The Lyman α Forest Toward Quasar HE 0940–1050: Investigation of the Proximity Effect and a Possible Void

von
Michelle Mizuno-Wiedner
aus Honolulu, Hawaii (U.S.A)

Diplomarbeit in Physik
vorgelegt dem
Fachbereich Physik der Universität Hamburg

im August 1998

Angefertigt unter der Anleitung von Prof. Dr. Dieter Reimers
an der Hamburger Sternwarte Bergedorf

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Untersuchung des Absorptionslinienspektrums des quasistellaren Objektes (oder kurz Quasar) HE 0940–1050. Für Quasarspektren typisch befindet sich auf der kurzwelligen Seite der Lyman-alpha-Emissionslinie eine Vielzahl von Absorptionslinien, die mit wenigen Ausnahmen der Lymanserie des neutralen Wasserstoffs zugeordnet werden können. Man spricht in diesem Zusammenhang vom sogenannten Lyman-alpha-Wald, der im Laufe dieser Arbeit genauer analysiert wird.

Das physikalische Bild hinter einem solchen Absorptionsspektrum ist das Folgende: Das vom Quasar ausgesandte Licht passiert auf dem Weg zum Beobachter interstellare Materie im Universum. Die Wellenlänge dieser Strahlung vergrößert sich hierbei kontinuierlich aufgrund der kosmologischen Ausdehnung des Weltalls. Wenn nun Strahlung der Wellenlänge 1215.67 Å auf eine genügend große Konzentration von neutralem Wasserstoff trifft, so kann ein Ly α -Übergang in diesem angeregt werden und es erscheint eine Ly α -Linie im Spektrum mit der gleichen Rotverschiebung wie die der absorbierenden Materie. Aufgrund der Beziehung zwischen Rotverschiebung und Entfernung nach dem Hubble-Gesetz ergibt sich dadurch ein Bild der Gaswolkenverteilung zwischen Quasar und Erde.

Solche Quasarabsorptionslinienspektren liefern nun nicht nur Daten für Modelle, die Beschreibungen über die Strukturbildung im Universum vornehmen, sondern auch für Theorien, die die metagalaktische Hintergrundstrahlung beschreiben. Diese ultraviolette (UV) Hintergrundstrahlung wird für den hohen Ionisationsgrad der Wolken im Lyman-α-Wald verantwortlich gemacht. Frühere Beobachtungen zeigten nun, daß in der Nähe eines Quasars deutlich weniger Absorptionslinien sind als man erwarten würde. In einem Modell von Bajtlik, Duncan und Ostriker (1988) wird dieser “Proximity Effect” auf die Photoionisation des Wasserstoffs durch den Quasar selber zurückgeführt. Das Hauptziel der vorliegenden Arbeit ist nun festzustellen, ob sich ein solcher “Proximity Effect” auch im Spektrum des Quasars HE 0940–1050 beobachten läßt.

Die untersuchten Daten wurden mit dem Echelle-Spektrographen CASPEC am 3.5-Meter-Teleskop der ESO in La Silla, Chile aufgenommen. Die 2-dimensionalen CCD-Bilder mußten in einem langwierigen Analyseprozeß in ein normiertes 1-dimensionales Linienspektrum umgewandelt werden. Der ausführlichen Beschreibung dieser Datenaufbereitung und der verwendeten Methoden sind die ersten Kapitel der Arbeit gewidmet. Für die Untersuchung des “Proximity Effects” selber wurden mehrere Monte-Carlo-Rechnungen unter Zugrundelegung unterschiedlicher Modelle durchgeführt und mit dem gemessenen Spektrum verglichen. Die Annahme von Bechtold und Dobrzycki

(1991b), daß die ionisierende Quasarstrahlung die schwächeren Linien im Spektrum stärker beeinflußt als die stärkeren, scheint durch die Ergebnisse bestätigt zu werden. Trotzdem sind die statistischen Ungenauigkeiten in einem einzelnen Spektrum zu groß, um endgültige Aussagen machen zu können. Sowohl Fits mit dem Modell von Bajtlik, Duncan und Ostriker (1988) für die metagalaktische UV-Hintergrundstrahlung als auch Fits mit der vereinfachten Annahme, daß die Strahlungsintensität an der Lyman-Grenze konstant für verschiedene Rotverschiebungen ist, geben vernünftige χ^2-Werte. Die mögliche Intensität des UV-Hintergrundes an der Lyman-Grenze wurde anhand der Daten auf $\log J_\nu = -21 \pm 0.7$ (alle Linien berücksichtigt) bzw. $\log J_\nu = -21.3 \pm 0.7$ (nur die schwächeren Linien berücksichtigt) abgeschätzt. Für den möglichen Fall, daß die Rotverschiebung des Quasars HE 0940–1050 um 1000–2000 km s^{-1} unterschätzt wurde, sollte sich der Wert von J_ν um einen Faktor 3 bis 5 reduzieren.

Im Laufe der Analyse des Absorptionslinienspektrums stellte sich heraus, daß es eine nahezu linienfreie Region (Void) bei der Rotverschiebung $z \approx 2.580$ gibt. Diese ließe sich z.B. mit der Anwesenheit eines zweiten Quasars und dessen "Proximity Effects" erklären. Zuerst einmal mußte die Signifikanz des Voids untersucht werden. Dies geschah ebenfalls mit Monte-Carlo-Methoden, die die Wahrscheinlichkeit des Autretens einer statistischen Fluktuation im Spektrum, die zu einem solchen Void führen könnte, simulierten. Je nachdem, ob man die stärkeren Linien berücksichtigt oder nicht, beträgt diese Wahrscheinlichkeit entweder $0.2 \pm 0.03\%$ (nur schwächere Linien) oder 5–6% (alle Linien). Hierbei ist es aber nicht klar, ob die starken Linien überhaupt berücksichtigt werden sollten, weil deren Einfluß auf das Spektrum durch Überdeckung schwächerer Linien genau auf einen Void nicht zutrifft. Deshalb scheint eine zukünftige Beobachtung mit dem Ziel des Auffindens eines zweiten Quasars durchaus sinnvoll. Zum Schluß wurde die mögliche Position und Helligkeit dieses weiteren Quasars berechnet. Er sollte wenige Bogenminuten außerhalb der Sichtlinie des Quasars HE 0940–1050 sein.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The spectra of quasistellar objects—more commonly called QSOs or quasars—are dominated by conspicuous emission lines, the largest of which is associated with the Lyman alpha transition of neutral hydrogen. On the long-wavelength side of this emission line, almost all of the not very numerous absorption lines can be readily identified as ionic transitions of metals, i.e., elements heavier than helium. Shortward of the Ly α emission line, however, quasar spectra are very densely populated with hundreds of absorption lines, the overwhelming majority of which are now known (after a suggestion from Lynds 1971) to be lines of the Lyman series of neutral hydrogen H I. The strongest and most numerous of these lines are due to the Ly α transition $1s \rightarrow 2p$, which has a large cross section and a rest wavelength of 1215.67 Å. The densely populated section of the spectrum between the Ly α and Ly β emission lines has been dubbed the "Lyman α forest," and its appearance in the spectrum toward quasar HE 0940–1050 will be presented in this work.

The physical picture corresponding to the absorption spectrum is as follows. Once radiation has been emitted by the quasar, its wavelength increases as a result of the cosmological expansion of the universe. If the wavelength reaches 1215.67 Å and the photons are absorbed by a high enough concentration of neutral hydrogen atoms along the line of sight, a Ly α absorption line appears in the observed spectrum with the same redshift z as that of the absorbing material. Since redshift is also a measure of distance according to the Hubble law, absorption spectra yield a picture of gaseous absorbers, or "clouds," scattered along the path from the quasar to Earth.

Traditionally the absorbers with column densities of up to $N_{\rm HI} \sim 10^{16}$ atoms per cm^2 along the line of sight have been called Ly α forest clouds. The metal lines present in the systems with the higher column densities become increasingly stronger up to $N_{\rm HI} \sim 10^{17}$, where the gas also becomes optically thick to ionizing radiation. The continuous absorption below 912 Å in the rest frame of these so-called Lyman-limit systems produces a characteristic discontinuity at the Lyman limit in the observed spectrum. Finally, the damped Ly α systems with $N_{\rm HI} \gtrsim 10^{20}$ produce such heavily saturated Ly α absorption lines that the line profiles are dominated by so-called damping wings.

The Ly α clouds with the lowest column densities arise in the intergalactic medium and were thought to constitute a primordial population because they appeared to lack metal lines and showed little clustering—as opposed to the systems with higher column densities and strong metal lines which seem to be associated with galaxies. However, as mentioned in chapter 7, recent detection of C IV absorption in clouds with column densities at least as low as $N_{\mathrm{HI}} \simeq 10^{14.5}$ has cast some doubt on this picture of different parent populations. From the other side, steadily improving theoretical simulations are blurring the distinction between discrete low-N_{HI} absorbers and other density fluctuations in a more diffuse medium.

Regardless of whether their chemical composition has undergone some metal enrichment, and regardless of whether future approaches will take a more general view of structures in the intergalactic medium, the fact remains that the systems producing Ly α absorption lines in quasar spectra provide an extremely useful probe of baryonic matter in the universe, its distribution over a wide range of densities and temperatures, and its evolution in time. The Ly α clouds appear at all directly observable redshifts and are by far the most numerous population at higher redshifts, where quasar lines of sight cover look-back times representing a significant fraction of the lifetime of the universe.[1]

Observational data provide input and constraints not only for theories of structure formation in the universe, but also for theories of the metagalactic ultraviolet radiation field that is assumed to be responsible for keeping the Ly α forest clouds in a highly ionized state (Haardt and Madau 1996). After it was noticed that the number density of Ly α lines in the vicinity of a quasar does not increase as strongly with redshift as expected, Bajtlik, Duncan and Ostriker (1988) proposed a specific model in which enhanced photoionization of the Ly α clouds by the quasar causes what they dubbed the "proximity effect." The principal parameter in this model is the ratio of the quasar's Lyman-limit flux density to that of the ambient UV background (UVB). Thus, attempts to determine the UVB by integrating over the contributions from all presumed sources could now be complemented with actual indirect measurements of the UVB at the Lyman limit.

The final aim of this work is to see if there is any detectable proximity effect in the Ly α forest toward HE 0940–1050, and if possible to obtain an estimate of the UVB under the assumption that photoionization is indeed responsible for the proximity effect. Before that could be attempted, however, the raw two-dimensional echelle data had to be converted into a one-dimensional normalized spectrum. Since this is a long and involved process for a first-time student, the individual steps will be described in great detail. Chapter 3 explains the preliminary steps required for the actual data reduction in chapter 4. Chapter 5 describes the additional corrections and normalization which produce a spectrum suitable for fitting absorption lines.

---

[1] The general properties and cosmological significance of quasar absorption systems are also presented in a brief and understandable introduction by Tytler (1997). For these and other topics concerning the Ly α forest, along with the developments in recent years, see the very thorough review by Rauch (1998).

The first part of the actual analysis, covered in chapter 6, begins by fitting Voigt profiles to all of the absorption lines in the Ly α forest in order to obtain the distributions of their redshifts, column densities and Doppler parameters. Chapter 7 describes the expected shapes of these distributions, introduces the theoretical model values chosen to parametrize these distributions, and explains how a preliminary examination of the data was used to define the sample used in the subsequent analysis.

The actual investigation of the proximity effect in the spectrum of HE 0940 is covered in chapter 8. A very brief introduction is followed by a description of the photoionization model in which the necessary formulas are also presented. The relevant quasar flux density is then determined via extrapolation to the Lyman limit.

After explaining the Monte Carlo methods and programs used to simulate the proximity effect, the data are compared to the Monte Carlo results for three different scenarios. The first represents the Ly α forest distributions in the absence of quasar radiation. The second and third include the presence of quasar radiation plus either the UVB model presented by Bajtlik, Duncan and Ostriker (1988) or a UVB with a Lyman-limit intensity that is constant with redshift. Trying different values of this constant leads to a more concrete estimate of the order of magnitude for the UVB and gives an impression of the range of errors involved. More general sources of error and systematic uncertainties are also mentioned.

The discovery of a possible void in the spectrum at $z \approx 2.58$ brings up the closely related topic of a possible two-quasar proximity effect, in which a foreground quasar produces a line-free region in the absorption spectrum of another quasar when it ionizes the Ly α clouds along the line of sight to the latter. After estimating the size of the void candidate in the spectrum of HE 0940, the region is examined more closely for the presence of weak lines, and two estimates of the region's statistical significance are made. Chapter 9 closes by considering a few of the conditions necessary for seeing such a foreground quasar and by providing rough estimates for its possible location and luminosity.

Chapter 2

Observations

2.1 The Hamburg/ESO Survey

The quasar HE 0940–1050 was discovered as part of the wide-angle Hamburg/ESO survey (HES), which began in 1990 with the objective of obtaining a large homogeneous sample of bright QSOs in the southern hemisphere (Reimers et al. 1996). The objective-prism photographs taken with the ESO 1-meter Schmidt telescope were digitized and examined according to a variety of selection criteria that included the ultraviolet excess, continuum slope and presence of strong emission lines typical of QSO spectra (Wisotzki et al. 1996). For the QSO candidates thus selected, follow-up slit spectroscopy was carried out on the 1.52 m, 2.2 m and 3.6 m ESO telescopes.

