that measure for the observed values and the cosmic variance for the simulated values (see text).
Figure 4.7: Selection of the pixels giving values $\delta_f(\theta, v_{||})\delta_f(\theta + \Delta \theta, v_{||}) > 0.4$ for the pairs S1S3 (left), T1T3 (centre) and T2T3 (right). The two upper plots show the values of the flux for the selected pixels in the indicated QSO while the lower plot report the value of the above product.
Lyα absorptions ($\log N(\text{HI}) \sim 15$) at this wavelength, the former with a clear associated C iv doublet while for the second there is the possible detection of the C iv $\lambda 1548$ line only. No corresponding H i Lyα line is observed at this redshift along the third line of sight of the triplet, T2. If we mask the redshift interval covered by the coincident lines and recompute $\chi^f_\times$ for T1T3 we get: $0.0075 \pm 0.0015$, which is in better agreement with the simulated value as shown in Fig. 4.6. In the right hand plots of Fig. 4.7, we show the case of pair T2T3 (at 8.9$'$ angular separation) where a coincident absorption system is present with characteristics very similar to the one in pair T1T3, and the cross-correlation coefficient has a value consistent with zero. This suggests that the presence of a coincident absorption system could be a necessary but not sufficient condition to explain a large value of $\chi^f_\times$. It is interesting to note that also the recently published two point correlation function of the C iv absorptions in the LP QSO spectra (Scannapieco et al. 2006) shows an excess of clustering signal at $\sim 500$ km s$^{-1}$ whose origin has not been satisfactorily explained. New observations are going to be carried out with UVES to increase the S/N in the two QSO pair spectra S1S3 and T1T3, in order to confirm the large value of the cross-correlation coefficient and possibly detect other coincident C iv systems. On the other hand, we will try and collect other QSO pairs at the same velocity separation to increase the statistics and we will look into physical mechanisms that could explain our results. In Fig. 4.8, we compare the cross-correlation coefficients with the cross-correlation function. The angular separation $\Delta \theta$ between two QSO lines of sight was transformed into a comoving spatial separation, $\Delta r$, with the formula:

$$\Delta r = \frac{c \Delta \theta}{H_0} \int_0^z \frac{dz'}{E(z')}$$

(4.3.5)
4.4 Discussion

Considering the error bars computed from the simulated pairs, there is a good agreement within 1σ. The large variations from one data point to the other in the cross-correlation coefficients should be due mainly to the small number of QSO pairs (between 1 and 3) contributing to each point. On the other hand, the smoothness of ξ′ is artificially increased by the fact that the values in the different bins are not independent.

4.4 Discussion

From the cosmological point of view, one of the most interesting and challenging applications of the kind of calculations carried out in this chapter is the constraint on the geometry of the Universe (in particular, the estimate of Ω_0Λ h^2) by linking angular separations and redshift differences in the hypothesis that the observed correlation properties are isotropic (Alcock & Paczynski 1979; McDonald & Miralda-Escudé 1999; Hui et al. 1999). A large number of QSO pairs at different angular separations (see McDonald 2003) is needed to obtain a measure of the cosmological parameters as accurate as those recently produced by the WMAP team (Spergel et al. 2003, 2007). This is why we do not attempt to derive a measure of Ω_0Λ with the present sample. Nevertheless, a comparison

![Figure 4.9: Comparison of the cross-correlation function for our sample of QSO pairs (squares) with the auto-correlation function computed for the LP QSOs sample (crosses) as a function of comoving spatial separation across and along the line of sight, respectively. The cross-correlation function is slightly shifted in Δr for clarity](image)

of the observed auto and cross-correlation functions is interesting to qualitatively evaluate the distortion due to peculiar velocities and the correctness of the adopted cosmological parameters. The two functions are shown in Fig. 4.9. Note that, due to the minimum angular separation between our QSO pairs, there is no transverse clustering signal corre-
sponding to the first bin of the auto-correlation (centred at $\Delta v = 50$ km s$^{-1}$ or $\Delta r \simeq 740$ kpc). The correspondence between $\xi^f_{\parallel}$ and $\xi^f_{\perp}$ is very good suggesting that the adopted cosmological parameters are reasonable and that peculiar velocities do not play a major role in the IGM gas on velocity scales $\gtrsim 100$ km s$^{-1}$.

4.4.1 Exploring the parameter space of the simulations

In this section we survey the behaviour of the simulated correlation functions when varying three parameters characterising the physical state of the gas.

4.4.1.1 Effective optical depth

The effective optical depth of a simulated spectrum can be rescaled to a new value simply by increasing or decreasing the optical depth of every pixel in the spectrum by a constant factor, which is equivalent to increase or decrease the column density of the absorption lines in the spectrum. Considering previous observations (Viel et al. 2004b) and the relations computed by Schaye et al. (2003) for the redshift evolution of the effective optical depth, we derived 3 $\sigma$ upper and lower limits for $\tau_{\text{eff}}$ at the redshifts of the simulated spectra, which are reported in Table 4.3. Then, we built two new samples of mock spectra starting from the original spectra produced from the simulation, one with the minimum and one with the maximum values of $\tau_{\text{eff}}$ and recomputed the auto and cross-correlation functions. Each sample is formed by 50 realisations of the observed sample of QSO pairs and we determined the average value for the correlation function. Our results are shown in Figs 4.10 and 4.11. At the smallest scale, probed only by the auto-correlation function, the agreement between observed and simulated clustering signal improves adopting the larger effective optical depths. However, at the Jeans scale, corresponding to the first bin of $\xi^f_{\perp}$ and to the bin centred at 100 km s$^{-1}$ of $\xi^f_{\parallel}$, the larger $\tau_{\text{eff}}$ possibly overestimates the clustering amplitude. Since the effective optical depth measured in the original simulation is in very good agreement with the one observationally determined from QSO Ly$\alpha$ forests, we infer that the discrepancies between simulations and observations are not due to a wrong estimate of $\tau_{\text{eff}}$ but, to the deficiency of strong lines in the simulations. In the spectra with the larger $\tau_{\text{eff}}$ the column density of all the lines is increased and the average flux is decreased, causing a general increase of the amplitude of the correlation function.

<table>
<thead>
<tr>
<th>QSO</th>
<th>$z$</th>
<th>$\tau_{\text{min}}$</th>
<th>$\tau_{\text{obs}}$</th>
<th>$\tau_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sestet</td>
<td>1.8</td>
<td>0.08</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td>Triplet</td>
<td>1.8</td>
<td>0.08</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td>Pair U</td>
<td>1.8</td>
<td>0.08</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td>Pair A</td>
<td>2.4</td>
<td>0.17</td>
<td>0.21</td>
<td>0.29</td>
</tr>
<tr>
<td>Pair Q</td>
<td>2.4</td>
<td>0.17</td>
<td>0.21</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 4.3: Effective optical depth adopted to build the three samples of mock spectra used to compute the correlation functions in Figs 4.10 and 4.11.
Figure 4.10: Auto-correlation function for the LP sample (crosses) related to the average auto-correlation functions of three samples of simulated spectra with different $\tau_{\text{eff}}$. The long dashed line traces the correlation function for the sample with the minimum optical depths, the dashed line is for the sample with the observed optical depths and the solid line is for the sample with the maximum optical depths (see Table 4.3).

Figure 4.11: Cross-correlation function for our sample of QSOs (solid squares) compared with the average cross-correlation functions of three samples of simulated spectra with different $\tau_{\text{eff}}$. The long dashed line traces the correlation function for the sample with the minimum optical depths, the dashed line is for the sample with the observed optical depths and the solid line is for the sample with the maximum optical depths (see Table 4.3)
partially compensating the small number of strong lines. In order to verify our statement, a quantitative analysis of the difference in the Lyα column density distribution function (the number of absorption lines per unit column density per unit redshift) of observed and simulated spectra (as done by Theuns et al. 2002, for small simulated boxes) is necessary. Furthermore, more pairs at small separations (< 1') would allow to decrease the error bars and test if the behaviour of clustering is the same as along the line of sight.

Figure 4.12: Cross-correlation function for our sample of pairs (solid squares) compared with the average cross-correlation functions of samples of simulated spectra with different temperature and exponent $\gamma$ of the temperature-density relation. Short dashed lines are for simulations in which $\gamma = 1.6$ and temperature is $T = 8000$, 22000 and 50000 K going from the upper to the lower curve. The long dashed lines represents simulations with $\gamma = 1.1$ while the temperatures are the same as before but they increase from the lower to the upper curve.

4.4.1.2 Temperature and $\gamma$

The influence on the flux cross-correlation function of the temperature and of the power-law index of the temperature-density relation adopted for the gas are explored. The aim of this analysis is to point out the modifications of the simulated correlation function when the parameters are varied, and not to derive precise constraints on the same parameters from the comparison with observations. With a running time of 2-3 weeks for simulation, it is impossible to run an extensive parameter study for different thermal histories, thereby we decided to rescale a posteriori the temperature-density relation at a given output by assuming that all the gas particles obey this relation. We then recomputed the ionization fractions for each gas particle with the proper UV background. This approximation does not account for the effects that the corresponding change in the gas pressure would have on
the gas distribution. However we explicitly checked that the differences in the correlation function for the fiducial run and for a rescaled temperature relation with the same $T - \gamma$ are of the order of 10 per cent (see Viel et al. 2004b, for the effect of the rescaling on the flux power spectrum). We built six sets of simulated spectra each one formed by 50 realisations of the observed sample of pairs. Every set is characterised by a value of $\gamma$ (1.1 or 1.6) and a value of the temperature (8000, 22000, and 50000 K), while the effective optical depth is the one used for the main sample. In Fig. 4.12, we plot together the observed $\xi_f$ with the average correlation functions computed from the six groups of simulated spectra. At the smallest separation probed by the observations, the predictions of the different models are well separated and could be tested with a larger sample of close QSO pairs. The models with $\gamma = 1.1$ give a larger clustering signal than those with $\gamma = 1.6$ for all temperatures and are in better agreement with the data at the small scales. The exponent of the temperature-density relation is determined by the ionization history of the gas. Hui & Gnedin (1997) demonstrated that instant reionizations of $\text{H} \, \text{I}$ occurring at redshifts decreasing from $z = 10$ to 5 would imply shallower and shallower $\gamma$ at redshifts $z \sim 3$. However, also for very late reionizations $\gamma$ should have increased above values $\sim 1.4$ by $z \sim 2$. Ricotti, Gnedin, & Shull (2000) studied the evolution with redshift of the equation of state of the gas and showed that a reionization of $\text{He} \, \text{II}$ occurring at $z \sim 3$ would cause a sudden increase in the gas temperature and a corresponding decrease in the value of $\gamma$ (their fig. 13). Observationally, this has been studied mainly by using the distribution of Doppler parameters and column densities of the Ly$\alpha$ forest lines (e.g. Schaye et al. 2000; Kim et al. 2002), but also with metal absorptions (e.g. Songaila 1998; Vladilo et al. 2003). Our result supports previous observational evidences of a second reheating of the Universe happening at $z \sim 3$, most likely due to the reionization of $\text{He} \, \text{II}$ and it requires confirmation by a larger sample of observed spectra and a set of more refined mock spectra.