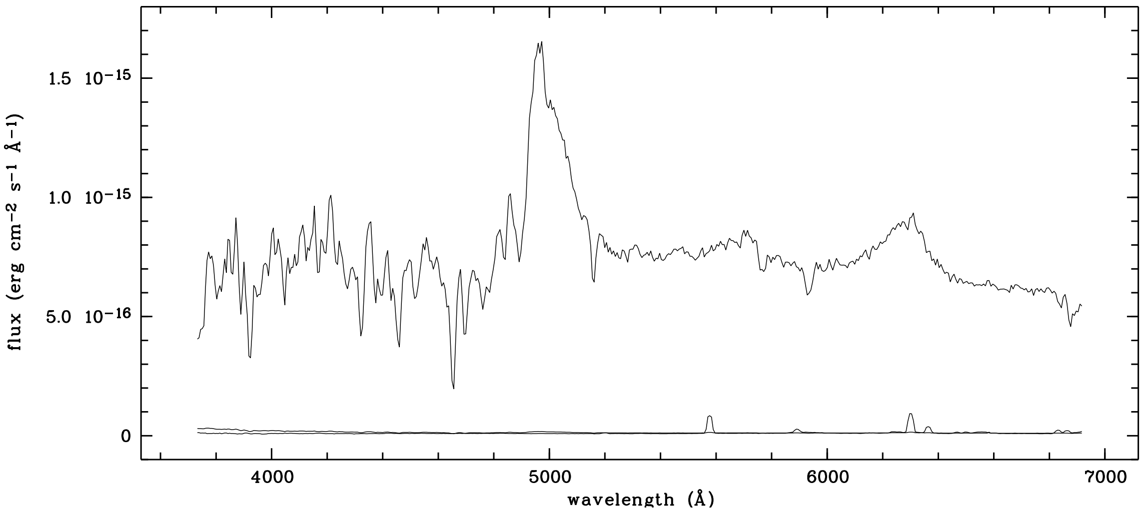


Figure 2.1: *Spectrum of HE 0940–1050 obtained from two 30-minute exposures taken with the ESO 1.52-meter telescope (courtesy of L. Wisotzki).*

The follow-up spectrum of HE 0940–1050 in Figure 2.1 was taken at the 1.52-meter telescope on March 21, 1993. The instrumental magnitude of $B = 17.2$ obtained via CCD photometry with the EFOSC1 spectrograph has an error smaller than $\Delta B = \pm 0.1$. Because the centroid of the Lyman α emission line can be distorted by the Lyman α forest, the QSO redshift of $z_{em} = 3.054$ was determined by taking the peak wavelength of the C IV emission line (Reimers et al. 1996).

HE 0940–1050			
Coordinates:	$\alpha(1950)$	=	$9^{\mathrm{h}}\ 40^{\mathrm{m}}\ 27.2^{\mathrm{s}}$
	$\delta(1950)$	=	$-10^{\circ}\ 50'\ 41''$
	$\alpha(2000)$	=	$9^{\mathrm{h}}\ 42^{\mathrm{m}}\ 53.5^{\mathrm{s}}$
	$\delta(2000)$	=	$-11^{\circ}\ 04'\ 26''$
Magnitude:	B	=	17.2 ± 0.1
Redshift:	z_{em}	=	3.054

Table 2.1: *Quasar HE 0940–1050*

2.2 Data Collection

The raw CCD images analyzed in this work were obtained by Susanne Köhler on February 3–5, 1994 at the 3.6 m telescope of the European Southern Observatory (ESO) at La Silla, Chile with the Cassegrain Echelle Spectrograph (CASPEC).

Echelle spectrographs use diffraction gratings with greatly increased blaze angles that allow them to observe higher diffraction orders and so achieve higher resolutions. But the resulting image consists of many short spectral orders superimposed on each other. They must be separated out by a second spectroscope, called a cross disperser, whose dispersion direction is perpendicular to that of the first (Kitchin 1995 has a clear explanation of the optics involved). Thus the final image on the CCD is a two-dimensional array of different segments of the spectrum.

In the case of the CASPEC observations, light passed through a slit 500 μm long and 280 μm wide (which corresponds to $2''$). Then it went through a collimator, an echelle grating with 31.6 lines mm^{-1}, and a cross disperser with 300 lines mm^{-1} and a maximum efficiency at 4224 Å (Pasquini 1993). Finally, the light was focused on the CCD by the Short Camera (which is no longer in use; Randich and Pasquini, 1996). The CCD (ESO #32) is a thinned, multi-phase-pinned Tektronix CCD with 512×512 pixels and a pixel size of 27μm $\times$27μm. No on-chip binning was done.

Table 2.2 lists all of the calibration and object exposures, for which the seeing was between $1.3''$ and $1.5''$. In column 2, BIAS AV5 means that the image stored on tape was the simple average of 5 bias frames. Flat-fields are labelled FLAT, and exposures of the thorium-argon lamp used for the wavelength calibration are labelled WLC. The STD exposure of μ Columbæ, an HST standard of spectral type O9V suitable for flux calibration, was used for the response correction in section 4.9. Where available, the universal time at the start of the exposure, the exposure time in seconds, and the telescope coordinates are listed in the last 4 columns of the table.

Frame name	Identifier	Date	U.T.	Sec	R.A.	Decl.
casp0004	BIAS AV5	00100	00:00:00	1	$0^h00^m00^s$	$00°00'00''$
casp0005	BIAS AV5	00100	00:00:00	1	$0^h00^m00^s$	$00°00'00''$
casp0006	BIAS AV5	00100	00:00:00	1	$0^h00^m00^s$	$00°00'00''$
casp0007	BIAS AV5	00100	00:00:00	0	$0^h00^m00^s$	$00°00'00''$
casp0021	FLAT	Feb 3	03:32:14	35	$9^h42^m38^s$	–11°02′46″
casp0022	WLC	Feb 3	03:34:56	2	$9^h42^m38^s$	–11°02′46″
casp0023	**HE0940** 90m	Feb 3	03:36:46	5400	$9^h42^m38^s$	–11°02′46″
casp0024	WLC	Feb 3	05:09:07	2	$9^h42^m38^s$	–11°02′42″
casp0025	FLAT	Feb 3	05:11:14	35	$9^h42^m38^s$	–11°02′43″
casp0029	FLAT	Feb 3	05:48:13	35	$9^h42^m38^s$	–11°02′58″
casp0030	WLC	Feb 3	05:51:04	2	$9^h42^m38^s$	–11°02′59″
casp0031	**HE0940** 2h50m	Feb 3	05:53:10	10200	$9^h42^m38^s$	–11°02′58″
casp0032	WLC	Feb 3	08:45:10	2	$9^h42^m38^s$	–11°03′00″
casp0033	FLAT	Feb 3	08:47:10	35	$9^h42^m38^s$	–11°03′01″
casp0034	DARK a.m.	Feb 3	08:55:38	3600	$13^h05^m37^s$	–29°25′40″
casp0035	BIAS	Feb 3	16:45:25	1	$20^h56^m40^s$	–29°25′40″
casp0036	BIAS	Feb 3	16:47:11	1	$20^h58^m27^s$	–29°25′40″
casp0037	BIAS	Feb 3	16:48:20	1	$20^h59^m36^s$	–29°25′40″
casp0038	BIAS	Feb 3	16:49:27	1	$21^h00^m44^s$	–29°25′40″
casp0039	BIAS	Feb 3	16:50:37	1	$21^h01^m53^s$	–29°25′40″
casp0067	DARK a.m.	Feb 4	09:17:21	3600	$13^h31^m22^s$	–29°25′37″
casp0069	BIAS AV5	00100	00:00:00	1	$00^h00^m00^s$	00°00′00″
casp0070	DARK p.m.	Feb 4	20:23:17	3600	$0^h39^m07^s$	–29°25′36″
casp0076	WLC	Feb 5	03:18:30	2	$5^h45^m48^s$	–32°18′40″
casp0077	FLAT	Feb 5	03:20:27	35	$5^h45^m48^s$	–32°18′41″
casp0078	**STD**	Feb 5	03:22:52	20	$5^h45^m48^s$	–32°18′41″
casp0087	FLAT	Feb 5	06:48:44	35	$9^h42^m37^s$	–11°02′51″
casp0088	WLC	Feb 5	06:51:15	2	$9^h42^m37^s$	–11°02′51″
casp0089	**HE0940** 2h30m	Feb 5	06:53:35	9000	$9^h42^m37^s$	–11°02′52″
casp0090	WLC	Feb 5	09:25:27	2	$9^h42^m38^s$	–11°02′53″
casp0091	FLAT	Feb 5	09:27:24	35	$9^h42^m38^s$	–11°02′53″

Table 2.2: *List of CCD exposures Feb. 3–5, 1994.*

Chapter 3

Preprocessing

Because the echelle spectrograph produces two-dimensional images on the CCD, many steps are required to produce the one-dimensional spectra in which the absorption lines can be seen and measured. To prepare the two-dimensional frames for this so-called data reduction, the digitized voltage signals—which are recorded by the CCD as data numbers (DN) or analog-to-digital units (ADU)—must be converted back into the number of photoelectrons or photons that created the signal in the first place.

3.1 ESO-MIDAS

The preprocessing and reduction of the data was performed almost entirely using the European Southern Observatory-Munich Image Data Analysis System. ESO-MIDAS, or MIDAS for short, is a software package containing general tools for image processing and data reduction (ESO-MIDAS User Guide, Vol. A).

The two-dimensional raw CCD images, also called frames in MIDAS, can be viewed in a display window, while plots of the pixels values versus the x- or y-coordinates can be generated in a graphics window. The preprocessing here made use primarily of the basic MIDAS commands for performing mathematical operations on the frames, such as the pixel-by-pixel subtraction of one frame from another. Later, during the data reduction, more use was made of higher-level command procedures with different options (see chapter 4).

3.2 Bias Subtraction

In order to keep readout noise from driving the analog-to-digital input below zero, the readout electronics of a CCD will always add a positive, constant offset to the signal from each pixel before it is digitized. This additive constant must therefore be determined and subtracted from all of the exposures before any further calibration steps can be done. Since the bias is present even when no photons are detected, the bias level can be determined by taking the shortest possible (nominally zero-second) exposures with the shutter closed.

For high-quality CCDs, the bias is constant in time, remaining the same from one exposure to another. It also shows no spatial dependence; i.e., the variation in the bias value from pixel to pixel is significantly smaller than the readout noise in each pixel (Wisotzki 1994). In that case, one can combine all of the bias frames and obtain a master bias frame or merely a constant value that is then subtracted from all of the other CCD exposures. According to simple error propagation formulas, averaging several bias frames has the advantage of reducing the estimated error by a factor of $\sqrt{n}$, where n is the number of frames used.

Unfortunately, the bias frames taken for HE 0940 show levels that vary not only from night to night but also over the surface of the CCD as well (see Figure 3.1). To compensate for the fact that the bias changed over time, a bias level was determined individually for each frame by averaging all of the pixel values in a stable part of the overscan region on the CCD. Between 191 and 210 ADU were thus subtracted from all of the frames in this first part of the bias correction.

After the information from the overscan region had been used for the first part of the bias correction, the edges were trimmed from all frames to avoid problems that could result from including areas with pathological behavior in the smoothing steps that followed. The START and STEP descriptors were then reset to (1.0, 1.0) so there would be no confusion between the so-called world coordinates and pixel coordinates.

Despite being scaled to about the same level, the spatial variations in the bias across each frame still remained. The strategy was to reduce them by subtracting a two-dimensional bias frame that had been strongly smoothed beforehand to minimize the effects of the readout noise.

The bias calibration frame for the exposures taken on the first night (Table 2.2) was created as follows. The top five curves in Figure 3.1 each represent the average of 5 individual bias frames. Since their pixel values and spatial structures overlapped well after subtraction of the overscan values, they were all averaged together and median filtered to remove most of the cosmics. The filtered frame was smoothed by fitting a polynomial of third degree in both the x and y positions on the CCD. This frame was then subtracted from frames casp0034.bdf and below.

The bias calibration frame for the second night's exposures corresponds to the bottom curve in Figure 3.1. It was created by averaging the remaining 5 bias frames, applying the MIDAS command `FILTER/COSMIC` (which was more effective than the median filter in removing a large cosmic covering four pixels), and again doing a two-dimensional, third-degree polynomial fit. This fitted frame was subtracted from frames casp0067.bdf and above.

The two-dimensional bias subtraction was reasonably successful in reducing most of the structure produced by a non-constant bias. The left panel of Figure 3.2 shows

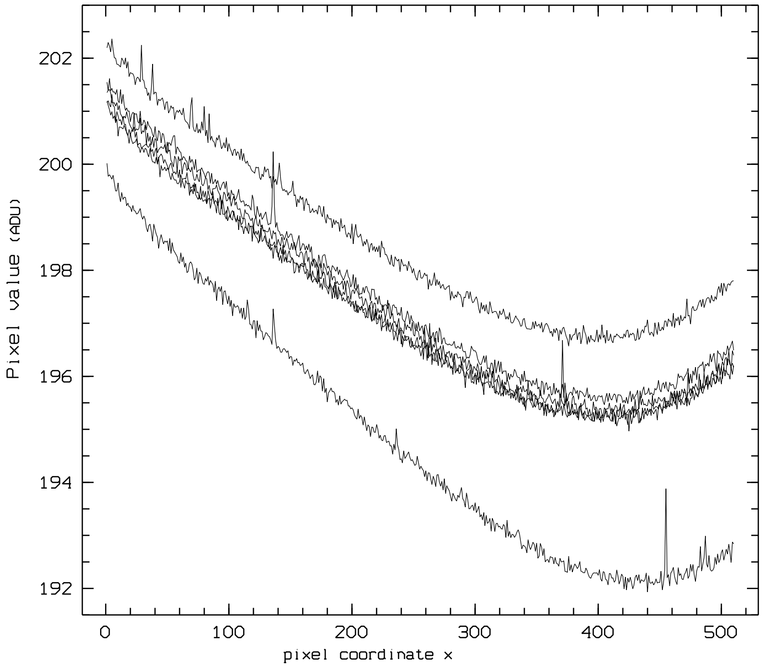


Figure 3.1: *Two-dimensional structure of bias frames as seen in the average column of each frame, shown before subtraction of the overscan value.*

the average column in each of four long dark frames before the bias correction (the cosmics have been filtered out for presentation purposes, to allow the structure to be seen). The right panel shows three of the average columns after the two-step bias correction. Note that casp0045.bdf is no longer present; it was discarded because its dark-current level was notably higher and its structure slightly different from that of the others. Since the dark frames were taken during the day and the detector—which is not completely light-tight—was left unattended, it is quite possible that scattered light degraded that exposure.

3.3 Dark-Current Correction

Independent of bias, the random thermal motions of atoms within the silicon lattice will produce enough energy to transfer some electrons from the valence band to the conduction band of the CCD, thus generating a non-zero signal even in the absence of light. Cooling the CCD or operating it in inversion mode (McLean 1997) greatly reduces but does not elimate this dark current, which adds charge linearly with exposure time and can be determined by taking long exposures with the shutter closed. The correction is then made to each non-bias frame by multiplying the dark current in ADU per second by the frame's exposure time in seconds, and subtracting the result from that frame.

As with any long exposures, the so-called dark frames show discrete bright spots

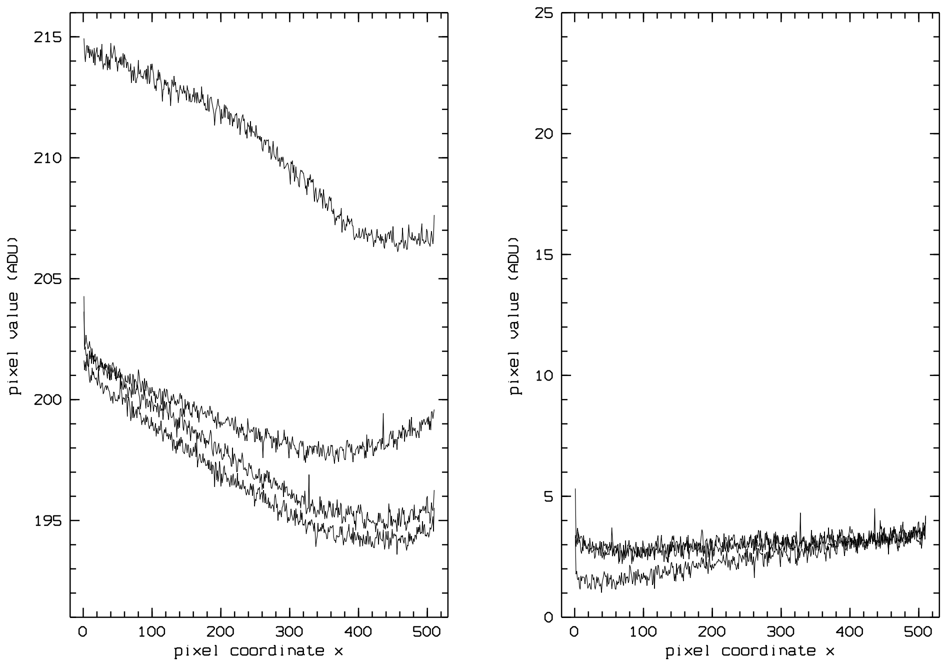


Figure 3.2: *Average filtered column of each dark frame. LEFT PANEL: before the two-step bias correction. RIGHT PANEL: after the bias correction (without casp0045).*

or small streaks caused by cosmic radiation—subatomic particles that deposit large amounts of energy when they are stopped in the CCD material. This happens up to a handful of times per cm^2 per minute, depending on the altitude. Given the low probability of two particles striking the same spot on the CCD in two different exposures, such cosmics can be removed by combining two or more long dark frames. But since casp0045.bdf had been discarded, only the long dark frame casp0034.bdf was available to calibrate the frames taken on the first night. Its counts differed from the long darks taken on the second night, so combining frames was not an option in this case.

Rather than attempting to remove all of the cosmics in casp0034.bdf by changing the pixel values manually with MIDAS commands such as `MODIFY/AREA`, a different approach was taken. The dark current was determined by making a histogram of the number of pixels vs. count rate in each of the remaining three long dark exposures (Fig. 3.3). Pixels containing cosmic events had much higher pixel values and were much less numerous than the 99.7% of lower-valued pixels that contained only dark-current counts, so the entries with cosmics were scattered well off to the right of the main distribution and thus neglected. A Gaussian was fitted to the histogram of dark-current counts, and its center value yielded the average dark-current value for each frame. The estimated dark current was 1.9 ADU per hour on the first night. The corresponding results for the other two dark frames were averaged to yield 2.5 ADU per hour for the second night.

Although the dark current seemed negligible at about 0.0006 counts per second, the long exposure times needed for the quasar images meant that it accounted for 8% and 11% of the count rates in the object frames taken on the first and second night, respectively.

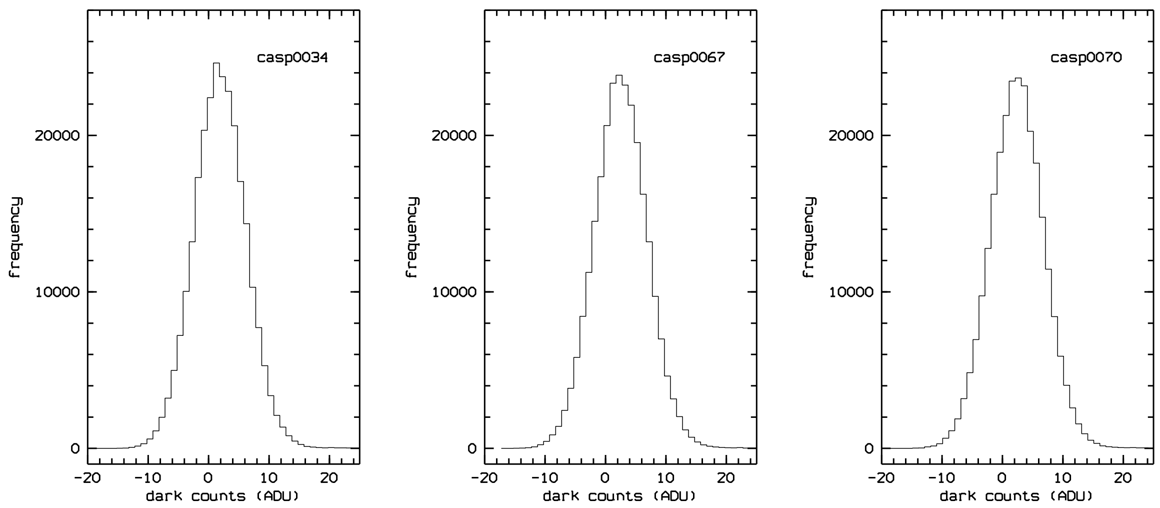


Figure 3.3: *Histograms of counts in long dark frames*

3.4 Conversion to Photoelectrons

When a photon strikes the CCD, its energy looses a certain number of electrons from the valence band into the conduction band, allowing them and the corresponding "holes" in the valence band to move through the material under the influence of an electrical field. The resulting voltage is read out and amplified by the detector electronics. The number of counts or ADU of the actual output signal S is thus

$$S = \frac{N_e}{g} + d\,\Delta t + b \tag{3.1}$$

where d is the dark current in units of ADU per second, Δt is the integration time in seconds, b is the bias in ADU, and g is the system's photon transfer gain factor in electrons per ADU. Clearly, after the bias and dark current have been subtracted, the remaining signal is directly proportional to the number of photoelectrons $N_e = \eta N_\gamma$, where N_γ is the number of photons that were absorbed in the time Δt. McLean (1997) describes in detail how the value of g can be either calculated from the overall amplifier gain and the CCD capacitance, or measured experimentally. In this case the conversion factor was simply taken to be 2.1 electrons per ADU, following the information contained in the MIDAS descriptors of the data frames.

Poisson statistics apply to photon or electron counts but not to ADU, so this conversion was needed in order to calculate the photon noise; namely $\sqrt{N}$ for a signal of N photoelectrons. On the other hand, the CCD electronics produce a certain amount of noise whenever they are read out, regardless of the size of the signal. In the case of CCD #32, this noise level was ~ 11 e$^-$ (Pasquini 1993). Assuming that the photon noise and the read-out noise σ_r are independent and random, the expression for the combined noise in the raw frames becomes

$$\sigma = \sqrt{{\sigma_N}^2 + {\sigma_r}^2} = \sqrt{N + {\sigma_r}^2} \tag{3.2}$$

Additional errors arise as changes are made to the frames during preprocessing and data reduction. Since the bias is read out without a photon signal, the error associated with a constant bias level is due to the read-out noise and is thus equal to σ_r. The two-dimensional bias correction is also subject to the errors in the read-out noise, but error propagation yields an error of $\sigma_r/\sqrt{n}$ because n frames were averaged to obtain each bias frame. In the end, the variances σ^2 for the raw frames, the bias correction, the dark-current subtraction and later the background subtraction all had to be added together and kept for later use in the optimal extraction of the spectrum (see section 4.5).

Chapter 4

Data Reduction

The purpose of the reduction is to convert the two-dimensional raw CCD images obtained at the telescope into a one-dimensional spectrum containing the measured intensity as a function of wavelength. Since this was one of the main tasks of this work—and certainly the most time-consuming—it will be described in detail in the following sections.

4.1 The Echelle Context in MIDAS

Most of the work described in this chapter was done within the MIDAS echelle context, which is an application package containing commands specially created for the reduction of echelle spectra. Information on the use of these commands was gleaned from chapter 7 and appendix D of the MIDAS manual, volume B.

An echelle session within MIDAS has a few additional features that should be mentioned briefly. During each reduction step that is performed within the echelle context, new information is written to temporary system tables which are automatically created in the current working directory. They are called order.tbl, back.tbl and (if any wavelength calibration steps were done) line.tbl. Since MIDAS refers to these particular tables in any subsequent reduction steps, it is usually necessary to preserve the information for use in a future MIDAS session by saving the tables under a session name.[1]

The echelle context uses a large number of variables, or MIDAS keywords, as session parameters to specify input frames, default command arguments, and reduction methods or options for each of the reduction steps. They are stored in the system tables and can be reviewed using the command `SHOW/ECHELLE`.

[1]For example, the command `SAVE/ECHELLE SESS` copies the temporary tables to the files SESSORDE.tbl, SESSback.tbl and SESSLINE.tbl. The information can later be made available via the command `INITIALIZE/ECHELLE SESS`, which copies the tables with this prefix back into the tables order.tbl, back.tbl and line.tbl.)

4.2 Order Definition

The first step in the data reduction is to define the positions of the spectral orders on the two-dimensional image. As described in chapter 7 of the MIDAS manual, this means solving the equation

$$y = f(x, m) \tag{4.1}$$

which relates the line number y to both the sample number x and the spectral order m on the raw image. The reduction software assumes that the images have been rotated and flipped so that they are oriented as follows: the spectral orders almost horizontal on the image, with a slight positive slope; the wavelength scale increasing from left to right; and the absolute order number increasing from top to bottom, so that the separation between spectral orders is smaller at the bottom that at the top.

The MIDAS command `DEFINE/HOUGH` was used to obtain the initial positions and slopes of the spectral orders on the CCD. In this algorithm, the center of each order is defined as the first moment of the pixel values above a certain threshold and is approximated by a polynomial

$$y \approx \sum_{j=0}^{J} \sum_{i=0}^{I} a_{ij}\, x^i m^j \tag{4.2}$$

The coefficients a_{ij} which define the order positions are stored in the table order.tbl. Points belonging to the background between the orders are also defined at this time and stored in the table back.tbl.

Because the quasar images of HE 0940 are so faint, the order definitions had to be done on the much brighter flat-field frames. This meant that when the spectrum was extracted from the wavelength-calibration (WLC), standard-star (STD) and quasar frames, the algorithm used the initial order positions obtained from the flat-fields taken just before or afterwards (see Table 4.1). All of the frames contain 43 echelle orders.

4.3 Background Subtraction

Even after subtraction of the bias and dark current, the pixels between the bright spectral orders still have non-zero values. This inter-order background, which is due to general scattered light from the instrumentation as well as diffuse light from adjacent orders, must be estimated and then subtracted from the entire frame.

The standard procedure uses the grid of background points identified and stored in back.tbl to estimate the inter-order background. This is done either by a polynomial fit of the selected points (method `POLY`) or by smoothing spline interpolation (method `SPLINE`), which divides the grid into discrete intervals and fits them with polynomials whose function values and derivatives are made to be continuous at the endpoints of the intervals where the adjacent polynomials meet. Both methods are implemented via the MIDAS command `BACKGROUND/ECHELLE`.

When applied to the quasar and flat-field frames, however, it turned out that the single polynomial in method `POLY` was not flexible enough to follow the background level in all cases, while method `SPLINE` produced artifacts and uneven patterns on the edges of the image. Only the command `BACKGROUND/SMOOTH`—which uses an iterative algorithm to heavily smooth the frame perpendicular to the orders and replace higher pixel values with the smoothed values—was able to produce background estimates that showed no undesirable features but had reasonable levels. Because read-out noise causes fluctuations even where there would otherwise be a constant level, the estimated background should not pass through the very lowest pixel values, but rather more or less through the middle of those fluctuations. Figure 4.1 shows an example of this.

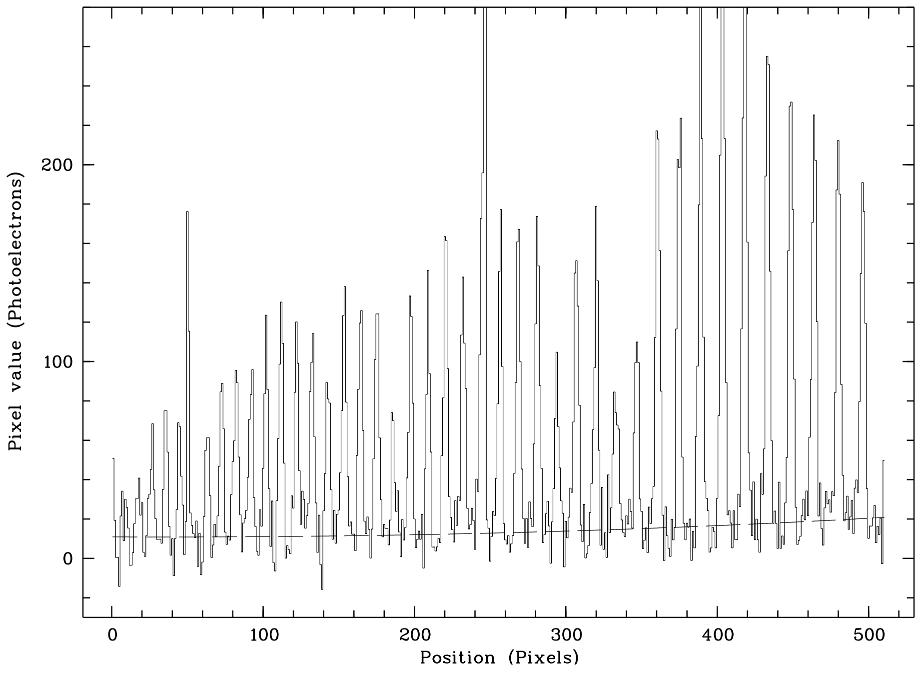


Figure 4.1: *Sample column in quasar frame casp0031 (solid line) and estimated inter-order background level (dashed line).*

The main disadvantage of `BACKGROUND/SMOOTH` was that it produced an extremely "lumpy" background for the wavelength-calibration frames; but since the identification of emission lines was not impaired, this was not judged to be critical. Thus the inter-order background in the WLC, standard-star and quasar frames was estimated via the smoothing method and then subtracted.

4.4 Flat-field Correction

The primary purpose of flat-field frames is to correct for the non-uniform sensitivity across the CCD. Flat-fields are created by exposing the CCD to a uniformly illuminated field, such as the one provided by the inside of the white telescope dome or a flat white screen uniformly illuminated by lamps; exposures of this kind are called dome flats.