### 4.5 Future Perspectives

As already said, the full exploitation of the potential offered by multiple QSO LOS is still limited by the dearth of groups of QSOs bright enough to make high resolution spectroscopy possible. Recent QSO surveys (2dFQRS, SDSS) have provided large and dense samples of QSOs at high redshifts. However, when looking for small angular separations the quasar magnitudes often exceed the $V \approx 20.5$ which are near to the limits for high spectral resolution observations at an 8-10 m telescope. To measure the correlation of the Ly$\alpha$ absorptions typical statistical tools are the two-point correlation function and the step optical depth correlation function (Miralda-Escudé et al. 1996; Kim et al. 2001). A relatively high resolution is necessary to clearly define the line-free continuum, which is essential for measurements of the transmitted flux. Spectral resolution and complete wavelength coverage including regions longward of the Ly$\alpha$ emission are also important to identify and remove metal lines falling in the Ly$\alpha$ forest, as well as to study the physical properties of the gas. In order to select a significant (>100) number of close pairs and groups (with separations up to 3') from inhomogenous catalogs (Véron-Cetty & Véron 2006) and from the new
Figure 4.13: Distribution of pairs of QSOs as a function of the V-Magnitude of the fainter one, selected from the Véron catalog (Véron-Cetty & Véron 2006) with the following criteria: declination $\delta < +15^\circ$, redshift and redshift difference between the QSOs in the pair $z \in [2,3]$, $\Delta z < 0.1$, separation $\theta \leq 3$ arcmin
2dFQRS and SDSS surveys at present it is necessary to push down to magnitudes $V \simeq 21$. Since for all the calculations we refer to AB-magnitudes, we recall that $V_{AB} \sim V$. Previous analysis of QSO spectra at redshift $z \sim 2 - 2.5$ show that in order to apply the standard methodology a $S/N \simeq 7$ per resolution element and a resolution $R \gtrsim 5000$ have to be reached. In this respect the UV sensitivity is important since it provides access to a lower redshift domain in which brighter and/or denser targets are easier to find. At present, the ESO instrumentation shows a gap in this area of the parameter space: UVES is not sensitive enough and FORS resolution is too low (and, with the higher resolution grisms, the spectral range is too limited), FLAMES/GIRAFFE would provide the desired resolution but the limited spectral coverage in the individual exposures and the lack of UV-blue sensitivity are major drawbacks (as it is also the case of FORS). All this demonstrates the necessity of a highly efficient and intermediate resolution ($R \gtrsim 5000$) spectrograph with a large spectral coverage extending into the UV. At the moment, the only instrument with similar capabilities in the world is ESI at Keck. Its UV efficiency is however very poor and this makes it impossible to carry out a systematic study of QSO pairs in the $z = 2 - 2.5$ range. In Chap. 6 it will be described a new instrument under construction at ESO, the so called X-shooter: this instrument will be a useful tool for the IGM tomography studies.
A wealth of observations has recently become available both on the cosmic microwave background (e.g. from the satellite WMAP) and on galaxies (e.g. from local surveys like 2dF or SDSS and deep fields like GOODS or COSMOS) which have made astrophysics and cosmology enter a new “high-precision” era. However, our understanding of the formation and assembly of galaxies is still poor, due to the high level of complexity of the physical processes involved in star formation, including its energetic feedback and metal pollution on the inter-stellar and inter-galactic media (ISM, IGM), not to mention the feedback from accreting supermassive black holes. In particular, feedback from star formation affects the IGM through the ejection of gas in galactic winds and the emission of ionising radiation. The two sources of feedback are not independent, as the removal of ISM through a wind can increase the escape fraction of ionising photons (Dove et al. 2000). Galactic winds have been commonly observed in local galaxies (the best studied cases are the Milky Way, M82 and NGC 3079) and their origin is ascribed to starburst events and/or AGN activity. At higher redshift it is not possible to directly observe the outflows; however evidence for galactic winds has now been found in a number of $z \sim 2 - 5$ galaxies. Low-ionisation interstellar absorption lines that are blueshifted by hundreds of kilometres per second relative to systemic velocities, and Ly-$\alpha$ emission lines similarly shifted redward, have been detected in most $z \sim 3 - 4$ Lyman break galaxies (LBGs) (Pettini et al. 2002; Adelberger et al. 2003; Shapley et al. 2003), in several gravitationally lensed Ly-$\alpha$ emitting galaxies at $z \sim 4 - 5$ (Franx et al. 1997; Frye et al. 2002), and in many luminous IR galaxies at $z \gtrsim 2$ (Smail et al. 2003; Swinbank et al. 2005). The physical processes responsible for the ejection of gas out of a galaxy, and then out of its host dark-matter halo, are extremely complex and at the moment it is not possible to model them in a completely self-consistent way (Monaco 2004). In particular, both the fraction of outflowing material and that of ionising radiation that can reach the IGM are very difficult to determine. The most recent cosmological hydro-simulations (Cen et al. 2005; Kollmeier et al. 2006) confirm the previous results (Bruscoli et al. 2003) that winds emerge along paths of least resistance, possibly avoiding large-scale filaments, and contribute to the metal enrichment and to the heating of the IGM. In order to constrain this remarkably important process, we are interested in studying the impact of
such winds on the surrounding low-density environment, in particular on the IGM. From the observational point of view, a very important and promising way to study galactic winds at high redshift is to observe their effect on the IGM as seen in absorption of a background QSO continuum light.

This chapter will describe our observational proposal on a pair of QSO, the observation we have done and the preliminary results we obtained. My own contribution to this project was to plan the observation, submit it to ESO, perform the observation at Paranal, reduce and analyze data, and plan new observations.

5.1 QSO absorption lines as probes of galactic winds

The first evidence of pollution of the IGM by galactic winds at \( z < 2 \) comes from the Mg\( ^{II} \) absorbers. Comparisons between the number densities of Mg\( ^{II} \) absorbers, star-forming galaxies and the properties of local outflows (Bond et al. 2001) suggest that there are enough Mg\( ^{II} \) absorbers to account for the expected properties of winds at \( 1 < z < 2 \). At larger redshift (\( z > 2 \)) many studies have investigated the correspondence between metal absorptions (mainly C\( IV \) and O\( VI \)) and LBGs and have tried to infer from this evidence the metal enrichment history of the IGM. The main results can be summarised as follows. There is a tight correlation between LBGs and C\( IV \) absorptions with \( N(CIV) \gtrsim 10^{12.5} \) (Adelberger et al. 2005) which however may not be enough to claim that galaxies are responsible for the metal enrichment in their neighbourhood (Porciani & Madau 2005). On the other hand, gas in the IGM, as traced by the Lyman-\( \alpha \) forest in QSO absorption spectra, is metal enriched down to rather low column densities (\( N(HI) \sim 10^{13.5} \) cm\(^{-2} \)), but this metallicity is found to increase in the proximity of galaxies, identified by strong H\( I \) or C\( IV \) absorptions (Aracil et al. 2004; Pieri et al. 2006). Remarkably, both the column density distribution of C\( IV \) absorbers and its integral (\( \Omega_{CIV} \)) are roughly constant over the interval \( 1.5 \lesssim z \lesssim 6 \) (Songaila 2001; Pettini et al. 2003; Ryan-Weber et al. 2006). One possible explanation is that most of the IGM metals are already in place by \( z \sim 6 \), perhaps from SN-driven outflows from low-mass subgalactic systems (Qian & Wasserburg 2005). However, this scenario does not completely explain why \( \Omega_{CIV} \) remains constant over this redshift range despite the known variations in the intensity and spectrum of the ionising background. Alternatively, the systems may be associated directly with outflows from star forming galaxies and the constancy of \( \Omega_{CIV} \) arises from the balance between the continual IGM enrichment and the changes in the ion fraction C\(^{3+} / C_{\text{TOT}} \) (Oppenheimer & Davé 2006). In summary, the present observations are not sufficient to put firm constraints on the nature of the galactic winds and on their effect on the IGM.

We proposed (ESO proposal, P79.A-0108, 'Who polluted the InterGalacticMedium? A pilot study in the field of a pair of QSOs', P.I. F.Saitta) to further investigate this topic adopting an approach, which is different from the previous studies in two main aspects:

1. we consider a pair of QSOs at close angular separation (\( \lesssim 1 \) arcmin) instead of a single QSO. The use of two adjacent lines of sight makes us benefit of the information
both in the longitudinal and transverse directions. In particular, once we detect a galaxy at a small impact parameter from one of the two lines of sight we will have the possibility to study the metal enrichment of the same galaxy at two different separations. In the ideal case of a galaxy sitting between the two lines of sight and giving rise to two metal absorption systems in the paired spectra, it would be possible also to obtain information on the dynamics of the gas from the relative positions of the absorbers.

2. we explore an intermediate redshift range, \( 1 \lesssim z \lesssim 1.7 \). The advantage with respect to the \( z < 1 \) range is that star formation is still very active at this epoch. On the other hand, we will be able to identify spectroscopically galaxies at \( L^* \) and fainter, which are better representatives of the bulk of the population than Lyman break galaxies at \( z > 2.5 \). Smaller galaxies are also better suited to study feedback effects since winds outflow more easily from their halos. Furthermore, galaxies can be studied in greater details, obtaining for examples rotation curves from the \([\text{O II}]3727 \) in emission (Vanzella et al. 2006) to be compared with the dynamical properties of the gas seen in absorption, and morphological information which is less affected by the \((1 + z)^4\) dimming.

The main idea of the project is to look for correspondence between metal systems (in particular \( \text{C IV} \) absorptions) observed in the spectra of a close pair of QSOs and galaxies. The steps of such a program would be:

- high resolution spectroscopy of the pair, to look for metal absorptions;
- imaging of the field to search for galaxies at the redshift of the coincident absorption (assuming to have found at least one of them);
- spectroscopy of the galaxy (if found) to determine precisely its redshift;

The proposal for the high resolution spectroscopy was accepted and in August 2007 our observations have been carried out in two nights of visitor mode at the ESO VLT with the UVES spectrograph.

### 5.2 Observations, Data Reduction, Preliminary Results and Work in Progress

The main characteristics of the observed QSO pair, separated by 1.03', are listed in tab. 5.1. Data have been reduced and analysed and we present here the preliminary

<table>
<thead>
<tr>
<th>Object</th>
<th>( \alpha(J2000) )</th>
<th>( \delta(J2000) )</th>
<th>hours allocated</th>
<th>Magnitude</th>
<th>emission redshift</th>
</tr>
</thead>
<tbody>
<tr>
<td>2QZ J225153-3146</td>
<td>22 51 53.2</td>
<td>-31 46 19</td>
<td>8.0</td>
<td>19.42</td>
<td>1.762</td>
</tr>
<tr>
<td>2QZ J225154-3145</td>
<td>22 51 54.9</td>
<td>-31 45 21</td>
<td>7.0</td>
<td>19.35</td>
<td>1.939</td>
</tr>
</tbody>
</table>

Table 5.1: Magnitudes are in \( B_J \) band from the \( GSC - 2 \).
results. In the spectrum of J225153 we found two associated systems \((z_{\text{abs}} \approx z_{\text{em}})\) identified through the NV doublet at \(z_{\text{abs}} \approx 1.7605\) and 1.7654. The latter system shows also the SiIV doublet. The spectrum of J225154 shows three intervening CIV absorptions, at \(z_{\text{abs}} \approx 1.357\) (strong), 1.655 (strong) and 1.707, but none of them has a corresponding system in the spectrum of J225153 (see Fig. 5.2). However, we also detect a weak SiIV at \(z_{\text{abs}} \approx 1.76794\) separated by only \(\sim 270\) km s\(^{-1}\) from the strong associated system in the paired spectrum (see Fig. 5.2).

![Figure 5.1: Coincident SiIV doublets at \(z \sim 1.77\) in the spectra of J225153-3146 (upper panel) and of J225154-3145 (positions defined by the tickmarks in the lower panel). The velocity separation between the two systems is \(\sim 270\) km s\(^{-1}\)](image)

### 5.2.1 New Observations

We still need observations to interpret these results; in September 2007 (ESO P81) we asked for new observations. In particular we asked for:

1. Completion of high resolution spectroscopy: in order to get more information on the physical and chemical properties of the detected coincident systems at \(z \sim 1.77\), it is necessary to extend the spectral range to the blue and to the red to observe the corresponding H1 Lyman-\(\alpha\) and CIV absorptions. We asked for 2 nights in visitor mode using again the high resolution spectrograph UVES at VLT.

2. FORS2 imaging of the pair field and identification of galaxies close to the lines of sight: the spectra of the observed pair present at least 4 interesting metal systems
5.3. Future Perspectives

In this chapter we have described a pilot study in the field of the interplay between galaxies and metal absorption in QSO spectra. Because of the lack of QSO pairs reachable with the current instrumentation available, at the moment is not possible to perform an intensive analysis with the method described above. To exploit a big observational campaign of

(one coincident and three single systems) whose constraining power on the properties (e.g., efficiency, isotropy) of galactic outflows and IGM metal pollution can be fully exploited only when the complementary information on the distribution of galaxies in the field is acquired. Based on the luminosity function at $z \sim 1.5$ (Giallongo et al. 2005), $L^*$ galaxies at this redshift have magnitude $I_{AB} \simeq 23.28$. In order to sample a representative portion of the population and thanks to the high efficiency of FORS2 in the I-band, we will be able to easily reach $0.2L^*$ (or $I_{AB} \simeq 25$). Based on the GOODS-MUSIC catalogue (Grazian et al. 2006), and considering a maximum impact parameter of $\sim 1$ arcmin from the two QSO sightlines as the region of interest for our study, we expect to find $\sim 50$ galaxies brighter than our magnitude limit with $1 \lesssim z \lesssim 2$. The selection of candidates for the follow up spectroscopy will be based on imaging in V, I and z band. We asked for 4hrs in service mode to have the imaging of the field.