Following McLean (1997), let I be the uniform illumination falling on a pixel in row i, column j whose quantum efficiency is η_{ij}. After the bias, dark current and inter-order background have all been subtracted, the observed signal in ADU from that pixel in the flat-field frame is

$$(X_{ij})_{FF} = \frac{1}{g}\ \eta_{ij} I. \tag{4.3}$$

The averaged, or sometimes merely strongly smoothed, flat-field has pixel values

$$\langle (X_{ij})_{FF} \rangle = \frac{1}{g}\ \eta_m I \tag{4.4}$$

where η_m is the mean quantum efficiency averaged over the entire image. When the original flat-field frame is divided by this average, the resulting "normalized" flat-field calibration frame

$$(X_{ij})_{cal} = \frac{(X_{ij})_{FF}}{\langle (X_{ij})_{FF} \rangle} = \frac{\eta_{ij}}{\eta_m} \tag{4.5}$$

has pixel values in the neighborhood of 1 which are essentially multiplicative factors representing the pixel sensitivity.

In a standard-star or science frame, the illumination I_{ij} is no longer constant and the observed signal from the pixel in row i, column j in a dark- and background-subtracted frame is

$$X_{ij} = \frac{1}{g}\ \eta_{ij} I_{ij} \tag{4.6}$$

The pixel-to-pixel variations represented by the factor η_{ij} can now be removed by dividing the frame in question by the normalized flat-field, yielding an image whose pixel values contain the relative intensity

$$\frac{X_{ij}}{(X_{ij})_{cal}} = \left(\frac{\eta_m}{g}\right)\ I_{ij} \tag{4.7}$$

which differs from the true illumination I_{ij} by only a constant factor. If necessary, the scale factor η_m/g can be determined by observing a source of known brightness and scaling the pixel values to match the known absolute flux.

Telescope flexure leads to changes in the CCD position of the infalling light over the course of the night, so in keeping with the usual practice flat-fields were taken immediately before and after each object exposure to ensure that the order positions would not differ much. Unfortunately, although the order centers obtained from the flat-fields matched those in the quasar frames well, the quasar orders themselves were wider because of slight shifts in position. Had the flat-fields been taken using a longer slit, the quasar and flat-field orders might have overlapped fully. But as it was, the division shown in equation (4.7) could not be carried out because it would have led to undefined regions in the QSO spectrum.

The flat-field correction therefore had to be neglected, and the flat-fields could be used only for order definition. Luckily, the CCD images contained no obvious features such as the fringes that sometimes appear in thinned CCDs due to interference from the night-sky emission lines.

4.5 Spectrum Extraction

Normally, the echelle orders on the CCD image are extracted by summing or averaging the pixel values in the y direction along a "numerical slit" one pixel wide, whose height is set by the user. The sum of the pixel values yields the flux as a function of x for each spectral order on the CCD. In practical terms, this means converting a two-dimensional raw image in pixel-pixel space to a two-dimensional "extracted" image in pixel-order space, i.e., whose x-axis is still the pixel coordinate but whose y-axis is now the relative order number. However, the final signal-to-noise ratio (S/N) is not always optimal. In particular, the many cosmics in long exposures must be removed manually.

In order to extract the spectrum in an optimal manner that maximizes the S/N, Sebastián López at the Hamburg Observatory has written a series of procedures that were used here instead of the standard MIDAS command `EXTRACT/ECHELLE`. In addition to a more flexible order definition, the main advantage of the Hamburg programs is that they take the variance of each pixel into account. The required input is therefore not only the usual frame containing the orders but also its variance frame, i.e., a frame whose pixel values equal the variance of the corresponding pixels in the frame that is to be extracted.

The variances for the frames processed here included not only the combined errors from photon noise and read-out noise, but also the estimated errors from the bias, dark-current and background corrections. Just before the extraction, cosmics were identified either as pixels for which the ratio of the input frame to a filtered version of the input frame exceeded a pre-defined threshold, or as pixels having more than the maximum count rate chosen for each frame by inspection. These outliers were given practically infinite variances.

For each echelle order, the extraction program used the positions provided by the MIDAS order definition (including any user-defined constant offset) as the initial values. At each x, a Levenberg-Marquardt algorithm fitted Gaussians to the order profile in the y direction along the slit; the pixels were weighted according to their variances. Once the Gaussian fits had yielded the order center in y at each x, a polynomial fit through the order centers was used to fix the position of the order as a whole. Then the spatial profiles were fitted once again, but this time the position and FWHM of the Gaussians were held fixed and only the amplitudes were allowed to vary. The area under the resulting curve at each x represented the extracted flux and was written to the output frame, while the error in the flux was written to a corresponding error frame. The flux errors proved extremely useful when trying to combine the three quasar spectra in a sensible way (see section 5.1).

The spectrum extraction achieved good results, as can be seen in Figure 4.2. Because of their long exposure times, the science frames in particular had been plagued by cosmics; but since cosmics had been tagged beforehand by means of variance frames, the Hamburg programs gave them correspondingly low weights in the Gaussian fits. Thus the optimally extracted flux values showed far fewer cosmics than those obtained with `EXTRACT/ECHELLE`.

Finally, all of the extracted frames were divided by their exposure times. This was in part because the MIDAS command used as the basis for the subsequent response correction assumed that the counts are normalized to an exposure time of one second.

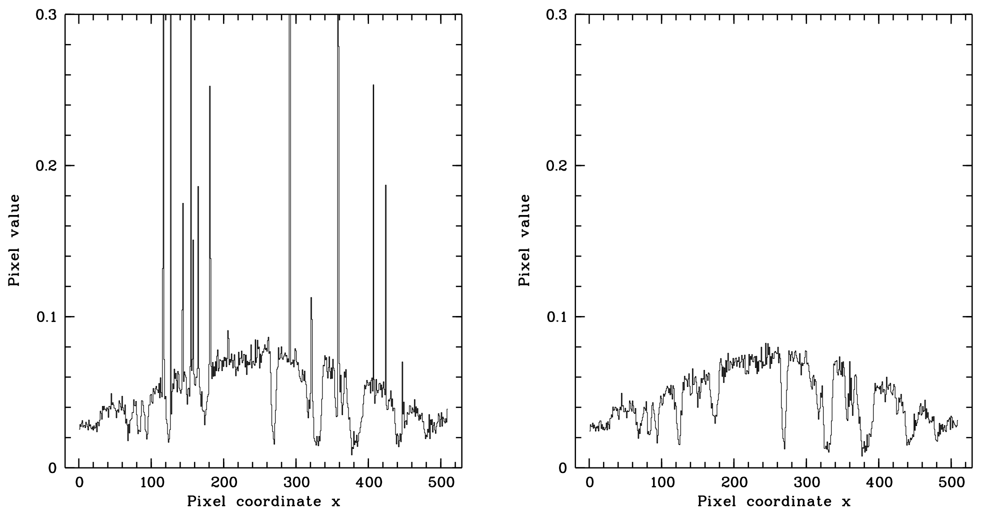


Figure 4.2: *Sample order (25) extracted from background-subtracted quasar frame casp0089.bdf. LEFT: result of standard MIDAS extraction. RIGHT: result of of Hamburg programs.*

4.6 Wavelength Calibration

The spectrum extraction provides flux values as a function of the sample position x on the CCD and the spectral order m. A wavelength calibration is still needed to translate the x position within each order into the actual wavelength λ; that is, the relation

$$\lambda = f(x, m) \tag{4.8}$$

must be solved together with equation (4.1) in order to obtain the desired transformation between (x, y) space and (λ, m) space. This is done with the help of wavelength calibration (WLC) frames, which in this case were exposures of a thorium-argon arc lamp.

The wavelength calibration in the MIDAS echelle context is done in two steps. The first step involves the extracted WLC frame, in which the thresholding algorithm in `SEARCH/ECHELLE` automatically detects the positions of emission lines. The user determines whether the center of each line should be estimated by fitting a Gaussian to the line profile (which was the method used in this case) or by finding its center of gravity. The line centers in x are then stored in the table line.tbl.

The second step in the wavelength calibration requires the interactive identification of two or more emission lines in the raw, or preferably background-subtracted, WLC frame. These lines must also have been successfuly detected by `SEARCH/ECHELLE` in the extracted frame. Since the user is asked to provide the laboratory wavelengths of the calibration lines as well as the absolute echelle order number of the first line, it is necessary to choose and identify these lines beforehand. This was done by plotting the intensities along an order in the two-dimensional frame (e.g., via `EXTRACT/TRACE`) and comparing them to the plots in the atlas of the thorium-argon spectrum (S. D'Odorico et al. 1987). Once certain groups of lines were recognized, it was easy to confirm that the CASPEC exposures at hand cover the spectral orders from 108 to 150.

The first wavelength calibration was performed interactively on frame casp0030.bdf using the `IDENTIFY/ECHELLE` method `PAIR`, which exploited the fact that the overlap in the wavelength regions covered by the spectral orders caused some emission lines at the red end of one order to appear again at the blue end of the next order. The algorithm uses the four locations of two such lines, chosen by cursor, to derive a set of global dispersion coefficients in an initial, third-degree polynomial approximation of the relation in equation (4.8). An iterative loop improves on the global relation while it identifies more lines by computing the estimated wavelengths and comparing them to the catalog wavelengths. When no more new lines can be identified, the algorithm goes on to compute dispersion coefficients and residuals for each order individually and writes them to the table line.tbl. The calculated wavelength step, i.e., the number of angstroms per pixel, is written to the line.tbl descriptor PIXEL as well as to the echelle session keyword SAMPLE, for use in the rebinning step (see below).

Since all of the WLC exposures were taken with the same instrumental set-up, the dispersion coefficients derived from casp0030.bdf could then be used as initial values in the other wavelength calibrations that followed. This was done using the `IDENTIFY/ECHELLE` method `GUESS`, which eliminated the need for further user interaction. As in the first WLC frame, the wavelength step was determined to be 0.1267 Å per pixel in all of these cases.

The wavelength calibration on each of the four WLC frames listed in table 4.1 yielded acceptable results. `SEARCH/ECHELLE` located between 647 and 713 emission lines in each of these frames, and the percentage of identifications among the brightest half of the lines used by `IDENTIFY/ECHELLE` was between 92% and 94% in every case. The differences between the catalog wavelengths and those obtained from the fit of the dispersion relation for the lines selected by `IDENTIFY/ECHELLE` were generally less than 0.02 Å, or less than 1/6 of a pixel. The residuals in each order of the first WLC frame are shown in figure 4.3, as are the histograms of these values.

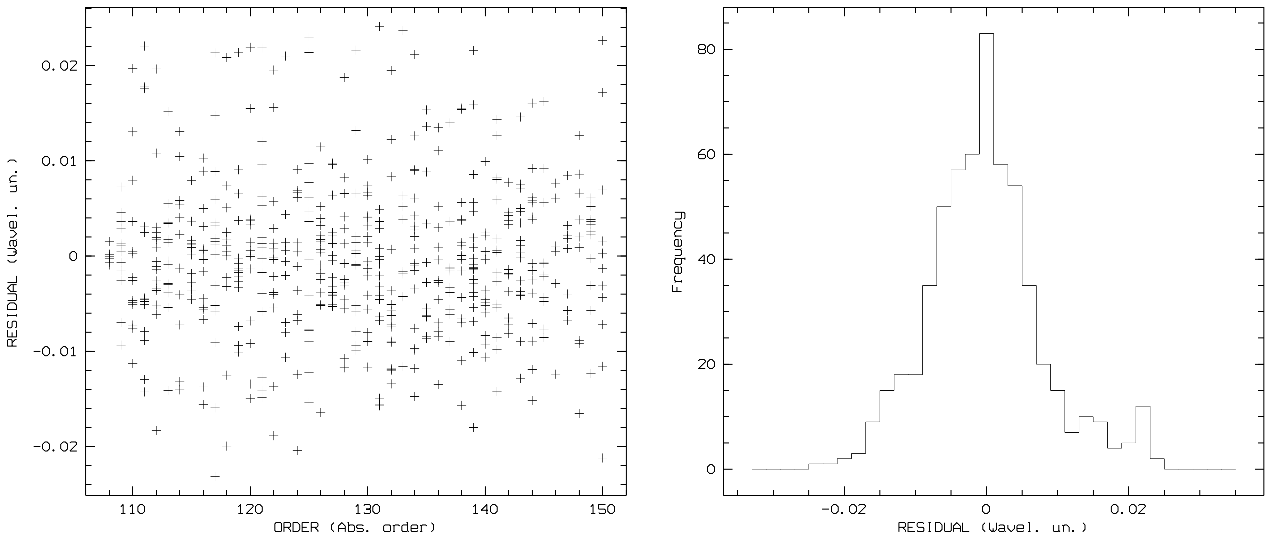


Figure 4.3: *Residuals from the wavelength calibration for casp0031 and histogram of residual values.*

4.7 Rebinning in Wavelengths

Once a wavelength calibration has been done, the stored dispersion coefficients and wavelength step can be used to convert any extracted frame into a so-called rebinned frame. The command `REBIN/ECHELLE` takes an extracted frame in pixel-order space and samples it in uniform wavelength steps, thus converting the pixel coordinates in each order of the input frame into wavelengths. The result is a rebinned frame in wavelength-order space, whose y-axis is still the relative order number but whose x-axis is now in wavelength units. The starting wavelength and the number of pixels in each order are written to the newly created descriptors WSTART and NPTOT in the output frame.

The two-dimensional rebinned MIDAS frame is, however, not the final result. Each spectral order that occupies a row in a rebinned frame has a different starting wavelength, yet shares the same pixel coordinates in x as all of the other orders. So the pixel coordinate x—though it is now in wavelength units—represents not the absolute wavelength but rather the relative wavelength, i.e., the distance in angstroms from the starting wavelength of each order. This is why a plot of one of the rows in a rebinned frame will show the flux values of that order on the y-axis, but not the absolute wavelength scale on the x-axis.

4.8 Order Merging (I)

The spectrum with the absolute wavelengths is obtained via the MIDAS command `MERGE/ECHELLE`, which takes a rebinned frame as input and produces a one-dimensional frame with the wavelengths along the x-axis and the measured fluxes contained in the pixel values. The algorithm first identifies the regions in which two orders overlap, then takes the weighted average of the pixels in those regions. Since no variance or error frames are used in the standard MIDAS echelle reduction, the weights are determined under the assumption that the closer a pixel is to the edge of an echelle order, the less reliable its value: the normalized weights are a linear function of position within the overlap, so that the pixel weights in the bluer order decrease with wavelength, in even steps from 1 to 0, while the pixel weights in the redder order increase steadily from 0 to 1. In the resulting complete spectrum, flux values in the overlapping regions covered by two orders are replaced by the weighted averages.