Figure 5.2: The two strong C IV absorption systems in the spectrum of J225154-3145 at $z \simeq 1.357$ (left lower panel) and at $z \simeq 1.655$ (right lower panel) without detected counterparts in the spectrum of J225153-3146 (upper panels).
QSO pairs at $z \sim 2 - 3$ we need to wait for new instruments able to reach the faint magnitudes of these objects and a high enough resolution to resolve metal systems. As we have already noticed, X-shooter will have enough sensitivity to observe many pairs of QSOs.
This chapter describes X-shooter, a new spectrograph built for the ESO-VLT, its importance for the study of the IGM, and the work done in the X-shooter project. My own contribution to the project was to

- collaborate in writing and updating the X-shooter science case for the IGM,
- perform and analyze laboratory measurements for calibration sources of the near-infrared (NIR) arm of the instrument,
- participate to the observational project to build a spectro-photometric catalogue of standard stars, reducing and analyzing the data.

6.1 The instrument and the IGM Science Case

X-shooter is a single target spectrograph for the Cassegrain focus of one of the VLT UTs covering in a single exposure the spectral range from the UV to the K band (320 - 2500 nm). It is designed to maximize the sensitivity in this spectral range by splitting the incoming light into three arms (UVB, VIS and NIR) with optimized optics, coatings, dispersive elements and detectors. X-shooter will be a unique instrument on 8 m class telescopes in that it is capable of recording - over such a large wavelength range - the spectrum of an astronomical target in a single exposure. It operates at intermediate resolutions (R=4000-14000, depending on wavelength and slit width) sufficient to address quantitatively a vast number of astrophysical applications while working in a background-limited S/N regime in the regions of the spectrum free from strong atmospheric emission and absorption lines. The instrument is currently undergoing subsystem assembly and commissioning is scheduled for summer 2008. Table 6.1 lists the main characteristics of X-shooter. Considering the emission wavelengths of the main elements used for the
Spectral Format | Prism cross-dispersed echelle (order separation ≥ 12″)
---|---
Wavelength range | 300-2500 nm, split in 3 arms by dichroics
Slit configuration | long slit (∼ 12″); widths: 1″ (standard), 0.6″ (high resolution), 5″ (flux calibration); IFU 1.8″ × 4″ input area
Detectors | 2k × 4k CCDs (UV and Visual-Red arms), 2k × 2k (1k × 2k used) Hawaii LPE MCT (IR)
Auxiliary functions | Calibration Unit; acquisition and guide unit with 1′ × 1′ field and filter set; ADC for the UV and Visual arms.

Table 6.1: X-shooter characteristics

analyses discussed in chap.4:

\[ z=2 \Rightarrow \text{Ly} \alpha @ \lambda \sim 3647 \text{ Å}, \text{Ly} \beta @ \lambda \sim 3077 \text{ Å}, \text{CIV} 1550 @ \lambda \sim 4652 \text{ Å} \]

\[ z=3 \Rightarrow \text{Ly} \alpha @ \lambda \sim 4862 \text{ Å}, \text{Ly} \beta @ \lambda \sim 4100 \text{ Å}, \text{CIV} 1550 @ \lambda \sim 6203 \text{ Å} \]

and the X-shooter wavelength range (all the numbers referring to the instrument’s characteristics in this section are taken from the Optical Design document of X-shooter)

<table>
<thead>
<tr>
<th>Arm</th>
<th>UVB</th>
<th>VIS</th>
<th>NIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range(nm)</td>
<td>300-550</td>
<td>550-1000</td>
<td>1000-2300</td>
</tr>
</tbody>
</table>

the interesting characteristics of X-shooter for this kind of project are:

<table>
<thead>
<tr>
<th>Arm</th>
<th>UVB</th>
<th>VIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (slit 1.0″ and 0.6″)</td>
<td>~ 5000 - 7000</td>
<td>~ 7500 - 11500</td>
</tr>
<tr>
<td>Limiting AB-Magnitudes (SNR 10, T=1hr)</td>
<td>~ 21.5</td>
<td>~ 21.6</td>
</tr>
</tbody>
</table>

As a consequence, this instrument will allow the scientific community to obtain spectra of the candidate QSOs at \( V \simeq 21 \) with the desired signal to noise ratio in a reasonable amount of exposure time (\( \simeq 1 \) h, see Fig. 6.1 and 6.2).

The long-term goal is to carry out the Alcock-Paczynski test (Hui 1999; McDonald & Miralda-Escudé 1999) that bases the measurement of the geometry of the universe (with particular sensitivity to \( \Omega_\Lambda \)) on the comparison between the 3-D correlation of neighboring QSO spectra and the clustering along the individual LOS. McDonald (2003) has shown...
Figure 6.1: A typical spectrum of the Lyman forest of a $z \sim 2.6$ QSO at $R \sim 5000$ at a S/N $\simeq 11$ per resolution element, obtained rebinning to the desired resolution a real QSO spectrum, with a predicted sampling of $\sim 7 \text{ pix}/fwhm$. This is a section of a 1hr exposure spectrum of a $V \simeq 21$ QSO that could be obtained with X-shooter.
Figure 6.2: Signal to noise ratio (SNR) expected for the UV arm (@ $\lambda = 330nm$) of X-Shooter as a function of the exposure time at different AB-Magnitudes.
that at least $13 \left( \theta/1' \right)^2$ QSO pairs with separation $< \theta$ (including separations $< 10'$) are needed to apply the Alcock-Paczynski test and determine the value of $\Omega_A$ with an accuracy of 10%.

A satisfactory estimation could therefore be reached with $\sim 80$ QSO pairs, at a mean separation of $\sim 2 - 3$ arcmin.

The time allocation to obtain the spectra of about 160 faint ($V \lesssim 21$) quasars with X-shooter would be of the order of 160 hr (see Fig. 6.2 and estimates from the ETC). X-shooter does provide the suitable trade-off between resolution, sensitivity and spectral range to carry out these observations in a reasonable amount of time, and would thus give the chance to get a unique cosmological result.

### 6.2 Laboratory Measurements of Wavelength Calibration Sources in the near IR

We have studied the properties of wavelength calibration sources for the near-IR (NIR: 1000-2500 nm) arm of X-shooter. In a novel approach we combine laboratory measurements from a Fourier Transform Spectrometer (FTS) and literature data with corresponding simulated data derived from a physical model of X-shooter. The sources studied are pen ray lamps filled with the noble gases Ne, Ar, Kr, and Xe and Th-Ar hollow cathode lamps. As a product we provide a quantitative order-by-order analysis of the expected properties of the calibration lamps during X-shooter operations. Based on these we found that a combination of Ne, Kr, and Ar is the best lamp combination for X-shooter calibration. To our knowledge this is the first time that such a detailed and quantitative analysis of a calibration system has been done prior to the operation of the instrument. The combination of laboratory measurements and instrument modeling provides a powerful tool for future instrument development.

#### 6.2.1 Wavelength Calibration during X-shooter operations

X-shooter is equipped with a dedicated calibration unit providing light for flat fielding and wavelength calibration across the entire operating range of the instrument. The unit consists of two arms:

- one employs a diffuser plate illuminated by faint sources
  1. D$_2$ lamp used for the Flat Field (FF) in the UVB arm and the Th-Ar Hollow cathod lamp (HCL) for wavelength calibration of the UVB and VIS.

- The other uses a 6-inch integrating sphere fed by bright sources
  1. 3 Quartz Tungsten Halogen (QTH) for UVB, VIS, and NIR FF
  2. up to 4 pen ray lamps (Ne, Ar, Kr, Xe) for NIR wavelength calibration
The light of the up-to-four pen ray lamps is superimposed in the integrating sphere and a combined spectrum is fed to the spectrograph. The four pen ray lamps will be operated as a single source; thus all lamps, although using separate power supply units (PSUs), will be burning simultaneously and for the same length of time. Hence, the intrinsically different intensity levels of the lamps need to be balanced by the positioning of the lamps inside the integrating sphere and by shielding cylinders mounted around the lamps. We have studied the properties of pen ray lamps with a noble fill gas, Ne, Ar, Kr or Xe in order to provide a basis for the selection of the best lamp combination. The results provided will be verified during testing and commissioning of X-shooter but will also help to optimize and speed up the laboratory tests by providing quantitative predictions of the performance of the calibration lamps.

### 6.2.2 Wavelength Calibration Sources for X-shooter

To fully realize the scientific potential of X-shooter, excellent wavelength calibration across all three arms is essential. For UVB and VIS, Th-Ar hollow cathode lamps have been chosen as calibration sources following the successful operations of such lamps in other ESO spectrographs, like FEROS, FLAMES, HARPS and UVES. For the NIR arm the solution is less obvious and we decided to conduct a dedicated program to select the best combination of calibration sources for this wavelength region which traditionally has relied on atmospheric features for wavelength calibration. Recently, ESO has gained significant experience with NIR wavelength standards in a collaboration with the US National Institute of Standards and Technology (NIST) as part of the CRIRES project (Kerber et al. 2007). Currently, there is no comprehensive database of emission line spectra of commercially available light sources. Based on experience, a combination of gas discharge lamps (Ne, Ar, Kr, Xe pen ray lamps) were envisaged. In addition we considered the possibility of utilizing a Th-Ar HCL - calibration source for X-shooter UVB and VIS - also for the NIR arm.

### 6.2.2.1 Pen ray lamps

These lamps are called “Pencil” lamps because of their size and shape. They are made of double bore quartz tubing with two electrodes at one end sealed into a handle. These lamps produce narrow, intense lines from the excitation of rare gases and metal vapors. They are commercial products widely used for wavelength calibration of spectroscopic instruments such as monochromators, spectrographs, and spectral radiometers e.g. in industrial and chemical applications.
6.2.2.2 Th-Ar Hollow cathode lamps

Modern commercial hollow cathode lamps (HCLs) are sealed-off glass tubes that contain a metal cathode, a metal anode and a fill gas at a given pressure. The lamp is operated by applying a voltage of a few hundred volts across cathode and anode. As a result, a discharge is formed in the low pressure (few hundred Pa) fill gas and positive ions of the plasma are accelerated towards the cathode where they release matter through sputtering. As a result a HCL emits a rich spectrum of narrow emission lines from both the gas and metal atoms and ions in the plasma; for a detailed technical review of HCL and their operations see Kerber et al. (2007). The Th spectrum was studied more than 20 years ago in the range from 278 nm to about 1000 nm at high resolution by Palmer & Engleman (1983). Its emission lines are very narrow and the spectrum is rich over a wide wavelength range. In nature Th has only one isotope, $^{232}$Th, which has zero nuclear spin. Thus the use of Th for calibration lines avoids complex and asymmetric line profiles attributable to isotopic or hyperfine structure. Th-Ar HCLs are widely used for wavelength calibration of high resolution spectrographs in the Visual, including many examples at ESO such as FEROS, FLAMES, HARPS, and UVES. A detailed account of the properties, design and operations of HCLs is given in Kerber et al. (2007). Two valuable studies of the Th-Ar spectrum in the near IR have recently been published, but neither is directly applicable to the operation of X-shooter:

- Hinkle et al. (2001) produced an atlas of the Th-Ar spectrum covering selected regions in the range 1000 nm to 2500 nm. Their list of about 500 lines contains significant gaps in wavelength coverage.
- A fundamental analysis of the Th-Ar spectrum was provided by Engleman et al. (2003). Their list contains more than 5000 lines. Their high-current source is very different from low-current commercially available lamps and is not well suited for operation at an astronomical facility. Although the spectrum from the high-current lamp is significantly different from commercial Th-Ar lamps, the line list is highly valuable for identification of the lines in low-current spectra.

The spectrum of low-current Th-Ar HCLs has been studied extensively at high spectral resolution by a collaboration of ESO and the US Institute of Standards and Technology (NIST) in preparation for ESO’s CRIRES. X-shooter directly benefits from this experience.

6.2.2.3 Experimental Setup

ESO operates a commercial Fourier Transform Spectrometer (FTS) (Thermo 5700) in its laboratory (Fig. 6.3). The spectrometer is equipped with an external port that allows one to feed the light from external light sources to the FTS for analysis. ESO’s Integration and Cryo-vacuum Department in Instrumentation Division has built a permanent set-up for the external feed which replicates part of the optical train of the FTS. Originally, we
had planned to emulate the X-shooter calibration unit as closely as possible using all four pen ray lamps inside an integrating sphere. The PSUs for lamp operations on X-shooter provide alternating current (AC) only. The resulting modulated output of the lamps is not suitable for analysis by the FTS. Hence we had to operate the pen ray lamps one at a time using a special PSU that is switchable between AC and DC. Although built for the purpose by the lamp manufacturer, the DC provided by this PSU is not very stable and the resulting interferogram of the FTS shows increased noise due to the residual modulation of the input signal. When using individual lamps inside the integrating sphere the signal from any of the lamps was significantly too low for the FTS to work. Hence we decided to mount the lamps individually in a way that they directly illuminate the FTS feed. The lamps were oriented parallel to the optical path and with their tip facing the FTS feed. Alignment and focusing of the lamps without the integrating sphere is rather critical; we used the reflected light from the He-Ne laser-source of the FTS itself to find the optimum position of the lamps.

Figure 6.3: Fourier Transform Spectrometer (FTS) in the ESO laboratory. Labeled are parts of the experimental set-up for external sources.

6.2.2.4 Measurements

**Th-Ar Lamp**  Wavelength range and resolution of the FTS were chosen so as to match the X-shooter NIR arm. The lamp (Fig. 6.4) was carefully aligned with the optical path in order to see and maximize the signal on the FTS interferogram. Then we took several spectra of this lamp, varying both current and exposure time as given in Table 6.2. Long exposure times are essential in order to reach a reasonable S/N. A sample spectrum taken at 10 mA is depicted in Fig. 6.5.
6.2. Laboratory Measurements of Wavelength Calibration Sources in the near IR

Figure 6.4: Picture of the Th-Ar lamp during operations at the FTS

Figure 6.5: Spectrum of the Th-Ar hollow cathode lamp measured with the FTS
6.2. Laboratory Measurements of Wavelength Calibration Sources in the near IR

<table>
<thead>
<tr>
<th>Current [ma]</th>
<th>Exposure Time [hrs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.2: List of operating current and exposure times used for observations of the Th-Ar lamp

Pen ray lamps The lamps were operated with a AC/DC switchable power supply. Figure 6.6 shows all four lamps during operation. Their distinctly different visual colours are readily apparent. For each pen ray lamp using the same FTS settings as for Th-Ar measurements, we took several spectra varying both the operating current and exposure time. Sample spectra of the lamps are displayed in Fig. 6.7.

![Pen ray lamps](image)

Figure 6.6: Picture of the pen ray lamps during operations at the FTS; clockwise from upper left: Ne, Ar, Xe, Kr

6.2.3 Physical Modeling and Simulated Data

6.2.3.1 Physical Model

Traditionally the wavelength calibration of spectrographs relies upon an empirical approach. An exposure of a source, usually an emission lamp, with clear, laboratory-
6.2. Laboratory Measurements of Wavelength Calibration Sources in the near IR

Figure 6.7: FTS spectra of the Ne, Ar, Kr, Xe pen ray lamps at X-shooter spectral resolution
6.2. Laboratory Measurements of Wavelength Calibration Sources in the near IR

<table>
<thead>
<tr>
<th>Current [mA]</th>
<th>Exposure Time [secs]</th>
<th>Resolution [cm$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>28800</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>25200</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>600</td>
<td>0.125</td>
</tr>
<tr>
<td>20</td>
<td>600</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 6.3: Summary of the operating currents and exposure times used for measurements of the pen ray lamps

<table>
<thead>
<tr>
<th>Lamp</th>
<th>Usage [hrs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>33</td>
</tr>
<tr>
<td>Xe</td>
<td>25</td>
</tr>
<tr>
<td>Ar</td>
<td>25</td>
</tr>
<tr>
<td>Kr</td>
<td>25</td>
</tr>
<tr>
<td>Th-Ar</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 6.4: Usage time of the lamps accumulated during the laboratory measurements

calibrated features, is obtained. The location of features on this wavelength calibration exposure are then matched to the catalogued wavelengths of the source, and a low order polynomial is fitted to the data points to provide an empirical relation between positions on the detector and wavelengths. A meaningful polynomial fit will require a sufficient density of useful lines distributed over the wavelength range of interest. Since such an empirical polynomial fit has zero predictive value outside the range defined by data points, a lack of calibration lines at the limits of the wavelength ranges and detector boundaries is particularly critical. We replace this empirical method of wavelength calibration by using our physical understanding of the instrument (Rosa 1995). In order to avoid computationally expensive, but for the present purpose unnecessary level of detail, those surfaces which do not affect the relative geometry on the detector (e.g. plain folding or pick off mirrors) are neglected. However, the ray tracing at all relevant surfaces is performed by the appropriate 3D matrix transformation (Ballester & Rosa 1997). Since the main goal for the model is to provide wavelength calibration, the orientation, offsets and other properties of dispersive components can be optimised via comparison to calibration data from the instrument. At this stage we do not yet have such data, however for present purposes it is sufficient to assume design values for these parameters. The physical model approach as used for X-shooter wavelength calibration is described in greater detail in Kerber et al. (2006).
6.2.3.2 Simulated 2D Echellograms

The present study requires not only computation of where a photon of a given wavelength arrives on the detector, but also the spectrograph throughput (which we combine with the spectral line intensity) at that wavelength. This is achieved via empirical functions (i.e., they cannot be optimised in the way the physical model parameters can) that describe the quantum efficiency of the detectors and the throughput of the dichroics at each wavelength. In addition, we compute the grating blaze efficiency, $\epsilon$, for incident angle $\alpha$ and diffracted angle $\beta$ as:

\[
\begin{align*}
\text{if } \alpha & \geq \beta : \\
\epsilon_g &= \left(\frac{\sin A}{A}\right)^2 \\
A &= \frac{\pi}{\lambda \sigma} \left(\frac{\cos \alpha}{\cos \alpha - \nu_g}\right) [\sin \alpha - \nu_g + \sin \beta - \nu_g] \\
\text{if } \alpha & \leq \beta : \\
\epsilon_g &= \left(\frac{\sin A \cos \beta}{A \cos \alpha}\right)^2 \\
A &= \frac{\pi}{\lambda \sigma} \left(\frac{\cos \beta}{\cos \alpha - \nu_g}\right) [\sin \alpha - \nu_g + \sin \beta - \nu_g]
\end{align*}
\]

where $\nu_g$ is the blaze angle, $\sigma$ the groove spacing and $\lambda$ the wavelength. We ignore groove shadowing. In addition, we require a description of the entrance slit. The entrance slit model is rectangular, having a profile in the dispersion direction that is a 1.0" box function convolved with a 0.02" Gaussian and a profile in the cross dispersion direction that is a 12.0" box function convolved with a 0.2" Gaussian. We use a Monte-Carlo approach to simulate the counts expected on the detector array during an exposure. In this way 2D simulated data containing many photons are produced by multiple calls of the physical model code. We begin with a line list for an emission lamp that specifies wavelengths and relative intensities of spectral features (note that FF lamps can also be treated if a suitably binned continuum is given). We appropriately scale the intensities for the signal-to-noise that we want to achieve. Then, for each feature, the slit model is used to distribute the appropriate number of virtual photons in the entrance slit plane. Dichroic and blaze functions are used to compute the probability of a photon at this wavelength reaching the detector. At high flux levels (>1000 virtual photons) the fraction of virtual photons arriving on the detector is simply the calculated throughput multiplied by the flux; at low flux levels each photon is treated individually using a pseudo-random number generator. Using the model in this way we have generated 2D simulated data for known sources (emission and FF lamps as well as stellar models) that are being used as test cases for the development of the X-shooter data reduction software. An example of such simulated 2D data for all four pen ray lamps is shown in Fig. 6.8.
Figure 6.8: Simulated 2D X-shooter echellograms of pen ray lamps; clockwise from upper left: Ne, Ar, Xe, Kr

Figure 6.9: Simulated 2D X-shooter echellogram of a Th-Ar lamp
6.2.3.3 Extracted 1D spectra

In the current project, we take the 2D simulated data for each lamp and extract a 1D spectrum along the loci traced on the detector array by photons arriving at the centre of the entrance slit, in all orders. To improve statistics we include the counts from pixels within 20 pixels in the cross-dispersion direction.

As a result we have, for each lamp and each order, a 1D extracted spectrum. There is considerable overlap between the orders with features in the centre of the spectrum for one order appearing at the edge of an adjacent order but with a much reduced signal due to the blaze function. Figure 6.10 shows the blaze function and Fig. 6.11 shows a sample 1D spectrum (order #20).

![Blaze function of X-shooter NIR: order #20 and adjacent orders.](image1)

Figure 6.10: Blaze function of X-shooter NIR: order #20 and adjacent orders.

![Simulated 1D spectrum of X-shooter NIR order #20. The lines from the different pen ray lamps are colour coded. The flux level of the lamps has been normalized such that the strongest line in all of the lamps yields 100000 counts.](image2)

Figure 6.11: Simulated 1D spectrum of X-shooter NIR order #20. The lines from the different pen ray lamps are colour coded. The flux level of the lamps has been normalized such that the strongest line in all of the lamps yields 100000 counts.
6.2.4 Analysis and Results

6.2.4.1 Line Identifications

For the identification of the lines we have used compilations available in the literature and some data provided by our colleagues at NIST prior to publication. In the wavelength range of the X-shooter NIR arm (4000-10000 cm\(^{-1}\), 1000 - 2500 nm) we have identified in the observed FTS spectra a total of about 700 lines distributed between the sources as shown in Table 6.5: A preliminary analysis based on our FTS measurements (see Table 6.5 and Fig. 6.7) suggests that a combination of Ne, Ar, Kr and Xe will provide a number of lines (~ 350) comparable to the Th-Ar lamp while providing a more even distribution of line coverage for a given dynamic range. The Th-Ar spectrum includes a large number of faint Th lines (Fig 6.9) most of which have not been recovered with our FTS measurements. In order to improve the distribution of lines from a Th based HCL a different fill gas such as Ne or a mix of Ne and Ar might be interesting. We plan to investigate this option in the future but did not consider it for the current study.

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of lines observed with the ESO FTS</th>
<th>Number of lines available in the literature</th>
<th>Reference</th>
<th>Dynamic range in Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>80</td>
<td>350</td>
<td>Sansonetti et al. (2004)</td>
<td>1.5 \times 10^4</td>
</tr>
<tr>
<td>Ar</td>
<td>155</td>
<td>360</td>
<td>Whaling et al. (2002)</td>
<td>10^9</td>
</tr>
<tr>
<td>Kr</td>
<td>130</td>
<td>380</td>
<td>Sansonetti &amp; Greene (2007)</td>
<td>4 \times 10^4</td>
</tr>
<tr>
<td>Xe</td>
<td>65</td>
<td>120</td>
<td>Saloman &amp; Sansonetti (2004)</td>
<td>5 \times 10^3</td>
</tr>
</tbody>
</table>

Table 6.5: Summary of the number of lines observed on the FTS spectra

6.2.4.2 Order-by-order analysis

We have studied the number of lines for each lamp and each order of X-shooter individually. The following assumptions have been made in the analysis:

- In agreement with the envisaged scheme of operations (6.2.1) we have assumed that all lamps have been brought to the same intensity level by appropriate measures. That is, the strongest line in each lamp has been normalized to a value of 100000. Whether the same level can actually be achieved and by what means remains to be seen and is discussed in Sect. 6.2.4.6

- The dynamic range achieved in practice for a typical wavelength calibration exposure is limited by the linearity/saturation behaviour of the detector and is around 1000. Saturation of the strongest lines is avoided by using a short detector integration time (DIT). Then the faintest lines having a S/N suitable for use in wavelength calibration
will have an intensity of a few 100 on the above scale. A total wavelength calibration exposure will therefore combine a small number of integrations times (NDIT) and we adopted an intensity of 500 as the lower limit for lines suitable for wavelength calibration purposes.