It was at this point that the complete arc-lamp spectra were obtained from the rebinned WLC frames. For the WLC exposures, this was the final step in the reduction process needed to produce usable one-dimensional spectra for later use (see section 6.1). For the QSO frames, further corrections required the use of the rebinned standard-star frames (see sections 4.9 and 4.10). The subsequent order merging for the QSo frames was also more complicated (sections 4.10 and 5.2).

4.9 Response Correction

Although flat-fielding corrects many of the variations in the pixel-to-pixel sensitivity of the CCD, two further factors affect the measured flux. The first is the echelle blaze effect, which results from variations in the efficiency of the spectrograph as a function of position along the echelle order; this is because the efficiency of the echelle grating is generally higher in the center of each order and decreases on edges. The second effect is the chromatic response of the instrument, which means simply that the detector sensitivity also varies as a function of wavelength. One reason is that photons of different wavelengths are absorbed at different depths in the CCD and the material is not homogeneous (McLean 1997). Because the color temperature of the artificial lamps used for dome flats differs greatly from that of the night sky, flat-fields of this type are not very effective in correcting for the blaze function and the instrument response, so the flux calibration must be done as an additional step.

Since the small slit length used to obtain the present exposures leads to a loss of light, there was no sense in attempting an absolute flux calibration. But the response correction was done anyway, to lessen the effects of the blaze function. This "flattened out" the orders, thus reducing the differences between the overlapping parts of adjacent orders and hopefully minimizing the size and occurrence of "bumps" at these overlaps when the spectrum was merged. The idea was that it would be easier to estimate and fit the continuum if the spectrum was flatter to begin with.

The method recommended in the MIDAS manual, which attempts to correct for both the blaze effect and the chromatic response at once, requires the use of a standard star whose known flux values are stored in a MIDAS table. Once the standard star has been observed with the same instrumental configuration and reduced in exactly the same way as the object, the correction factors can be calculated by comparing the relative flux values obtained from the standard-star frame with the absolute flux values listed in the table. Specifically, the command `RESPONSE/ECHELLE` begins by taking a standard-star frame as input. It subtracts the background, optionally performs the flat-field correction, extracts the spectrum, divides the extracted frame by the exposure time in seconds, and does the rebinning. Then a median filter is applied (in which pixel values that differ too much from the median of the surrounding pixels are replaced by that median value), followed by a smoothing filter (in this case similar to the median filter, but using the average rather than the median).

The algorithm continues by converting the values in the flux table into a frame, then dividing this frame by the rebinned and smoothed standard-star frame to produce the so-called response frame. The pixel values in the response frame represent the known flux divided by the observed flux; they are essentially the factors by which each pixel in an object frame should be scaled to yield the corrected flux.

As noted in the MIDAS manual, one source of possible inaccuracies is the difference in sampling step between flux table and standard-star frame. In this case the available flux table for μ Columbae was sampled at intervals of 9 Å, so its flux values had to be interpolated to the rebinned standard-star frame's smaller sampling step of 0.1267 Å before the frame division could be carried out.

A much more critical factor was the presence of absorption lines in the standard-star frame which are not in the table. These regions would have led to an incorrect response and had to be edited manually. This, and the fact that the standard-star extraction was done with the Hamburg procedures rather than with the usual MIDAS commands, required several modifications to the response correction described above.

First, the standard-star frame casp0078.bdf was background-corrected, extracted according to the Hamburg procedures, divided by the exposure time, and rebinned. The individual order spectra were obtained as 43 separate frames via `MERGE/ECHELLE NOAPPEND`. These order frames were converted into tables, where the regions without any absorption lines could be selected interactively and fit using fifth-degree polynomials. (The polynomials produced smoother curves and a closer fit than the median filter and smoothing normally done in `RESPONSE/ECHELLE`.) These order fits were free of both cosmics and absorption lines and were therefore suitable for calculating the response correction; but since the response frame is normally a rebinned frame, all 43 fits were copied into a two-dimensional frame with the same format and descriptors as a rebinned frame. The procedure `RESPONSE/ECHELLE` was modified so that it would use this optimally extracted, rebinned frame with the fitted orders rather than using the standard MIDAS procedures to reduce a raw standard-star frame from scratch. Thus it created the frame with the absolute fluxes as usual, but then divided that frame by the fitted rebinned frame to yield a response frame much less corrupted by extraneous absorption.

Figure 4.4 shows a sample standard-star order, along with the fifth-degree polynomial fit used in the response correction. Figure 4.5 shows the corresponding rows in the response frame, compared to what the response would have been if the absorption regions had not been replaced by the polynomial fits.

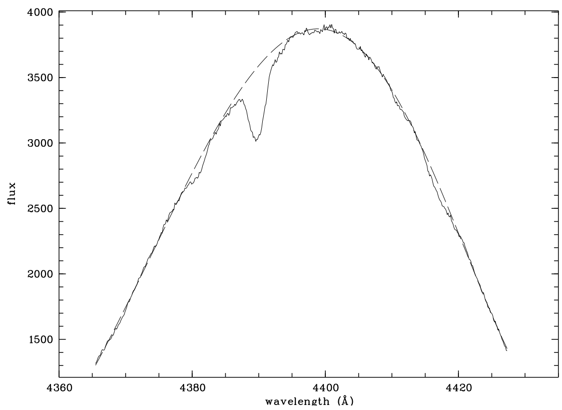


Figure 4.4: *Sample standard-star order 22 (solid line) and fifth-degree polynomial fit (dashed line).*

4.10 Reduction of the QSO Frames

After the wavelength calibrations had been done and the response frame created, the quasar frames could finally be reduced in one go.

First, the initial order definitions for each quasar frame were chosen from one of the appropriate flat-fields. As mentioned earlier, the flat-field and wavelength-calibration

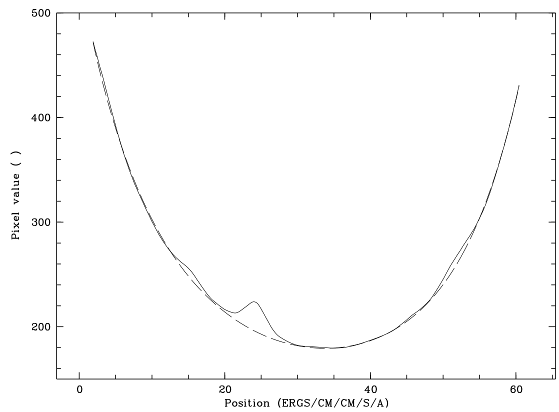


Figure 4.5: *Sample response function for order 22, uncorrected (solid line) and corrected for absorption (dashed line).*

frames used to make corrections were taken just before or after each quasar frame, to avoid introducing additional errors as a result of the positional shifts that take place over the course of the night. Table 4.1 summarizes the calibration frames used for each object exposure.

	QSO frame casp0023	QSO frame casp0031	STD frame casp0078	QSO frame casp0089
BIAS frames	casp0004–casp0007 casp0035–casp0039	casp0004–casp0007 casp0035–casp0039	casp0069	casp0069
DARK frames	casp0034	casp0034	casp0067 casp0070	casp0067 casp0070
FLAT (order def. only)	casp0025	casp0029	casp0077	casp0087
WLC	casp0024	casp0030	casp0076	casp0088
response	casp0078	casp0078	n.a.	casp0078

Table 4.1: *Object and Calibration Frames*

After the order definitions had been chosen, the inter-order background on all three quasar frames was estimated using `BACKGROUND/SMOOTH` and then subtracted. At this point the background-corrected frames would normally have been divided by the normalized flat-field, but this was not done for the reasons explained above.

A variance frame was now created for each of the quasar frames, and any cosmics that had been identified were tagged with practically infinite variances. The programs written in Hamburg made use of these variance frames to optimally extract the spectral orders from each of the background-corrected quasar frames. The program output was not only a frame in pixel-order space containing the extracted fluxes, but an equivalent frame containing the corresponding errors for each flux value. In all of the steps following the optimal extraction, the error frames were processed in parallel with the data frames, so that these error estimates were available at each step.

The extracted spectrum and the extracted errors obtained from the quasar frame casp0031.bdf were the first to be rebinned. It was first necessary to initialize the session with the appropriate wavelength calibration—and thus ensure that the correct dispersion coefficients were available in line.tbl—before executing the command `REBIN/ECHELLE`. The other two quasar spectra and their errors were also rebinned using the dispersion coefficients from the WLC session appropriate to each of them, but in those two cases the echelle keyword `SAMPLE` was used to define the first rebinned image as a reference frame. This guaranteed that all three spectra would have the same starting wavelength and wavelength step.

The response correction was applied to the three rebinned quasar frames (as well as their rebinned error frames) by simply multiplying each of them by the response frame. The effects of this correction on the fluxes can be seen by the example in Figure 4.6.

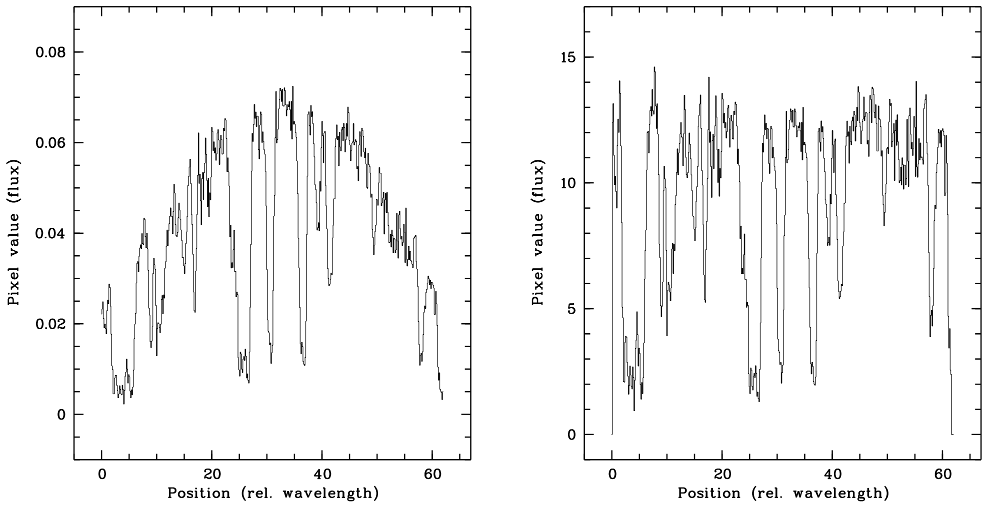


Figure 4.6: *Sample echelle order (22) before and after response correction.*

The final task in the normal MIDAS echelle reduction is to produce one-dimensional spectra from the response-corrected, rebinned frames. For reasons explained below, this was done in several steps. The first of these steps was to apply `MERGE/ECHELLE` to the data and error frames, but this time with the option NOAPPEND, so that nothing more was done than to take each of the orders from a rebinned frame, convert the relative wavelengths along the x-axis to absolute wavelengths, and store each order in a separate frame.

This concluded the main part of the data reduction, which produced three one-dimensional spectra for each of the 43 echelle orders from the three quasar exposures listed in Table 2.2.

Chapter 5

The Spectrum Toward HE 0940

Data reduction was defined as the production of one-dimensional spectra from two-dimensional CCD images. However, a number of technical steps were still necessary in order to achieve a spectrum suitable for fitting absorption lines. A single, normalized spectrum was needed, i.e., a spectrum in which the different quasar exposures had been combined, the orders merged, and the fluxes at each wavelength divided by the estimated continuum level there.

5.1 Combination of QSO Orders

At this point the spectra from each of the three quasar exposures were combined into one spectrum per order. This not only reduced the errors and increased the signal-to-noise-ratio, but also further minimized any traces of cosmics that might have survived the data reduction. Since the error frames produced during the extraction were also available, it made sense to take advantage of them and use the errors at each wavelength to determine the weighted average for each order.

If each of three data points x_i has been measured with the precision σ_i, then the most probable value is the weighted average of the data points, where the weight w_i given to each point is inversely proportional to the variance of that data point (Bevington and Robinson 1969):

$$\mu = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3}{w_1 + w_2 + w_3} \qquad \text{with} \quad w_i = \frac{1}{{\sigma_i}^2} \qquad \text{for } i = 1,\, 2,\, 3 \tag{5.1}$$

$$\mu = \frac{(x_1/{\sigma_1}^2) + (x_2/{\sigma_2}^2) + (x_3/{\sigma_3}^2)}{(1/{\sigma_1}^2) + (1/{\sigma_2}^2) + (1/{\sigma_3}^2)} \tag{5.2}$$

Assuming that the fluctuations in the x_i's are uncorrelated, the variance of the weighted average is the inverse of the sum of the weights:

$${\sigma_\mu}^2 \simeq {\sigma_1}^2 \left(\frac{\partial \mu}{\partial x_1}\right)^2 + {\sigma_2}^2 \left(\frac{\partial \mu}{\partial x_2}\right)^2 + {\sigma_3}^2 \left(\frac{\partial \mu}{\partial x_3}\right)^2 \tag{5.3}$$

$${\sigma_\mu}^2 \simeq \frac{1}{(1/{\sigma_1}^2) + (1/{\sigma_2}^2) + (1/{\sigma_3}^2)} \tag{5.4}$$

These formulas could not be applied directly to the data, however, because the flux levels were not the same in all three spectra. For most orders, the fluxes in casp0031.bdf were higher than in the other two spectra. In an attempt to correct for this effect, the three median values were used to scale the casp0023 and casp0089 spectra to the level of the casp0031 spectrum before all three were averaged.

So the median value m_i of each order spectrum was used to scale that spectrum by a factor of $c_i = m_1/m_i$. The weighted average of the *scaled* spectra thus became

$$\mu = \frac{w_1(c_1x_1) + w_2(c_2x_2) + w_3(c_3x_3)}{w_1 + w_2 + w_3} \quad \text{with} \quad w_i = \frac{1}{{c_i}^2{\sigma_i}^2} \tag{5.5}$$

$$\mu = \frac{x_1/(c_1{\sigma_1}^2) + x_2/(c_2{\sigma_2}^2) + x_3/(c_3{\sigma_3}^2)}{1/({c_1}^2{\sigma_1}^2) + 1/({c_2}^2{\sigma_2}^2) + 1/({c_3}^2{\sigma_3}^2)} \tag{5.6}$$

with a variance of

$${\sigma_\mu}^2 \simeq \frac{1}{1/(c_1\sigma_1)^2 + 1/(c_2\sigma_2)^2 + 1/(c_3\sigma_3)^2} \tag{5.7}$$

After each of the orders were combined according to equation (5.6), they were kept separate at this point in case order-by-order corrections might later become necessary.