These assumptions are in agreement with the exposure times envisaged for X-shooter wavelength calibration exposures of order a few seconds total integration time. For the analysis and comparison of the lamps we have made the following reasonable but somewhat arbitrary assumptions:

- A line suitable for wavelength calibration has an intensity in the range 500 - 100000.
- A minimum of 10 lines is required for wavelength calibration in any order.
- A goal of 30 lines is set for wavelength calibration in any order.

In the following tables we summarize the result of this analysis for each lamp and order. We used a standard colour code to visualise the results (Table 6.6). Full details of the analysis and a spectral atlas are presented in the appendix. In the tables the number

<table>
<thead>
<tr>
<th># of lines per lamp</th>
<th># of lines-complete calibration system</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-3</td>
<td>&lt;10</td>
</tr>
<tr>
<td>4-6</td>
<td>10-19</td>
</tr>
<tr>
<td>7-9</td>
<td>20-29</td>
</tr>
<tr>
<td>≥10</td>
<td>≥30</td>
</tr>
</tbody>
</table>

Table 6.6: Requirement and goals for number of calibration lines provided

of lines available from each lamp is given. The total includes blended lines, while the net gives the number of unblended lines accounting also for blending across lamps. This net number is the best current estimate of the lines usable for wavelength calibration. While the simulation provides quantitative results, the actual numbers observed by X-shooter may differ depending on the actual performance of the lamps and instrument. Validation of the numbers will be one objective of the laboratory testing and commissioning phases.

6.2.4.3 Summary of two lamp solutions

No combination of any two lamps provides a spectrum that satisfies the calibration needs of X-shooter. This is illustrated by the distribution of colours in tables 6.8,6.9,6.10.
## Laboratory Measurements of Wavelength Calibration Sources in the near IR

### Table 6.7: Summary of the number of lines provided by each pen ray lamp for a given order

<table>
<thead>
<tr>
<th>Order# ($\lambda$ [nm])</th>
<th>Ne</th>
<th>Ar</th>
<th>Kr</th>
<th>Xe</th>
<th>Total</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 (2250-2500)</td>
<td>17</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>32</td>
<td>27</td>
</tr>
<tr>
<td>12 (2060-2270)</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>13 (1910-2110)</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>14 (1770-1960)</td>
<td>18</td>
<td>2</td>
<td>10</td>
<td>2</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td>15 (1650-1825)</td>
<td>5</td>
<td>3</td>
<td>16</td>
<td>5</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>16 (1545-1710)</td>
<td>0</td>
<td>8</td>
<td>14</td>
<td>6</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>17 (1455-1610)</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>6</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>18 (1375-1520)</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td>6</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>19 (1300-1440)</td>
<td>1</td>
<td>16</td>
<td>15</td>
<td>3</td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>20 (1235-1370)</td>
<td>5</td>
<td>21</td>
<td>10</td>
<td>1</td>
<td>37</td>
<td>30</td>
</tr>
<tr>
<td>21 (1180-1305)</td>
<td>15</td>
<td>14</td>
<td>6</td>
<td>2</td>
<td>37</td>
<td>30</td>
</tr>
<tr>
<td>22 (1125-1245)</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>23 (1075-1190)</td>
<td>12</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>24 (1040-1140)</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>25 (1015-1095)</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>26 (995-1055)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6.8: Summary of the lines provided by a combination of Ne and Ar pen ray lamp for a given order

<table>
<thead>
<tr>
<th>Order# ($\lambda$ [nm])</th>
<th>Ne</th>
<th>Ar</th>
<th>Total</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 (2250-2500)</td>
<td>17</td>
<td>2</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>12 (2060-2270)</td>
<td>7</td>
<td>3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>13 (1910-2110)</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>14 (1770-1960)</td>
<td>18</td>
<td>2</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>15 (1650-1825)</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>16 (1545-1710)</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>17 (1455-1610)</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>18 (1375-1520)</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>19 (1300-1440)</td>
<td>1</td>
<td>16</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>20 (1235-1370)</td>
<td>5</td>
<td>21</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>21 (1180-1305)</td>
<td>15</td>
<td>14</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>22 (1125-1245)</td>
<td>9</td>
<td>10</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>23 (1075-1190)</td>
<td>12</td>
<td>6</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>24 (1040-1140)</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>25 (1015-1095)</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>26 (995-1055)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Laboratory Measurements of Wavelength Calibration Sources in the near IR

<table>
<thead>
<tr>
<th>Order# (λ [nm])</th>
<th>Ne</th>
<th>Kr</th>
<th>Total</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 (2250-2500)</td>
<td>17</td>
<td>4</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>12 (2060-2270)</td>
<td>7</td>
<td>3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>13 (1910-2110)</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>14 (1770-1960)</td>
<td>18</td>
<td>10</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>15 (1650-1825)</td>
<td>5</td>
<td>16</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>16 (1545-1710)</td>
<td>0</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>17 (1455-1610)</td>
<td>5</td>
<td>12</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>18 (1375-1520)</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>19 (1300-1440)</td>
<td>1</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>20 (1235-1370)</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>21 (1180-1305)</td>
<td>15</td>
<td>6</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>22 (1125-1245)</td>
<td>9</td>
<td>6</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>23 (1075-1190)</td>
<td>12</td>
<td>4</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>24 (1040-1140)</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>25 (1015-1095)</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>26 (995-1055)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.9: Summary of the lines provided by a combination of Ne and Kr pen ray lamp for a given order

<table>
<thead>
<tr>
<th>Order# (λ [nm])</th>
<th>Ar</th>
<th>Kr</th>
<th>Total</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 (2250-2500)</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>12 (2060-2270)</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>13 (1910-2110)</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>14 (1770-1960)</td>
<td>2</td>
<td>10</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>15 (1650-1825)</td>
<td>3</td>
<td>16</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>16 (1545-1710)</td>
<td>8</td>
<td>14</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>17 (1455-1610)</td>
<td>7</td>
<td>12</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>18 (1375-1520)</td>
<td>7</td>
<td>11</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>19 (1300-1440)</td>
<td>16</td>
<td>15</td>
<td>31</td>
<td>28</td>
</tr>
<tr>
<td>20 (1235-1370)</td>
<td>21</td>
<td>10</td>
<td>31</td>
<td>28</td>
</tr>
<tr>
<td>21 (1180-1305)</td>
<td>14</td>
<td>6</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>22 (1125-1245)</td>
<td>10</td>
<td>6</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>23 (1075-1190)</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>24 (1040-1140)</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>25 (1015-1095)</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>26 (995-1055)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6.10: Summary of the lines provided by a combination of Ar and Kr pen ray lamp for a given order
6.2.4.4 Summary of three lamp solutions

A combination of Ne, Ar and Kr is the best solution using three calibration lamps. Since Xe is not a good substitute for any of the other three lamps we only provide details (Table 6.11) for this best solution. Comparison with Table 6.7 shows that the combination of Ne, Ar and Kr yields a spectrum almost as good as the full four lamp solution.

![Table 6.11: Summary of the lines provided by a combination of Ne, Ar and Kr pen ray lamp, the best three lamp solution.](image)

6.2.4.5 Relative Intensities

The FTS is not an ideal tool for measuring overall intensities of lines. For the current purpose we only need intensities of the lamps relative to each other and we derived these from the integral of the line fluxes in the region of interest. There is a factor of more than an order of magnitude in intensity between the faintest and the most intense source. In X-shooter the fluxes need to be brought to the same level in the integrations sphere since all calibration sources will be operated simultaneously, see Section 6.2.1. One way to achieve this is to use different operating currents since each lamp has its own PSU.

6.2.4.6 Operation Current

We observed all sources at various currents in order to quantify its effect on the spectral output. There is a pronounced difference in the behaviour of the pen ray lamps and the
6.2. Laboratory Measurements of Wavelength Calibration Sources in the Near IR

<table>
<thead>
<tr>
<th>Lamp</th>
<th>Relative Integral Intensity (Ne=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>1.0</td>
</tr>
<tr>
<td>Ar</td>
<td>3.3</td>
</tr>
<tr>
<td>Kr</td>
<td>1.9</td>
</tr>
<tr>
<td>Xe</td>
<td>0.3</td>
</tr>
<tr>
<td>Th-Ar at 10 ma</td>
<td>1.05</td>
</tr>
</tbody>
</table>

Table 6.12: Relative overall intensities of the lamps based on the integrated line flux

HCLs. This is caused by the fact that pen ray lamps are pure gas discharges while in the HCL the gas discharge drives sputtering of the methods which then causes emission from metal lines. Higher currents lead to higher sputtering rates and hence to an increase in the intensities of the Th lines. A more detailed explanation is given in Kerber et al. (2007). For all pen ray lamps some increase in output is observed for higher currents. In the case of Kr this amounts to a factor of 2 in intensity for a change of current from 5 to 20 mA. For the other three lamps the effect is considerably smaller. We find that the operating current is a parameter that can be easily tuned to vary the relative intensity of the lamps but the effect is limited and most of the attenuation required to bring the individual lamps to the same intensity will have to be achieved by other means. Based on these results we decided to take spectra with long exposures with an operating current of 10-15 mA for all pen ray lamps.

![Figure 6.12: Variation of the relative line intensities in a Th-Ar HCL as a function of current](image)

6.2.4.7 Th-Ar

Most of the strong lines in the Th-Ar spectrum are from Ar, hence the overall distribution of lines closely follows the one of the Ar pen ray lamp, see Fig. 6.7 and Table 6.7. In
Figure 6.13: Intensity variation of selected lines as a function of operating currents in pen ray lamps
addition there is a multitude of faint lines mostly from Th (Fig. 6.9 & 6.14). While operations at higher current would result in a very significant increase in the intensity of the faint Th lines, this will also result in blending of many lines at X-shooter resolution. In summary a combination of Ne, Ar, and Kr provides a much more even distribution of lines than the Th-Ar lamp. If the faint Th lines are enhanced by using a higher current, blending of lines will become a problem at X-shooter resolution in the NIR. For the UVB and in particular the VIS arm a similar analysis will be valuable to assess the possible impact of blending in these wavelength domains as well. A previous study de Cuyper & Hensberge (1998) suggests that blending will become a major issue for resolutions below 20,000.

Figure 6.14: Sample spectrum of Th-Ar showing a forest of faint Th lines. At the spectral resolution of X-shooter many of these lines will be blended.

6.2.4.8 Results

The following result have been derived from our measurements:

- The spectra of pen ray lamps filled with Ne, Ar, Kr, and Xe have been measured at a spectral resolution equivalent to the one of X-shooter in the NIR. All lamps show sharp, narrow lines in emission. Continuum emission is minimal with the exception of Xe.
- The number of lines and their intensities available from each lamp have been measured and recorded in electronic form. Analysis shows that almost all lines are from neutral atoms and that line intensities are similar to the ones reported in the literature.
The relative intensities of the lamps have been derived as an integral of the line fluxes measured with the FTS. Ne, Ar and Kr are within a factor of three in intensity while Xe is another 3 times fainter than the next faintest source.

No combination of any two lamps yields a satisfactory spectrum for X-shooter wavelength calibration.

A combination of Ne, Ar and Kr offers the best three lamp solution meeting the requirement of 10 lines per order for all but orders 13 and 26. It also approaches the goal of 30 lines per order for many orders and provides a suitable spectrum for X-shooter calibration.

The addition of Xe (four lamp solution) will only bring a small overall improvement in the number of lines and coverage. For some orders the improvement would be significant. Since Xe is the faintest source its addition depends on practical considerations such as feasibility of attenuation and added complexity.

The combined spectrum of Ne, Ar and Kr provides a number of lines suitable for X-shooter wavelength calibration while providing a more even distribution across the spectral range than a Th-Ar lamp. In addition, Th has a large number of faint lines many of which are blended at X-shooter resolution.

Operating current is a straightforward but not very effective means to control the relative intensity of the lamps. Clearly, most of the attenuation needed to achieve equal intensity has to be done by hardware measures.