Of course it would have been possible to merge the orders from each QSO exposure into a full spectrum first, and then combine the full spectra from the three exposures. But it would have been much more difficult to scale the full spectra appropriately. This was because the flux levels differed not only from one exposure to another, as seen above, but also within the different wavelength regions of each exposure. For example, the second QSO exposure had slightly more flux at lower wavelengths, but clearly less flux at higher wavelengths, than the third QSO exposure.

5.2 Order Merging (II)

Normally, the next step would have been to estimate the continuum in each of the orders, then to normalize the orders and merge them to produce the complete spectrum. However, when it was noticed that the measured flux did not drop to zero in the middle of obviously saturated lines, it was clear that not all of the sky background had been effectively subtracted and that a sky-subtraction step had to be added.

In order to enable a global estimate of the sky background, the 43 combined orders had to be merged into a complete spectrum at this point. This was not such a simple matter, however. Since `MERGE/ECHELLE` requires a rebinned frame as input, the procedure necmerge.prg and the source-code file necmerge.for (MIDAS release 95NOV) were modified so that many separate order frames would be accepted instead. The second, more important difficulty was that the `MERGE/ECHELLE AVERAGE` algorithm had to be adapted to enable it to merge the error frames as well as the data frames. The fluxes could be averaged as described in section 4.7:

$$\mu = w_1 x_1 + w_2 x_2, \qquad \text{with } w_1 + w_2 = 1 \tag{5.8}$$

The corresponding errors, on the other hand, could not be averaged in the same way as the fluxes. According to the rules of error propagation (cf. equation 5.3), the error σ_μ for the weighted average is given by

$${\sigma_\mu}^2 = {w_1}^2 {x_1}^2 + {w_2}^2 {x_2}^2 \tag{5.9}$$

That is, the weights used for the variances must be the *squares* of the weights used for the data.

The modified versions of necmerge.prg and necmerge.for were therefore given an additional input parameter to specify whether the pixel values being averaged were fluxes or variances; the modified algorithm assigned the normalized weights from 0 to 1 accordingly. Thus the full spectrum (Fig. 5.1) resulted from the equivalent of `MERGE/ECHELLE AVERAGE`, while the corresponding error frame was obtained by squaring all of the individual order frames, merging them using the variance option of the modified procedure, and then taking the square root of the resulting output frame.

5.3 Heliocentric and Vacuum Corrections

It is standard practice to correct the wavelengths in a spectrum for the Earth's orbital motion along the line of sight, as well as to convert the measured wavelengths to their vacuum values. These corrections were made for the 43 individual order frames as well as for the merged spectrum described in section 5.2.

An object's heliocentric velocity v_H (i.e., with respect to the Sun) is determined by removing the velocity component resulting from the Earth's orbital motion. For an object with right ascension α_o and declination δ_o, the heliocentric radial velocity is related to its geocentric velocity v_G (i.e., measured from the center of the Earth) by

$$v_H = v_G + v_E[\sin\delta_A \sin\delta_o + \cos\delta_A \cos\delta_o(\alpha_o - \alpha_A)] \tag{5.10}$$

where v_E is the Earth's orbital velocity (29.78 km s^{-1} on average),

$$\alpha_A = \arcsin(\cot\delta_s \sin\alpha_s \sin\varepsilon) \tag{5.11}$$

$$\delta_A = \alpha_s - \arccos(-\tan\delta_A \tan\delta_s) \tag{5.12}$$

are the right ascension and declination of the apex of the Earth's orbital motion at the time of the observation, ε is the obliquity of the ecliptic ($23°27'$), and α_s and δ_s are Sun's right ascension and declination at the time of the observation (Kitchin 1995). Clearly, the difference between the geocentric and heliocentric velocities—which is the

Earth's heliocentric radial velocity—causes a shift $\Delta\lambda$ in the wavelength according to the Doppler effect,

$$\frac{\Delta\lambda}{\lambda} = -\frac{\Delta\nu}{\nu} = \frac{v}{c} \tag{5.13}$$

where c is the speed of light.

The calculations were done automatically with the MIDAS commands `COMPUTE/PRECESSION`, which precessed the celestial coordinates of HE 0940 (listed in Table 2.1) to the day of observation, followed by the command `COMPUTE/BARYCORR`, which computed the Earth's heliocentric radial velocity at the desired time. The total heliocentric correction to the radial velocity ranged between 7.6 km s^{-1} and 8.7 km s^{-1}, which led to a wavelength correction of approximately 0.1 angstroms.

Converting the wavelengths to their vacuum values is a larger correction that requires an expression for the index of refraction n of the Earth's atmosphere. For convenience, an estimate was used that applies to optical wavelengths (2000–7000Å) and is independent of the time and location of the exposure:

$$(n-1) \times 10^7 = 643.28 + \frac{294981}{146-\sigma^2} + \frac{2554.0}{41-\sigma^2} \tag{5.14}$$

where σ is the wave number in vacuum ($1/\lambda_{\rm vac}$), in units of reciprocal microns (Lang 1980). The vacuum wavelength is merely

$$\lambda_{\rm vac} = n\ \lambda_{\rm air} \tag{5.15}$$

In the case of the HE 0940 spectra, the correction $\Delta\lambda = \lambda_{\rm vac} - \lambda_{\rm air}$ resulted in shifts between 1.07 and 1.47 Å, also comparable to the approximate value $n \approx 1.000297$ that is often used (Lang 1980).

On a technical note, the preceding corrections represented a change in the wavelength axis of the spectra and resulted in a non-uniform wavelength step from one pixel to the next. But MIDAS frames currently allow only a single step size. So during these corrections, a table was created to hold the wavelengths, fluxes and errors from each input frame and its corresponding error frame. The data were subsequently kept in that form, and all subsequent steps were performed on these tables.

5.4 Sky Subtraction

Because of the small slit size used for the object exposures, the quasar most probably filled the entire slit. In any case, the spatial profile along the slit was examined before and during the spectrum extraction and no sky background had been recognizable; so the sky background could not be extracted separately. There was no other choice but to estimate the level of the sky background at this point by examining the line centers of saturated lines.

It was assumed that the sky background would be a slowly varying function of the wavelength. A global fit was therefore thought to be preferable to order-to-order fits, which would not only necessitate smoothing of discontinuous fits in the overlaps between two orders but probably lack sufficient data points in some of the orders anyway. That is why the combined orders were merged as described in section 5.2, even though the continuum definition was later to be done on each of the orders individually.

Points at the center of lines judged to be obviously saturated were selected manually. The sky background level was estimated by fitting a fourth-degree polynomial through these 492 out of a total of 11,710 data points. The standard deviation of the fit was 0.45, and the resulting sky level ranged between 0.67 and 2.54, in the arbitrary flux units of the spectrum shown in Figure 5.1.

Since there were no saturated lines redward of the line at 4910 Å, there was no way to know how the sky background should be extrapolated to higher wavelengths. The fourth-degree polynomial fit was used up to 4910.3 Å, and the fit value there (1.9557) was used for all wavelengths beyond that point. This value is doubtless a bit high in the area of the C IV λ1548 lines around 5155 Å; six pixel values fall below the estimated sky there. But in the absence of any information about the sky level, it was thought best to use the constant value rather than making any further assumptions about the shape of the sky spectrum or allowing a discontinuity by assuming no sky background above 4910.3 Å. In the end, this was also the estimate that was subtracted from all of the flux values in the individual orders, just before a fit of the continuum was attempted.

5.5 Spectrum Normalization

The next step consisted of normalizing the spectrum, i.e., scaling the flux so that the pixels in which there was no absorption would have pixel values that fluctuate (at levels consistent with the noise there) about a constant level of 1.

In addition to the reasons mentioned in section 5.1, the orders were first combined (rather than merged) because it proved much easier to estimate the continuum for each of the 43 echelle orders individually. This also allowed more flexibility in fitting the edges of the orders, where the lower fluxes meant that the background subtraction and response correction had more strongly affected the determination of the flux than in the center of the order.

The most common method of estimating the continuum in echelle spectra is to begin by fitting polynomials to the spectrum and discarding the pixel values that differ more than a certain number of standard deviations from the fit. The process is then repeated iteratively, usually until the resulting residuals correspond to levels consistent with the expected noise. The aim is to remove from consideration any pixels

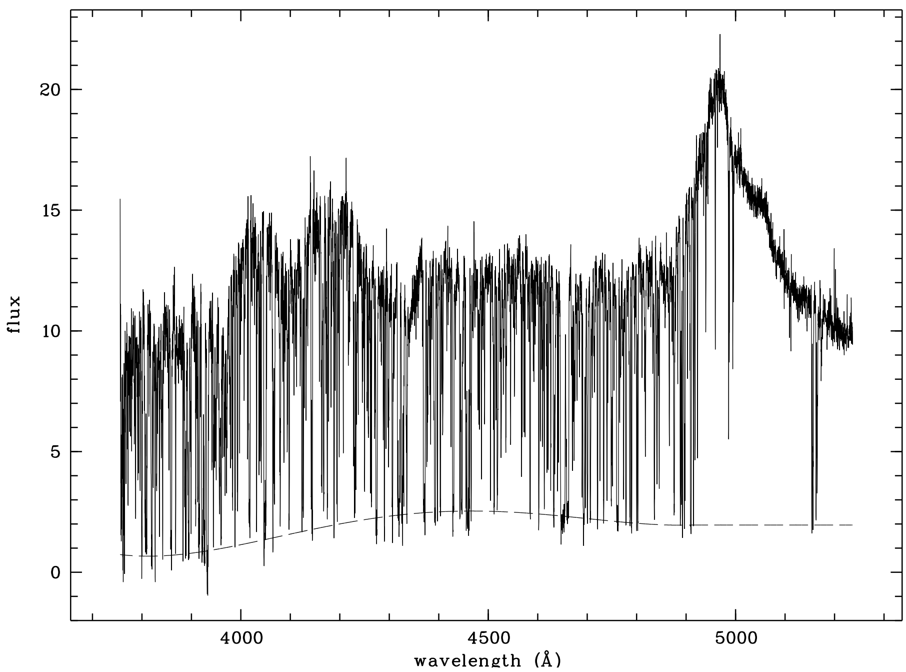


Figure 5.1: *Preliminary merged spectrum with estimated sky background.*

belonging to absorption or emission lines, noise spikes, or cosmics before performing the final fit that represents the estimated continuum.

Strictly speaking, normalizing the flux values in the spectrum is not quite the same as estimating the continuum, because a continuum fit in the above sense would aim to ignore emission lines as well as absorption lines. When a spectrum is normalized, however, any emission lines that might have been present are fit along with the continuum to produce the absorption spectrum, so only pixel values below a certain cut-off would be excluded from the fit. Nevertheless, such a fit is also referred to as the estimated continuum, and the normalized spectrum is created by dividing the flux by this continuum.

In this particular case, no automated fitting procedure was used. Instead, points in each echelle order were first selected as belonging to the continuum and then used to determine the fit. The principal difficulty lay of course in making the subjective judgment as to which pixels were unaffected by absorption. Under the necessary assumption that at least some points in each order did belong the continuum, an attempt was made to select all of the pixels that did not appear to be part of absorption lines. Groups of adjacent pixels were selected rather than only the very highest values, to account for the fluctuations due to noise. The merged spectrum obtained in section 5.2 was also consulted during the selection process since it provided some visual guidance as to the global continuum level.

The selection and fitting were made more difficult by the fact that the large number of absorption lines in most of the orders left very little continuum available, so that usually only very small groups of points could be chosen. The number of values used in each fit reached a maximum of 514 in order 39, and included well over 60% of the pixels in all of the other echelle orders redward of Lyman α emission. The number of fit points dropped to a minimum of only 24 pixels in order 7 (see Table 5.1); but such low values are no surprise in orders 1 to 15, which lie blueward of Lyman β emission and thus contain additional lines from Ly β and even Ly γ absorption. In most cases within the primary region of the Ly α forest, however, between 80 and 100 values were used in each fit. The continuum is estimated to cover between 15% to 20% of the spectrum in most orders. The markedly higher percentages in orders 20 and 21 are due to the presence of a possible void, which will be investigated in chapter 9.

Once the input values had been selected, second- or third-degree polynomial fits were then performed on the chosen pixels. Low-degree fits were purposely used because some orders lacked fit points on one or both edges—and in such cases higher-degree fits tend to fluctuate widely, curving sharply upward or downward depending on the degree. Only in three echelle orders did polynomials of degree 4 or 5 seem necessary.

It should be noted in passing that since this type of method assumes that the highest pixel values in each order do indeed represent continuum levels, it will tend to consistently underestimate the continuum level by ignoring smoothly fluctuating absorption such as that from weak, very broad lines or, on a cosmological scale, from diffuse intergalactic hydrogen (the Gunn-Peterson effect). Of course it would be difficult to argue that the interactive fitting was able to correct more for this bias than automated fitting. The latter would probably have yielded more consistent, or at least more quantifiable, results from one order to the next. Admittedly, however, the continuum fitting was done by hand less for consistency than for educational purposes—to gain first-hand experience with the problems inherent in continuum estimates for spectra so densely populated with absorption lines.

5.6 Order Merging (III)

The normalized orders and their correspoding errors were now merged using the same algorithm as described in section 5.2, although some changes were necessary now that the input was present in the form of tables rather than one-dimensional images. The algorithm was also modified further in order to obtain the greatest possible wavelength coverage in the merged spectrum.

When the MIDAS command `MERGE/ECHELLE AVERAGE` merges two echelle orders, it assumes that the overlap between them is bounded by the minimum of the redder order on the low-wavelength end and by the maximum of the bluer order on the high-wavelength end. This is usually the case, so usually the interval whose length corresponds to the echelle keyword `DELTA` is excluded from the appropriate end of each order. This exclusion feature was retained in the modified procedures because despite normalization, the flux values on the ends of each order showed occasional spikes and were generally not very reliable. An interval of 5 Å was chosen to avoid these regions without sacrificing the good overlap already present.

Although the Hamburg procedures had allowed the optimal extraction of a greater spectral region than would have been possible otherwise, only a small part of order 43 was present and its maximum wavelength was less than the minimum wavelength in order 42. In this case, the normal MIDAS algorithm would have ignored the highest wavelengths in order 42 and kept the much less reliable values at the end of order 43. The modified algorithm, on the other hand, first checked the positions of the order boundaries when deciding at which ends to exclude values before averaging the overlaps. The final normalized and merged spectrum of HE 0940 was obtained via this modified algorithm.