Allowing saturation of a few of the strongest lines will result in a net increase in the number of lines available in a given dynamic range of the detector.

**6.2.4.9 Remaining Issues**

Two orders - #13 and #26 - fall short of the requirement of at least 10 lines available for calibration purposes.

Order 26 is not intrinsically poor in lines but is affected by the performance of the dichroic at the short wavelength end. With longer exposure times it will be possible to reach several fainter lines. This needs to be assessed during Assembly, Integration and Testing (AIT).

For order 13 the situation is different. The region between 1880 and 2015 nm (orders 13 and 14) is almost devoid of lines leaving no easy option to the calibration. For a dispersion solution based on an empirical 2D polynomial fit this could be a difficult situation. Since X-shooter will be using a physical model to drive the wavelength calibration in the pipeline we don’t expect any significant negative impact on the accuracy of calibrations. Furthermore, the region is strongly affected by atmospheric absorption limiting its scientific use.

Blending of lines is significant in a Th-Ar source in the NIR at X-shooter resolution. Further work is needed to assess this issue in the UVB and VIS region.

It is necessary to investigate the lifetime of the pen ray lamps e.g. by means of an accelerated aging test.
6.2.4.10 Lessons Learned

A combination of laboratory measurements and simulated data based on a physical model provides a very powerful tool to estimate the performance of a calibration system for an instrument. To our knowledge this is the first time such an analysis has been done at this level of detail. There is no comprehensive study or data-base of emission line spectra provided by commercially available calibration lamps. For the heavily used and well studied noble gases the input data were available from NIST. A spot-check for other elements of interest showed very significant gaps in our knowledge of their spectra in particular at near IR wavelengths. In the context of an Extremely Large Telescope (ELT) instrumentation, establishment of such a database will be highly valuable to provide input for studies similar to this one. With our approach a quantitative prediction of wavelength calibration data can be made before an instrument is build. This opens up new opportunities to optimize a calibration system during the design phase and validate it during testing and commissioning of the instrument. This also has the benefit of reducing the pressure on the very busy AIT and commissioning phases.

6.3 Building up a database of spectro-photometric standard stars from the UV to the near IR

While there is a well-established and widely used set of high-quality optical spectro-photometric standard stars (in particular Oke (1990); Hamuy et al. (1992, 1994)), we still lack an equivalent set for the NIR. Observations of white dwarfs with HST spectrographs and atmospheric modeling have established 3 stars as primary standards (1% accuracy) for the wavelength range 115-1800 nm (Bohlin 2007). A set of NIR spectro-photometric standards adequate for routine operations of ground based NIR spectrographs simply does not exist at the moment. The current used methods for estimating the absolute flux calibration of NIR spectra are:

(i) use the NIR broad band magnitudes of the telluric standard that have been observed for a given science exposure (eg. from 2MASS);

(ii) alternatively, when it is known, use the broad-band magnitude of the scientific target itself to scale the NIR spectrum.

These methods are usually no more precise than 20-30%. In addition, from a point of view of operational efficiency and for consistency of the calibration over the 3 arms of X-shooter (UV-blue, Visual-Red and NIR) it would be much preferable to calibrate the whole wavelength range using one single spectro-photometric standard. We therefore proposed to extend the wavelength range of a set of well established optical spectro-photometric standards to the NIR by combining dedicated high-quality SINFONI observations with state-of-the-art stellar atmosphere models. Note that undertaking such a program has been marked as crucial for the scientific operations of X-shooter both by the Preliminary
Design Preview and the Final Design Review boards. It is important to stress that with only a modest investment in observing time, the quality of the scientific products of X-shooter, but also of all the present and future ESO NIR spectrographs (ISAAC, SINFONI, NACO, K-MOS), will be significantly improved. When published, this set of standards is aimed at becoming the main reference for NIR spectro-photometric calibration in the southern hemisphere and will therefore benefit to the entire astronomical community. In the longer term, it is clear that establishing a robust set of NIR spectro-photometric standard stars is absolutely necessary in order to be able to calibrate future instrumentation associated with ELTs that are expected to carry out their core science in this wavelength range.

6.3.1 Choice of the Instrument

The two main difficulties in reliably measuring absolute flux in the NIR are the variable telluric absorption and the strongly variable sky OH emission lines. There are actually very few regions completely free of telluric lines. Even the classical J, H and K atmospheric windows are seriously affected by telluric absorptions as can clearly be seen in Fig. 6.15 (J band) and 6.16 (H+K bands). OH sky lines on the other hand are present throughout the NIR wavelength range up to \( \sim 2 \mu m \). In order to peer through these forests of lines and perform robust photometric measurements we need to observe at a sufficiently high spectral resolution. By carefully examining the J, H and K band spectra using a \( R \sim 100000 \) Gemini atmospheric transmission spectrum with the list of identified OH sky lines (Rousselot et al. 2000), we found that a resolution better than \( R \sim 1000 \) is required. The other constraint to the choice of our instrument is the need for a wide aperture to reliably collect the whole flux. In order to avoid slit-losses and to minimize the effect of centering uncertainties, spectro-photometric standards are usually observed using wide 5" slits or even in slitless mode. Combining the requirements of resolution with that of wide aperture led us to select SINFONI IFU as the instrument of choice for our program. It provides a resolution of \( R \sim 2000 \) in the J band and \( R \sim 1500 \) in the H+K bands. It is a rather efficient setup since only 2 shots are needed to cover the whole wavelength range. SOFI does not have any slit wider than 2" at the moment and the highest resolution achievable would be \( \sim 130 \). Similarly, with its widest 5" slit, ISAAC would provide way too low resolution (\( \sim 180 \)). Note that the observations do not require good seeing conditions (1.4" acceptable) and can be performed during bright time and even twilight.

6.3.2 Observational Strategy

In order to establish our standard stars in the NIR we need to reliably measure the fluxes in the few carefully chosen atmospheric windows available. In addition, we need to apply flux models in order to cover the full wavelength range. The best options are hydrogen white dwarfs (WDs) and very hot stars which have very few absorption lines almost exclusively from hydrogen and helium. State-of-the-art models of stellar atmosphere have
Figure 6.15: Telluric absorption in the J band at the resolution of SINFONI (R=2000). Shaded areas show regions almost free of telluric absorption that will be used to perform our absolute flux measurements. The dashed line shows the featureless spectrum of one of our primary standards GD71.

Figure 6.16: Similar to Fig. 6.15 in the H+K wavelength range (SINFONI resolution R=1500).
been demonstrated to quantitatively describe the physical properties of such stars with high accuracy (Bohlin 2007).

Our strategy is therefore the following:

- Use as primary Standards (PrimS) three HST white dwarfs observable from the southern hemisphere: GD71, GD153 and BD+17D4708. These HST PrimS provide a uniquely robust reference frame because the measured relative flux between these three stars agrees with the ratio of their model atmosphere SEDs to $\sim 1\%$ from 115 to 1800 nm. The models are normalized to the precision Landolt V mag. (Bohlin 2007). These will be our reference to calibrate the response of SINFONI.

- We have selected a set of 12 secondary standards (SecS) by choosing WDs and very hot stars from Oke (1990) and Hamuy et al. (1992, 1994) optical spectro-photometric standards. We have restricted ourselves to stars brighter than 14th mag in the V band in order to limit the amount of exposure time required. Based on the predicted efficiencies of X-shooter, we have also checked that none of these stars would saturate X-shooter detectors for integration time ($\sim 5$s) suitable for flux calibration. We have selected targets to provide a homogeneous distribution in right ascension (RA). Therefore, collecting data for the whole sample requires observations spread over a full year.

- Each SecS should be observed at 5-10 times to build up a robust observational base for their use as standard stars and to allow a proper statistical treatment of the errors on the measured flux. For each observation, one or more PrimS must be observed.

- All observations will be reduced using the SINFONI pipeline and total flux spectra will be extracted from data cubes using wide apertures. For each run, at least one PrimS will be observed to deduce a response function by comparing flux levels in wavelength regions free of atmospheric absorption and sky emission with tabulated values in the HST calibrations data-base. This response will then be used to calibrate the SecS using the same clean wavelength bins.

- In a related effort, a Co-I of the proposal (E. Mason) and a student are finalizing a study of the NIR photometric zero points for Paranal based on ISAAC data taken 2000-2005. We plan to supplement this effort by observing one star at several airmasses on one occasion. In combination, these data sets will enable us to properly describe the effect of extinction as a function of airmass for our standard stars.

- All the proposed targets have been observed with UVES. These spectra will be reduced and measures of their absorption-lines will allow an accurate determination of stellar parameters to be given as an input to planeparallel, static, and chemically homogeneous models computed with the Tübingen NLTE Model Atmosphere Package (Werner et al. 2003; Rauch & Deetjen 2003). These models will then be used to interpolate in between observed fluxes to derive an absolute flux table for each SecS across the full wavelength range.

In Tab. 6.13 we give our sample of standard stars.
### Table 6.13: Standard stars to be observed for the project together with their coordinates and V band magnitudes

<table>
<thead>
<tr>
<th>Name</th>
<th>( \alpha (\text{J2000}) )</th>
<th>( \delta (\text{J2000}) )</th>
<th>Mag.</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>GD71</td>
<td>05 52 27.6</td>
<td>+15 53 13</td>
<td>13.032</td>
<td>PrimS</td>
</tr>
<tr>
<td>GD153</td>
<td>12 57 02.3</td>
<td>+22 01 52</td>
<td>13.346</td>
<td>PrimS</td>
</tr>
<tr>
<td>BD+17D4708</td>
<td>22 11 31.4</td>
<td>+18 05 34.17</td>
<td>9.47</td>
<td>PrimS</td>
</tr>
<tr>
<td>Feige110</td>
<td>23 19 58.4</td>
<td>-05 09 56</td>
<td>11.82</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT1020</td>
<td>01 54 50.3</td>
<td>-27 28 36</td>
<td>11.54</td>
<td>SecS</td>
</tr>
<tr>
<td>EG21</td>
<td>03 10 29.0</td>
<td>-68 35 54</td>
<td>11.38</td>
<td>SecS</td>
</tr>
<tr>
<td>Hiltner600</td>
<td>06 45 13.4</td>
<td>+02 08 15</td>
<td>10.37</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT45-46</td>
<td>07 40 20.8</td>
<td>-17 24 49</td>
<td>12.98</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT3218</td>
<td>08 41 36.7</td>
<td>-32 57 41</td>
<td>12.00</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT4364</td>
<td>11 45 22.0</td>
<td>-64 50 14</td>
<td>11.50</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT4816</td>
<td>12 38 48.1</td>
<td>-49 49 00</td>
<td>13.96</td>
<td>SecS</td>
</tr>
<tr>
<td>CD-32 9927</td>
<td>14 11 46.3</td>
<td>-33 03 14</td>
<td>10.33</td>
<td>SecS</td>
</tr>
<tr>
<td>EG 274</td>
<td>16 23 33.3</td>
<td>-39 13 45</td>
<td>10.98</td>
<td>SecS</td>
</tr>
<tr>
<td>LTT7998</td>
<td>20 10 56.8</td>
<td>-30 13 06</td>
<td>12.18</td>
<td>SecS</td>
</tr>
<tr>
<td>EG131</td>
<td>19 20 34.9</td>
<td>-07 40 00</td>
<td>12.28</td>
<td>SecS</td>
</tr>
</tbody>
</table>

The project is approved and observations are being taken. Once released, data are reduced using the ESO pipeline for the SINFONI instrument and analyzed to get the efficiency of the instrument, necessary to calibrate the absolute flux of the stars. The
6.3. Building up a database of spectro-photometric standard stars from UV to NIR

<table>
<thead>
<tr>
<th>Atmospheric window</th>
<th>$\lambda_{\text{start}}(\mu m)$</th>
<th>$\lambda_{\text{end}}(\mu m)$</th>
<th>$\lambda_{\text{center}}(\mu m)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>1.2329</td>
<td>1.2474</td>
<td>1.24015</td>
</tr>
<tr>
<td>H-1</td>
<td>1.5500</td>
<td>1.5600</td>
<td>1.55500</td>
</tr>
<tr>
<td>H-2</td>
<td>1.5885</td>
<td>1.5950</td>
<td>1.59175</td>
</tr>
<tr>
<td>H-3</td>
<td>1.6250</td>
<td>1.6300</td>
<td>1.62750</td>
</tr>
<tr>
<td>H-4</td>
<td>1.6720</td>
<td>1.6890</td>
<td>1.68050</td>
</tr>
<tr>
<td>K-1</td>
<td>2.1345</td>
<td>2.1480</td>
<td>2.14125</td>
</tr>
<tr>
<td>K-2</td>
<td>2.1550</td>
<td>2.1780</td>
<td>2.16650</td>
</tr>
<tr>
<td>K-3</td>
<td>2.1890</td>
<td>2.1940</td>
<td>2.19150</td>
</tr>
</tbody>
</table>

Table 6.14: Atmospheric transmission windows used to compute photometric points

efficiency is measured from the PrimS by means of the following relation:

$$\epsilon = \frac{F_{\text{obs}}}{F_{\text{exp}}S\Omega\Delta\lambda}.$$  \hspace{1cm} (6.3.1)

That is the ratio between the number of observed and expected photons by HST. $F_{\text{obs}}$ is the number of observed photons from the ground, $F_{\text{exp}}$ is the number of photons above the atmosphere, measured by HST, $S$ the collecting area, $\Omega$ the solid angle subtended by the integration element, and $\Delta\lambda$ the spectral bin. Fig. 6.17 shows the status of the observations at the moment. For each star, dark green shows validated photometric data, dark red shows observations that were carried out in non-photometric conditions, light green shows observations that were not yet released to us, and light red shows the number of observations remaining to reach the goal of 10 observations per Secondary Standard. Tab. 6.14 shows the set of eight photometric windows where the atmospheric transmission is greater than 99% and free of very strong OH sky lines.