Relat. order	Echelle order	Minimum wavelength	Maximum wavelength	Data values	Contin. values	Contin. coverage	poly. degree
1	150	3755.98	3808.70	417	71	17 %	3
2	149	3781.20	3834.17	419	94	22 %	2
3	148	3806.80	3860.03	421	90	21 %	3
4	147	3832.53	3886.26	425	61	14 %	3
5	146	3858.76	3913.13	430	52	12 %	2
6	145	3885.37	3939.87	431	40	9 %	3
7	144	3912.37	3967.37	435	24	6 %	2
8	143	3939.74	3994.00	437	54	12 %	2
9	142	3967.37	4023.13	441	123	28 %	2
10	141	3995.50	4051.65	444	154	35 %	5
11	140	4024.02	4080.54	447	71	16 %	2
12	139	4052.92	4109.82	450	57	13 %	2
13	138	4082.32	4139.73	454	47	10 %	2
14	137	4112.10	4169.89	457	78	17 %	2
15	136	4142.14	4200.56	462	101	22 %	2
16	135	4172.93	4231.61	464	133	29 %	2
17	134	4203.98	4263.04	467	93	20 %	3
18	133	4235.54	4295.10	471	93	20 %	2
19	132	4267.60	4327.67	475	67	14 %	4
20	131	4300.17	4360.62	478	229	48 %	3
21	130	4333.25	4394.21	482	217	45 %	4
22	129	4366.83	4428.30	486	83	17 %	2
23	128	4400.80	4462.77	490	90	18 %	2
24	127	4435.40	4497.88	494	98	20 %	2
25	126	4470.63	4533.49	497	148	30 %	2
26	125	4506.37	4569.73	501	81	16 %	2
27	124	4542.74	4606.61	505	92	18 %	2
28	123	4579.62	4644.13	510	90	18 %	2
29	122	4617.00	4682.02	514	77	15 %	2
30	121	4655.28	4720.67	517	74	14 %	2
31	120	4693.93	4760.09	523	55	11 %	2
32	119	4733.35	4800.01	527	52	10 %	2
33	118	4773.39	4840.69	532	108	20 %	2
34	117	4814.08	4881.88	536	85	16 %	3
35	116	4855.64	4923.95	540	103	19 %	2
36	115	4897.85	4966.79	545	284	52 %	3
37	114	4940.68	5010.39	551	384	70 %	3
38	113	4984.41	5054.62	555	393	71 %	2
39	112	5028.76	5099.61	560	514	92 %	2
40	111	5074.01	5145.61	566	475	84 %	2
41	110	5120.14	5192.38	571	362	63 %	2
42	109	5167.03	5239.90	576	374	65 %	3
43	108	5214.81	5236.35	171	171	100 %	2

Table 5.1: *Echelle orders and continuum fits*

Chapter 6

Data Analysis

Now that a normalized spectrum had finally been obtained, the actual analysis of the data could begin. The principal task here was to determine the parameters of all of the lines in the Ly α forest of HE 0940 and thus produce a suitable sample for studying the proximity effect. As part of that process, an attempt was made to identify as many metal lines as possible so that they could be removed from the Ly α sample.

The present spectrum has complete wavelength coverage of the primary region between the quasar's Ly β and Ly α emission lines at 4158.3 Å and 4928.3 Å, respectively. Lines in the spectrum short of Ly β emission were not fitted unless they belonged to systems containing lines in the primary region of interest. This was because the many, usually stronger Ly α lines and the large amount of blending in the Ly β forest make it difficult to reliably identify all of the higher-order Lyman lines. The resulting contamination in the Ly α sample is difficult to quantify. That is why studies of this type invariably omit the spectral region below 1025.7223 Å in the quasar's rest frame.

Before the labor-intensive process of interactive line fitting could begin, however, two last technical steps were required. The first was to measure the width of the instrumental profile, which is needed as input to the fitting programs. The second was to determine the detection limit for absorption lines in terms of a minimum column density, both for consistency and to avoid wasting time trying to fit noise.

6.1 Instrumental Profile

The observed spectrum is the convolution of the true spectrum with the so-called instrumental profile, so that the observed intensity at a given wavelength λ is

$$I(\lambda) = \int_0^\infty T(\lambda')\, R(\lambda - \lambda')\, d\lambda' \tag{6.1}$$

where $T(\lambda')$ is the intensity in the true spectrum at λ' and $R(\lambda - \lambda')$ is the response of the spectroscope at a wavelength distance $(\lambda - \lambda')$ from the center (Kitchin 1991).

An estimate of the instrumental profile could be made by measuring the FWHM (full width at half the maximum intensity) of the emission lines in the wavelength-calibration spectra obtained in section 4.8. They spectra were used because an ideal spectroscope and a monochromatic source would produce an infinitely thin emission line at a single wavelength, while diffraction within a real detector causes each emission line to be spread over an entire range of wavelengths. This spread describes the instrumental profile and determines the resolution that can be achieved with the instrument at hand. For a real but optically perfect spectroscope, the instrumental profile would have the shape of the Fourier transform for a regularly spaced finite set of delta functions, modulated by the Fourier transform for a single aperture (Kitchin 1995). Errors in the grating or misalignments in the detector components will obviously lead to devations from this profile and increase the spread even more.

This information is vital for the subsequent measurements of absorption lines because the fitting algorithm must be able to separate out how much of the line width is due to instrumental effects rather than from the instrinsic line width, Doppler broadening and pressure broadening. This can be done on the basis of equation (6.1) if the appropriate input is provided, e.g., the FWHM values at each wavelength.

The FWHM measurements were performed on the three merged WLC frames from which the dispersion coefficients for the three quasar exposures had been obtained. A total of 203 emission lines were chosen in roughly equal intervals along the entire spectrum, and an effort was made to choose only lines that appeared unblended. To avoid selection effects that would lead to differences from one spectrum to another, the exact same transitions were measured in all three spectra.

As often done, the instrumental profile was assumed to be Gaussian, so the line widths were measured by fitting a Gaussian to each emission line using the MIDAS command `CENTER/GAUSS`. The command requires as input the endpoints of the interval in which the fit is to be performed. Strangely enough, it produced different results depending on whether this user input was provided via cursor or via an input table. Measurements of artificially generated Gaussians showed that the former method gave the correct results, so the measurements were done interactively in the graphics window. Fig. 6.1 shows the individual FWHM measurements for one of the three frames. The average FWHM of all the measurements in the three frames, over the wavelength range of the entire spectrum, is 0.26 Å. Using the Doppler relation in equation (5.13) with $\Delta\lambda$ = FWHM and λ_o now the observed (redshifted) wavelength λ_c of the center of each line, these values yielded an average FWHM resolution of 17.7 km s^{-1}.

Fig. 6.1 also shows the linear least-squares fit of the FWHM as a function of the vacuum wavelength in the heliocentric frame. The corresponding estimates for the instrumental resolution as $R = \Delta\lambda/\lambda$ are included merely to show that the linear least-squares fit through the estimated instrumental resolution is, as expected, nearly constant with wavelength. Because of the substantial scatter in the measured FWHM values, the instrumental FWHM assigned to each wavelength in the normalized spectrum was taken as the simple average of the three linear fits at that point.

Some fitting programs (section 6.4) require the line width in terms of the instrumental b value. As with a Gaussian distribution

$$I(\lambda) = I_{\max}\ \exp\left[-\frac{1}{2}\left(\frac{\lambda-\lambda_c}{\sigma}\right)^2\right] \tag{6.2}$$

(where λ_c is the line center), the relation between the full width at half-maximum and standard deviation σ is

$$\sigma = \frac{\frac{1}{2}\,\mathrm{FWHM}}{\sqrt{2\ln 2}} \approx \frac{\mathrm{FWHM}}{2.35482} \tag{6.3}$$

The instrumental b value, which is defined as

$$b = \sigma\sqrt{2} = \frac{\mathrm{FWHM}}{2\sqrt{\ln 2}} \tag{6.4}$$

is most often expressed in units of km s^{-1}. This can be calculated from the FWHM, again with the help of the Doppler relation in equation (5.13), by setting $\Delta\lambda$ equal to b in wavelength units and Δv equal to b in velocity units.

$$b\ [\mathrm{km/s}] \quad = \quad \frac{\sigma\sqrt{2}}{\lambda_c}\,c \quad = \quad \frac{\mathrm{FWHM}\ [\text{Å}]}{2\sqrt{\ln 2}}\ \frac{c\ [\mathrm{km/s}]}{\lambda_c\ [\text{Å}]} \tag{6.5}$$

where λ_c is again the observed central wavelength of the line in question.

The physical meaning of the b value can be envisioned when assuming a Maxwell-Boltzmann velocity distribution (which applies, e.g., to the random thermal motions of atoms of a gas in thermal equilibrium). The resulting line profile is a Gaussian with a full width at half-maximum of

$$\mathrm{FWHM} = 2\sqrt{\ln 2}\,\frac{\lambda_c}{c}\sqrt{\frac{2kT}{m}} \tag{6.6}$$

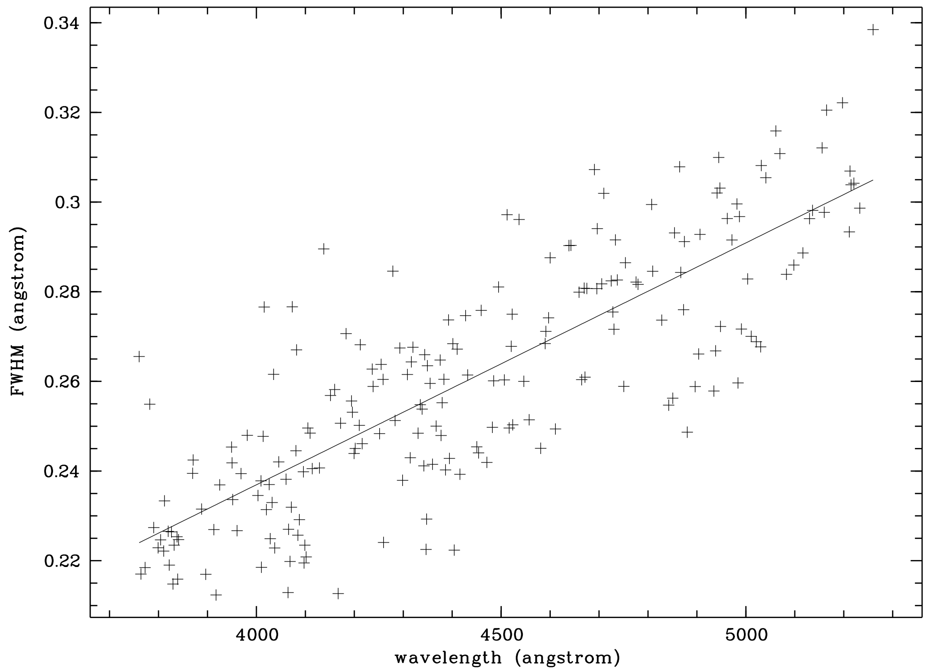
From equation (6.5) we see that

$$b\ [\mathrm{km/s}] = \sqrt{\frac{2kT}{m}}, \tag{6.7}$$

which is also the most probable speed. So, very loosely speaking, b in these units corresponds to the most probable speed in the Maxwell-Boltzmann distribution that would be needed to produce the same line width via Doppler broadening. The difference between the instrumental b value and the b value of a spectral line is that in the former case, the broadening is caused not by particle motions but by the finite spectral resolution of the instrument.

6.2 Astrophysical Quantities

The lines seen in an absorption spectrum can be characterized by the three astrophysical parameters of redshift z, column density N and Doppler width b. Since the main



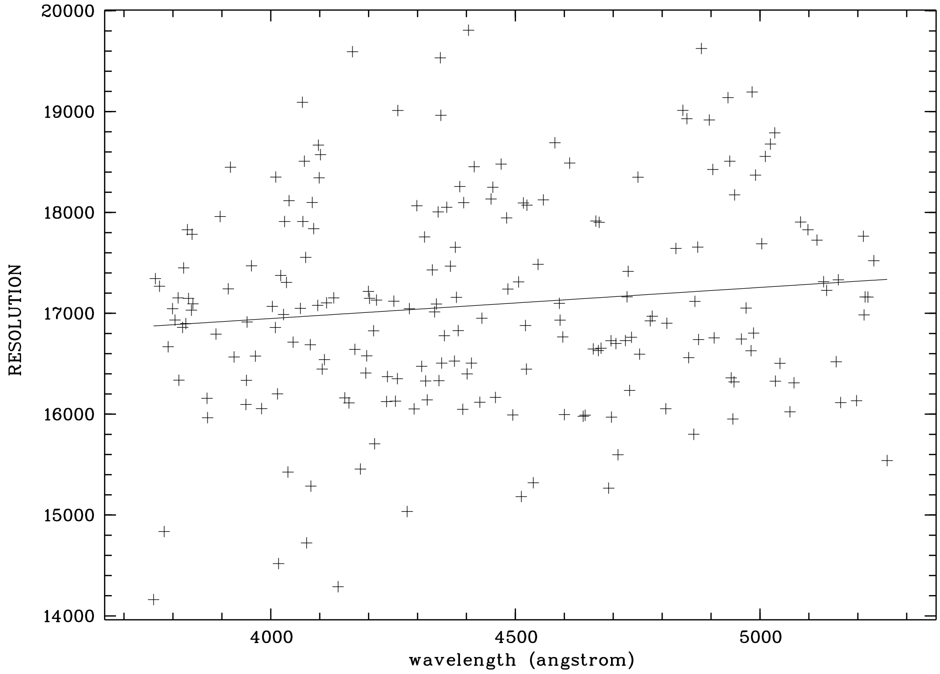


Figure 6.1: *Measurements of instrumental FWHM values for selected arc-lamp emission lines in one of three WLC frames, and corresponding estimates for the instrumental resolution.*

task in the next part of this work was to determine these parameters for the lines in the spectrum of HE 0940, they will be introduced briefly here.

The first absorption-line characteristic to be determined is the wavelength, where we distinguish between the observed wavelength λ and the rest wavelength λ_o in the rest frame where the absorption took place. To facilitate comparison between different spectra, it is often more convenient to use the redshift z, defined by

$$1 + z = \frac{\lambda}{\lambda_o} \tag{6.8}$$

The strength of an absorption line clearly depends on the number of atoms taking part in the absorption. But because the distances between observers and astronomical objects are difficult to determine, it is impractical to use actual densities. Instead, the observable quantity is the column density N, defined by

$$N = \int n_j \, ds \tag{6.9}$$

where n_j is the number density of the absorbing atoms, more precisely the atoms occupying the lower level of the atomic transition $j \rightarrow k$. The integral is evaluated over the pathlength ds along the line of sight, so that the column density is equal to the number of absorbing atoms per unit area over the entire distance between the source and the observer.