### 6.3.3.1 Efficiency of SINFONI

Using all the photometric measurements of our Primary Standards, we have derived the overall system efficiency of SINFONI+Telescope in our eight atmospheric transparent windows. Results are displayed in figure 6.18. Our results are in reasonably good agreement with Commissioning+detector upgrade measurements in the J and H bands. Our measured efficiency in the K band is significantly higher ($\sim30\%$ vs. $\sim25\%$).

### 6.3.3.2 Photometry

So far, only one Secondary Standard star, Hiltner 600, has been observed more than three times in photometric conditions. For this one star, our first analysis gives very good results. The 6 measurements we have obtained in the J window agree within 8% and within 5% in H and K windows peak to valley. Dispersion ($1\sigma$) is within 3.5% in all windows. These results—although very preliminary—clearly demonstrate feasibility of our
6.3. Building up a database of spectro-photometric standard stars from UV to NIR

Figure 6.18: Overall system efficiency. Red dots show our measurements. Black line shows efficiency measured at Commissioning (with engineering grade detector) corrected to take into account the improved efficiency of the science grade detector.

program. Our initial goal of 10% accuracy should easily reached. Fig. 6.19 and fig. 6.20 show the results for this star.

Figure 6.19: Spectrum of our secondary standard Hiltner600 from the UV to the NIR. Our photometric points are the squares while the UV/Optical points are from Hamuy et al. (1992, 1994).
Figure 6.20: The 6 flux measurements we have obtained in our H-3 window for Hiltner600. The solid line indicates the mean flux and the dotted lines mark ±5% deviation from the mean.
Conclusions

During my thesis work I have studied the properties of the IGM, using both the Lyα and Metal-line forests, observed in high resolution spectra of QSOs at emission redshift $z_{em} \sim 2 - 3$. The two main points I have addressed are:

1. the cosmological nature of the Lyα forest and the properties of the underlying hydrogen density field as a tracer of the matter distribution in the universe at those redshifts;
2. the metal enrichment of the IGM, looking at the connection between metal absorption systems in QSO spectra and galaxies in the same field.

I have also participated to an instrumental project, the building of a new spectrograph for the ESO VLT, called X-shooter, studying the contribution of this new instrument to the IGM science, carrying out laboratory measurements of calibration sources for the Near-Infrared (NIR) arm (1-2.5$\mu$m) and participating to an observational project to build a spectro-photometric flux catalogue of standard stars for the instrument.

The Classical Lyα forest statistics

We have considered a sample of 22 high resolution, high SNR QSO spectra covering the Lyα redshift range between $\sim 1.7$ and $3.5$. Each spectrum has been fitted with Voigt profiles resulting in the largest homogeneous collection of fitted Lyα lines ever studied ($\sim 8435$ fitted Lyα lines). I have applied the classical statistics of the Lyα forest to this sample obtaining:

1. the line number density evolution with redshift, $dn/dz \simeq (166 \pm 4)((1+z)/3.5)^{2.8 \pm 0.2}$. While the redshift evolution is consistent with previous results, the normalisation is higher by a factor ranging from $\sim 0.03$ in $\log(dn/dz)$ at $z \sim 2$ to 0.1 at $z \sim 3$. 
This difference is due to the improved treatment of the contamination by metal lines (amounting to $\sim$ 9 percent of the redshift interval covered by the Ly$\alpha$ forests), which is made possible by the high resolution and signal-to-noise ratio of our spectra. Consistently with Kim et al. (2002), we also find a steeper evolution for the stronger lines ($14.5 \leq \log N(\text{H}i) < 17$) compared to the weak ones ($13 \leq \log N(\text{H}i) < 14$).

2. the TPCF, which shows a significant clustering signal up to $\sim 2$ comoving Mpc for strong lines ($13.8 \leq \log N(\text{H}i) < 17$), and, at variance with previous results, also for weak lines ($12 \leq \log N(\text{H}i) < 13.8$) although on smaller scales, $\lesssim 1.5$ Mpc. We then calculated the TPCF by grouping all the lines closer than the local Jeans length (the assumed typical size for the hydrogen absorbers in the IGM). The signal is still significant in the first bin ($r \lesssim 2.5$ Mpc), indicating that Ly$\alpha$ absorbers cluster among themselves and not only inside themselves.

3. the clustering evolution with redshift of strong lines. The QSO sample has been divided in two sub-samples, the first one formed by objects with $z_{\text{em}} \lesssim 2.5$, for which the average Ly$\alpha$ forest redshift is $\langle z_{\text{Ly}\alpha} \rangle = 2.07$, and the second one formed by objects with $z_{\text{em}} > 2.5$, with $\langle z_{\text{Ly}\alpha} \rangle = 3.02$. The TPCF shows a trend of increasing clustering with decreasing redshift; this is an apparent evolution, due to the fact that the relation $\delta - \log(N(\text{H}i))$ is $z$-dependent, see eq.3.1.1. Indeed, the evolution signal is no longer significant when lines arising in the same kind of structures ($\delta > 3$) are selected.

**FLO, a new approach to the IGM**

We have described a new algorithm to transform the measured H$\text{i}$ column densities of the Ly$\alpha$ lines detected along a line of sight, into the underlying total H density field (and in particular, the density contrast, $\delta \equiv n_\text{H}/\langle n_\text{H} \rangle - 1$). The method is based on the assumption that Ly$\alpha$ absorbers are in local hydrostatic equilibrium and, as a consequence, the Jeans length corresponds to their characteristic size. The aim of this study is to find a robust statistical estimator which allows a direct link to the physical properties of the gas and an easy comparison with the results of simulations. To test the effects of the transformation, we have used a set of lines of sight obtained from a large N-body hydrodynamical simulation run. For every line of sight we have both the density and velocity field pixel per pixel and the list of Voigt fitted Ly$\alpha$ lines with central redshift, column density and Doppler parameter. Our results can be summarised as follows:

1. FLO recovers well (within 30 percent) the over-densities up to $\delta \sim 30$ while it is not reproducing correctly the under-densities (more than 50 percent of $\delta$ values are not recovered correctly) even in the range above our resolution limit. This result suggests that the hypothesis of hydrostatic equilibrium is not valid for the under-dense regions that are likely still expanding. On the other hand, as long as the computation of statistical properties of the IGM is concerned, the resulting $\delta$ field gives satisfactory results when the two-point correlation function and the 1D power spectrum are considered.
When applied to the observed data sample, the FLO algorithm gives the following results:

2. the redshift distribution of the average hydrogen density is consistent with the evolution of the cosmic mean hydrogen density in the redshift range covered by our QSO sample, supporting the fact that the Ly$\alpha$ forest arises from fluctuations of the IGM close to the mean density;

3. the correlation function of the density field obtained from the observed spectra shows a significant clustering signal up to $\sim 4$ comoving Mpc and is consistent with the analogous result obtained for the recovered density field in simulations.

4. the one-dimensional power spectrum of the $\delta$ field obtained from the observed spectra is in agreement with that obtained from the recovered density field from the simulation based on the concordance cosmological model, on scale lengths between $\sim 2.5$ and 63 comoving Mpc.

5. the proximity effect is observed in our data sample at distance $r \lesssim 15$ Mpc from the QSO as a decrease in the density contrast field.

6. when taking into account the effect of the QSO ionizing flux in the computation of the density contrast field, FLO recovers the overdense environment hosting the QSO.

**Tomography of the IGM**

We exploited the capabilities of the largest sample of high-resolution UVES spectra of QSO pairs to study the properties of the 3-dimensional distribution of baryonic matter in the IGM as traced by the transmitted flux in the QSO H$\textup{i}$ Ly$\alpha$ forests, using a sample formed by 21 QSO pairs evenly distributed between angular separations of $\sim 1$ and 14 arcmin, whose Ly$\alpha$ forests are at a median redshift $z \simeq 1.8$. We selected also 8 UVES QSO spectra from the ESO Large Program ‘The Cosmic Evolution of the IGM’ (Bergeron et al. 2004) to compute the correlation function along the line of sight and to be used as a control sample for the cross correlation function. We compared the observed sample with a set of mock spectra drawn from a cosmological hydro-simulation run. The simulated sample reproduces 50 different realisations of the observed sample. From this analysis we obtained the following results:

1. The clustering properties of matter in the IGM are the same in the direction parallel and transverse to the line of sight when using the parameters of the concordance cosmology to transform the angular distance into velocity separation. As an implication, peculiar velocities in the absorbing gas are likely smaller than $\sim 100$ km s$^{-1}$.

2. Matter in the IGM is clustered on scales smaller than $\sim 200$ km s$^{-1}$ or about 3 $h^{-1}$ comoving Mpc. We verified that the clustering signal is significant also for the slightly over-dense gas ($\tau_{\textup{H}1} \gtrsim 2.3$ or $\rho/\bar{\rho} \lesssim 6.5$) although on smaller scales.

3. The simulated correlation functions are consistent with the observed analogous quantities at the $3\sigma$ level, although they systematically predict lower clustering at the
smaller scales. The agreement becomes better when only the lower density regions are selected for the computation or when the effective optical depth of the simulated spectra is fixed to a larger value (marginally consistent with previous extensive observational results on the redshift evolution of the effective optical depth). These are indications of a deficiency of strong absorption lines in the simulated spectra that needs further investigation.

4. We observed an improved consistency between observations and simulations also when a lower $\gamma$ is adopted in the equation of state of the gas, $T = T_0 (\rho/\bar{\rho})^{\gamma-1}$ ($\gamma = 1.1$ instead of the standard 1.6). This result hints to a late He II reionization epoch whose effects on the IGM could still be measured at the redshifts investigated here.

### The metal enrichment of the IGM

We have proposed to observe with UVES at the ESO VLT a close pair of QSOs at redshift $z_{\text{em}} \sim 1.8$ (J225153 and J225154), to get their high resolution spectra, look for matching metal systems and eventually find the corresponding galaxies in the pair field, to study the metal enrichment mechanisms in the IGM. The proposal was accepted and I went to Chile, at the ESO Paranal observatory, to perform the observation. In the spectrum of J225153 we found two associated systems ($z_{\text{abs}} \approx z_{\text{em}}$) identified through the N V doublet at $z_{\text{abs}} \approx 1.7605$ and 1.7654. The latter system shows also the Si IV doublet. The spectrum of J225154 shows three intervening C IV absorptions, at $z_{\text{abs}} \approx 1.357$ (strong), 1.655 (strong) and 1.707, but none of them has a corresponding system in the spectrum of J225153. However, we also detected a weak Si IV at $z_{\text{abs}} \approx 1.76794$ separated by only $\sim 270$ km s$^{-1}$ from the strong associated system in the paired spectrum. To understand these results we still need observations, so we asked for UVES time to get high resolution spectra to understand the physical and chemical properties of the detected system at $z \sim 1.77$; moreover we asked for FORS2 imaging of the pair field to identify galaxies close to the line of sight and study their possible influence on the found metal absorptions.

### X-shooter

I have presented X-shooter, the new ESO spectrograph to be commissioned in Summer 2008, whose capabilities will be unique to study the Ly$\alpha$ forests of close pair of QSOs at $z \sim 2 - 3$. The sensitivity of this instrument will allow a big observational campaign to collect enough data to apply the Alcock-Paczyński test and constrain the geometry of the universe (through the cosmological parameter $\Omega_A$). I have reported also the work done at ESO for the X-shooter project:

1. The laboratory measurements of calibration sources for the NIR arm of X-shooter. A set of 4 pen ray lamps (Ne, Ar, Kr and Xe) has been analyzed to look for a combination of them where lines are better distributed and less blended (at the X-shooter resolution) than a classical Th-Ar lamp. Combining laboratory measurements with a physical model of the instrument we found that a combination of three
lamps (Ne, Ar and Kr) fills our requirements, and we have produced an order by order prediction for the spectrum that X-shooter should see once in operation.

2. The observational project to build a spectro-photometric flux catalogue of standard stars. Such a catalogue does not exist at the moment and the methods for estimating the absolute flux calibration in the NIR are no more precise than 20-30%. Combining HST and ground-based measurements to calibrate the flux is our solution to establish a robust set of NIR spectro-photometric standard stars. Observations of 15 stars, 3 “primaries” for which we have HST spectra, and 12 “secondaries”, are going on at the VLT using the SINFONI Integral Field Unit. Very preliminary results seem to confirm that when conditions are photometric, we reach an accuracy in the absolute flux calibration of better than 5%.

**Longer term perspectives: the ESPRESSO and CODEX projects**

ESPRESSO and CODEX will be high resolution and very stable spectrographs for an ELT and the VLT, respectively. These instruments will be extraordinary tools to study the universe through QSO absorption lines. On the one hand, CODEX will be able to measure directly the expansion of the universe looking at shift in the central wavelength of Ly$_\alpha$ lines in the forest; on the other hand, ESPRESSO will be the precursor of CODEX, working both with 1 UT (“pre-CODEX mode”) and with the combination of 4 of them (“post-UVES mode”). In both cases ESPRESSO will be able to address in a substantial way many problems from the metal enrichment of the IGM, the chemical elements in Damped Ly$_\alpha$ systems, to fundamental questions like the variation of the fine structure constant with redshift. The main characteristics of these two instruments are listed in tab 7.1 and 7.2. More information on the projects can be found in Pasquini et al. (2005); Cristiani et al. (2007); Liske et al. (2008). I am currently involved in discussions on the realization of these projects and in the possible applications of these instruments on the IGM physics.

<table>
<thead>
<tr>
<th>Telescope diameter</th>
<th>30→60 mt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale at F/2</td>
<td>1 arcsec=0.582 mm</td>
</tr>
<tr>
<td>Feed</td>
<td>Fibre(single or multiple)</td>
</tr>
<tr>
<td>Entrance aperture</td>
<td>1 arcsec</td>
</tr>
<tr>
<td>Wavelength range</td>
<td>380-680 nm</td>
</tr>
<tr>
<td>Spectral Resolution</td>
<td>~110000-140000(depends on feed)</td>
</tr>
<tr>
<td>Main disperser</td>
<td>5×R4 echelle 42 l/mm 160 × 20 cm</td>
</tr>
<tr>
<td>Cross disperser</td>
<td>10 × VPHG 1500 l/mm 15 × 15 cm</td>
</tr>
<tr>
<td>Camera</td>
<td>10 × F/1.4-2.8</td>
</tr>
<tr>
<td>CCD</td>
<td>10 × ~ 8K × 8K (15 µm pixels)</td>
</tr>
<tr>
<td></td>
<td>360 Mpix or 810 cm$^2$</td>
</tr>
</tbody>
</table>

Table 7.1: basic parameters of the CODEX instrument
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Standard: 1 UT</th>
<th>Faint-Object: 4 UT</th>
<th>High Efficiency: 4UT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral coverage</td>
<td>350-780 nm</td>
<td>350-780 nm</td>
<td>350-780 nm</td>
</tr>
<tr>
<td>Resolving Power</td>
<td>160.000</td>
<td>40.000</td>
<td>80.000</td>
</tr>
<tr>
<td>Aperture in the SKY</td>
<td>1.05 arcsec</td>
<td>1.05 arcsec</td>
<td>1.05 arcsec</td>
</tr>
<tr>
<td>Sampling</td>
<td>4 pixels average</td>
<td>16 pixels average</td>
<td>8 pixels average</td>
</tr>
<tr>
<td></td>
<td>(3-5 edges)</td>
<td>(12-20 edges)</td>
<td>(6-10 edges)</td>
</tr>
<tr>
<td>Simultaneous Calibration</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Sky Fibre</td>
<td>YES (either SimCal)</td>
<td>YES</td>
<td>YES (Either SimCal)</td>
</tr>
<tr>
<td>Spatial Pixels</td>
<td>24</td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>RV accuracy (Long Term)</td>
<td>2 cm/sec</td>
<td>100 cm/sec</td>
<td>6 cm/sec (TBD)</td>
</tr>
<tr>
<td>Peak Efficiency</td>
<td>19% including</td>
<td>19% including</td>
<td>19% including</td>
</tr>
<tr>
<td></td>
<td>20% slit losses</td>
<td>20% slit losses</td>
<td>20% slit losses</td>
</tr>
</tbody>
</table>

Table 7.2: Main characteristics of the three modes of ESPRESSO
In this appendix we show the spectra of the pen ray lamps in an order by order analysis of the X-shooter NIR arm. The format is the same for all orders:

- Upper plot: emission line spectrum of a given order showing all lines as a function of the wavelength. The intensity is normalized in such a way that the strongest line of each lamp (over the whole X-shooter NIR spectrum) is set to 100000.

- Lower plot: the same sample zoomed in intensity to clearly show all the lines contributing to the wavelength calibration (intensity > 500).

- The accompanying table gives the number of lines for each lamp. The first column reports the number of lines with intensity above 500 disregarding any possible blends within the spectrum of the same lamp; the second column reports the net number of lines with intensity above 500 available for wavelength calibration taking into account possible blending of lines; the third column reports the number of lines stronger than 20000. This number gives in indication of how many lines would be lost if the operational scheme would allow for saturation of very strong lines. Total values are also shown. The requirement is 10 lines per order, while the goal is 30 for wavelength calibration.

- The appendix can also serve as a quick look tool or atlas during testing and commissioning of X-shooter. While it provides quantitative predictions based on a high fidelity physical model of X-shooter as designed one has to keep in mind that many of the input data are preliminary at this stage and the optimization of the model of X-shooter as actually built is one of the goals of the testing and commissioning phase. Hence caution has to be administered when making quantitative comparisons between simulations and observed spectra.
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.1: Simulated 1D spectrum of X-shooter order #11

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt;20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>17</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Ar</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>9</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>27</td>
<td>2</td>
</tr>
</tbody>
</table>

Table A.1: Order #11
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.2: Simulated 1D spectrum of X-shooter order #12

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Xe</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>19</td>
<td>19</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A.2: Order #12
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.3: Simulated 1D spectrum of X-shooter order #13

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A.3: Order #13
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

**Figure A.4:** Simulated 1D spectrum of X-shooter order #14

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt; 500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>18</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Ar</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>10</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Xe</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>32</td>
<td>24</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table A.4:** Order #14
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.5: Simulated 1D spectrum of X-shooter order #15

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>16</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Xe</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>29</td>
<td>22</td>
<td>7</td>
</tr>
</tbody>
</table>

Table A.5: Order #15
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.6: Simulated 1D spectrum of X-shooter order #16

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt; 500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>8</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>14</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Xe</td>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>28</td>
<td>22</td>
<td>3</td>
</tr>
</tbody>
</table>

Table A.6: Order #16
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity \(>500\) can be expected to be recorded in a single wavelength calibration exposure.

Figure A.7: Simulated 1D spectrum of X-shooter order #17

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: (&gt;500) blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity (&gt;20000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>7</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>12</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Xe</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>30</td>
<td>24</td>
<td>2</td>
</tr>
</tbody>
</table>

Table A.7: Order #17
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.8: Simulated 1D spectrum of X-shooter order #18

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>11</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Xe</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total(req 10; goal 30)</td>
<td>25</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

Table A.8: Order #18
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.9: Simulated 1D spectrum of X-shooter order #19

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt; 500 blending ignored</th>
<th>Net # of Lines: blending ignored</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>16</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Kr</td>
<td>15</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Xe</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Total(req 10; goal 30)</td>
<td>34</td>
<td>27</td>
<td>6</td>
</tr>
</tbody>
</table>

Table A.9: Order #19
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.10: Simulated 1D spectrum of X-shooter order #20

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt;20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Ar</td>
<td>21</td>
<td>17</td>
<td>3</td>
</tr>
<tr>
<td>Kr</td>
<td>10</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total(req 10; goal 30)</td>
<td>37</td>
<td>30</td>
<td>4</td>
</tr>
</tbody>
</table>

Table A.10: Order #20
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.11: Simulated 1D spectrum of X-shooter order #21

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>15</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>14</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Kr</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>37</td>
<td>30</td>
<td>3</td>
</tr>
</tbody>
</table>

Table A.11: Order #21
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.12: Simulated 1D spectrum of X-shooter order #22

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt;20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>9</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Ar</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Xe</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>26</td>
<td>26</td>
<td>6</td>
</tr>
</tbody>
</table>

Table A.12: Order #22
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.13: Simulated 1D spectrum of X-shooter order #23

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500</th>
<th>Net # of Lines:</th>
<th>Lines: Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>blending ignored</td>
<td>blending included</td>
<td>&gt;20000</td>
</tr>
<tr>
<td>Ne</td>
<td>12</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Ar</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>26</td>
<td>24</td>
<td>5</td>
</tr>
</tbody>
</table>

Table A.13: Order #23
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000.

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure.

Figure A.14: Simulated 1D spectrum of X-shooter order #24

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt;500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Ar</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>8</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>19</td>
<td>6</td>
</tr>
</tbody>
</table>

Table A.14: Order #24
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.15: Simulated 1D spectrum of X-shooter order #25

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt; 500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Ar</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Kr</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Total(req 10; goal 30)</td>
<td>16</td>
<td>16</td>
<td>3</td>
</tr>
</tbody>
</table>

Table A.15: Order #25
(a) Ne, Ar, Kr and Xe are superimposed and color-coded. Intensities are normalized such that the strongest line in the whole NIR range has 10000

(b) Expanded view emphasizing fainter lines. Lines with intensity >500 can be expected to be recorded in a single wavelength calibration exposure

Figure A.16: Simulated 1D spectrum of X-shooter order #26

<table>
<thead>
<tr>
<th>Source</th>
<th>Lines: &gt; 500 blending ignored</th>
<th>Net # of Lines: blending included</th>
<th>Lines: Intensity &gt; 20000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ar</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kr</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Xe</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total (req 10; goal 30)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table A.16: Order #26
Bibliography

Barkana, R. & Loeb, A. 2001, Physics Reports, 349, 125
Engleman, R. J., Hinkle, K. H., & Wallace. 2003, JQSRT, 78
Fontana, A. & Ballester, P. 1995, The Messenger, 80, 37
Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168
Kerber, F. e. a. 2008, in preparation
Kurek, A., Hrycyna, O., & Szydlowski, M. 2008, Physics Letters B, 659, 14


Padmanabhan, T. 2003, Physics Reports, 380, 235


Rosa, M. R. 1995, in Calibrating and Understanding HST and ESO Instruments, ed. P. Benvenuti, 43
Sansonetti, C. J. & Greene, M. B. 2007, Physica Scripta, 75, 577
Santos, R. C., Cunha, J. V., & Lima, J. A. S. 2008, PhRvD, 77, 023519
Weinberg, S. 1989, Reviews of Modern Physics, 61, 1