As already hinted in section 6.1, the b value—which goes by various names, such as Doppler parameter, velocity dispersion or Doppler width—is a measure for the width of a spectral line. However, the b value is not the sole determining factor in the line width. Indeed, the measurements of absorption lines are often complicated by the fact that the column density and the b parameter cannot be determined independently of one another. In order to demonstrate this, some basic concepts of radiative transfer are mentioned below. (Unless otherwise noted, this and the formulas in the rest of section 6.2 were taken from Unsöld and Baschek, 1988).

Optical depth

Given an intial beam of radiation $I_\nu(0)$, defined as the energy passing through the unit area per time interval, frequency and solid angle. As its photons interact with matter along the distance ds, the specific intensity I_ν changes by the amount

$$dI_\nu = -\kappa_\nu I_\nu \, ds + \varepsilon_\nu \, ds \tag{6.10}$$

where κ_ν and ε_ν are the so-called absorption and emission coefficients, respectively. The optical depth is defined by

$$\tau_\nu = \int \kappa_\nu \, ds, \tag{6.11}$$

with $\tau_\nu = 0$ at the observer. Thus the specific intensity reaching an observer through a region of total optical thickness τ_{ν_r} is

$$I_\nu = I_\nu(0)\, e^{-\tau_\nu} + \int_0^{\tau_{\nu_r}} \frac{\varepsilon_\nu}{\kappa_\nu}\, e^{-\tau_\nu} d\tau_\nu \tag{6.12}$$

(Spitzer 1978). In the case of pure absorption, the integral on the right is zero. If we also neglect stimulated emission (because otherwise it would be included as negative absorption), the optical depth can be expressed in terms of atomic constants in SI units as

$$\tau_\nu = \pi r_e c\, N f_{jk}\, \phi(\nu) \tag{6.13}$$

where $r_e = e^2/(4\pi\epsilon_o m_e c^2) = 2.817941$ fm is the classical electron radius, f_{jk} is the upward oscillator strength for the transition $j \rightarrow k$, and $\phi(\nu)$ is the line profile function.

Line Profiles

The integral of the line profile function $\phi(\nu)$ over the line is normalized to unity,

$$\int \phi(\nu)\, d\nu = 1, \tag{6.14}$$

while the profile function itself is a complicated function of the distance $(\nu - \nu_c)$ from the center ν_c of the line. It determines the shape of the line and depends on the physical processes taking place in the gas along the line of sight.

The line profile due to natural line broadening, which results from the intrinsic width of energy levels due to their finite lifetimes, is the Lorentz or damping profile

$$L(\nu) = \frac{\gamma}{4\pi^2\,(\nu - \nu_c)^2 + \gamma^2/4} \tag{6.15}$$

where γ is the damping constant and ν_c is the frequency of the line center. Pressure broadening due to nearby atoms also produces this profile in weak absorption lines, in which case γ becomes the sum of the damping constants for each process.

In addition to the natural line broadening and the pressure broadening, there is also the Doppler broadening due to the thermal, possibly turbulent velocities of the absorbing atoms along the line of sight. Assuming these atoms have a Maxwell-Boltzmann velocity distribution, then the profile function assumes a Doppler profile

$$D(\nu) = \frac{c/\nu_c}{b\sqrt{\pi}}\, e^{-\left[\frac{c(\nu-\nu_c)}{b\nu_c}\right]^2} \tag{6.16}$$

For purely thermal motions, $b^2 = {b_k}^2 = 2kT/m$. But Gaussian-distributed turbulent motion with a velocity dispersion b_T can also be accomodated, by using

$$b^2 = {b_k}^2 + {b_T}^2 \tag{6.17}$$

instead.

The combined effects of natural line broadening, pressure broadening and Doppler broadening are described by convolving the two profiles to yield the profile function

$$\phi(\nu) = \int L(\nu - \nu')\, D(\nu')\, d\nu' \tag{6.18}$$

which, with a slightly different normalization, is called the Voigt profile. As mentioned earlier in equation (6.1), the observed spectrum will itself be the convolution of the intrinsic line profile with the instrumental profile.

Near the center of the line, where the Doppler profile is much wider than the damping profile, the Doppler core dominates with the term obtained by expressing equations (6.13) and (6.16) in wavelength units:

$$\tau_\nu = \sqrt{\pi}\, r_e\, N f_{jk}\, \frac{c\lambda_c}{b}\, e^{-\left[\frac{c(\lambda-\lambda_c)}{b\lambda_c}\right]^2} \tag{6.19}$$

If the column densities are small, the lines remain relatively narrow and the area covered by the line increases proportionally to the column density; the line is said to be on the linear part of the so-called curve of growth. But at sufficiently large column densities, the line center becomes optically thick, and the growth in line strength with column density slows considerably. Finally, at even higher column densities, the rate increases a bit because pressure broadening in the saturated line leads to the appearance of the so-called damping wings, where the Lorentz profile dominates. Neglecting the second term in the denominator of equation (6.15), the optical depth in the damping wings is

$$\tau_\nu \approx \frac{r_e}{4\pi}\, N f_{jk}\, \gamma\, \frac{{\lambda_c}^4}{c\,(\lambda - \lambda_c)^2} \tag{6.20}$$

where the approximation $\lambda^2 {\lambda_c}^2 \approx {\lambda_c}^4$ can be made for optical wavelengths.

Equivalent width

One last concept should be mentioned, that of the equivalent width W_λ. It is defined as

$$W_\lambda = \int \left[\frac{I_\lambda(0) - I_\lambda}{I_\lambda(0)}\right] d\lambda = \int \left[1 - e^{-\tau_\lambda}\right] d\lambda \tag{6.21}$$

where the specific intensity I_λ and optical depth τ_λ are now defined in wavelength terms and the integral is evaluated over the entire absorption line. If we visualize a rectangle in a normalized spectrum whose height is equal to 1 and whose width in wavelength units is equal to W_λ, it will have the same area as that covered in the spectrum by the line itself. Lower-resolution spectra, which cannot resolve lines sufficiently to fit the Voigt profile and thus determine N and b, measure W_λ instead. The equivalent width is independent of the instrumental resolution because it represents the fraction of energy removed from the spectrum by an absorption line; so it is a good measure for the line strength, regardless of the resolution. That is why it was used to determine the minimum detection limit in this spectrum (section 6.3).

6.3 Minimum Detection Limit

In order to distinguish small lines from noise, the significance of an absorption line was assessed according to its equivalent width W_λ. A $5\,\sigma$ detection limit was chosen, where the $1\,\sigma$ level was given by

$$\sigma_{\mathrm{W}} = \frac{\mathrm{FWHM}}{\langle\,\mathrm{S/N}\,\rangle} \tag{6.22}$$

(Caulet 1989). The FWHM is the instrumental resolution, and $\langle\,\mathrm{S/N}\,\rangle$ is the mean local signal-to-noise ratio, at the location where a given absorption line is thought to appear.

The FWHM values used were those assigned to each wavelength as described in section 6.1. The mean signal-to-noise ratio was determined by selecting featureless regions at regular intervals wherever possible throughout the spectrum; the standard deviation of the flux values within that region was taken as the local measure of the noise, while the mean signal-to-noise ratio there was simply the inverse standard deviation because the spectrum was already normalized. Finally, the sizes of the selected regions were plotted against the $\langle\,\mathrm{S/N}\,\rangle$ to ensure that there was no correlation between the size of the region chosen and the standard devation, as might have been present had the regions been too small.

This method was chosen rather than the simpler one of merely dividing the flux by the errors in the normalized frames. Because the errors calculated according to error propagation included offset errors such as that for the bias level or background subtraction, they were significantly higher—by a factor of about 1.5 on average—than the actual fluctuations in the spectrum's continuum.

One additional step was necessary to relate this minimum equivalent width to the column densities delivered by the line-fitting programs. If we define an equivalent width W_ν in frequency units,

$$W_\nu \;=\; \int \left[\frac{I_\nu(0) - I_\nu}{I_\nu(0)}\right] d\nu \;=\; \int \left[1 - e^{-\tau_\nu}\right] d\nu \tag{6.23}$$

then it is possible, in principle, to insert equations (6.13) and (6.18) for τ_ν and try to evaluate this integral for W_ν. In practice, however, it was much easier to use an approximation (Spitzer 1978) which takes advantage of the fact that the weak lines being considered here have optical depths $\tau \ll 1$. In this case, the equivalent width in equation (6.21) becomes

$$W_\lambda \;=\; \frac{{\lambda_o}^2}{c} \int \left[1 - e^{-\tau_\nu}\right] d\nu \;\approx\; \frac{{\lambda_o}^2}{c} \int \tau_\nu \, d\nu \tag{6.24}$$

where the middle term in brackets could be expanded and approximated by $e^{-\tau_\nu} \approx 1 - \tau_\nu$ because $\tau_\nu \ll 1$. Along with equations (6.13) and (6.14), this yields

$$W_\lambda \approx \pi r_e \, N f_{jk} \, {\lambda_o}^2 \tag{6.25}$$

Therefore the column density and the equivalent width of optically thin lines are related by

$$N \approx \frac{1}{\pi r_e f_{jk}} \frac{W_\lambda}{\lambda_o{}^2} = 1.130 \times 10^{20} \text{ Å cm}^{-2} \frac{W_\lambda}{\lambda_o{}^2 f_{jk}} \tag{6.26}$$

$$N_{\rm HI} \approx 1.838 \times 10^{14} \text{ cm}^{-2} \text{ Å}^{-1} W_\lambda \tag{6.27}$$

where the oscillator strength $f_{jk} = 0.416$ and rest wavelength $\lambda_o = 1215.6701$ Å used in equation (6.27) are specifically for Ly α lines (Verner et al. 1994).

Since equations (6.26) and (6.27) are valid in the rest frame of the absorption line, care must be taken to correct for the effects of cosmological expansion. Just as with the observed wavelength, the observed equivalent width is a factor of $(1 + z)$ larger than the rest-frame equivalent width (Weedman 1986). Thus the only features in the spectrum toward HE 0940 that were identified as Ly α lines were those at the $5\,\sigma$ significance level, which equation (6.27) says are those with column densities

$$N_{\rm HI} \geq \frac{1.838 \times 10^{20} \text{ cm}^{-2} \text{ Å}^{-1}}{1+z} \frac{5 \times \text{FWHM}}{\langle\, \text{S/N}\,\rangle} \tag{6.28}$$

The minimum column density for Ly α lines in the spectrum as a whole ranged from log N_{min} = 12.6 at 3800Å to log N_{min} = 12.1 at 5000 Å. In the pure Ly α forest, the detection limit went from log N_{min} = 12.5 at the short-wavelength end down to log N_{min} = 12.2 at the long-wavelength end.

6.4 Line-Fitting Programs

Two fitting programs were used to fit Voigt profiles to the absorption lines in the normalized spectrum of HE 0940. The first was a version of Xvoigt, copyright 1994, written by David Mar at the University of Sydney. Using a text file containing the normalized spectrum as input, Xvoigt asks for the central wavelength (in Å), the size (in km s^{-1}) of the spectral region to be displayed, and the instrumental b value (in km s^{-1}) for that spectral region. It then plots this region on the screen, along with an arbitrary overlayed fit and its current N and b values. These values can be changed by moving the appropriate scroll bars with the mouse. Thus the fit is adjusted interactively until the user obtains what is judged (by eye) to be the best fit. The residuals of the fit can also be plotted to assist the user in making this decision.

Xvoigt allows the fitting of absorption lines by a superposition of more than one Voigt profile. Selecting an on-screen button and typing the central wavelength of the new component makes Xvoigt calculate and display the composite profile. Different atomic transitions of the same ion can also be fit simultaneously and viewed by toggling the check marks in the appropriately labeled boxes.

Although fitting absorption lines manually using Xvoigt provided a good "feel" for how N and b values affect the line profile, the main fitting program used in the analysis was FITLYMAN, a package in the Lyman context of MIDAS whose code was developed at the Osservatorio Astronomico di Roma. Like Xvoigt, FITLYMAN fits Voigt profiles to absorption lines in a normalized spectrum; but rather than relying on user judgments, the fit is obtained through the numerical minimization of χ^2 (Fontana and Ballester 1995).

The line parameters used by FITLYMAN are the observed central wavelength λ_c, the column density N and the Doppler broadening b (which actually consists of two parameters, b_k and b_T; see equation 6.17). The measured flux in the spectrum is written in FITLYMAN as

$$I_\lambda = I_\lambda(0)\, e^{-\tau_\lambda} \tag{6.29}$$

$$\text{with} \quad \tau_\lambda = \sqrt{\pi}\, r_e\, N f_{jk}\, \frac{c\lambda_o}{b\sqrt{2\ln 2}}\, H(a,u) \tag{6.30}$$

where $H(a,u)$ is the equivalent of the line profile in equation (6.18), but expressed here in terms of wavelengths and with a different normalization.

$$H(a,u) = \int_{-\infty}^{\infty} \frac{e^{-y^2}}{a^2+(u-y)^2}\, dy \tag{6.31}$$

$$a = \frac{\gamma\lambda_c}{4\pi b} \tag{6.32}$$

$$u = \frac{c\,(\lambda_c-\lambda)}{b\lambda_c\sqrt{2\ln 2}} \tag{6.33}$$

The atomic parameters needed (rest wavelength λ_o, oscillator strength f_{jk} and damping constant γ) are taken from the table atompar.tbl in the same directory. The function is fit to the data in such a way as to minimize the chi-square, which is defined for a set of binned data in j histogram bins as

$$\chi^2 \equiv \sum_{i=1}^{j} \frac{[\,h(x_i)-nP(x_i)\,]^2}{\sigma_i(h)^2} = \sum_{i=1}^{j} \frac{[\,h(x_i)-nP(x_i)\,]^2}{nP(x_i)} \tag{6.34}$$

where n is the total number of measurements, $h(x_i)$ is the frequency or observed number of counts in bin i, $P(x_i)$ is the probability of observing the value x_i in a random measurement, $nP(x_i)$ is the expected number of measurements in bin i, and $\sigma_i(h)$ is the standard deviation associated with the uncertainty $h(x_i)$ in the observed counts. Since fluctuations in the observed counts result from the statistical probabilities of taking random samples of a finite number of items, they follow a Poisson distribution with $nP(x_i)$ as mean (Bevington and Robinson 1992). And since the variance in a Poisson distribution is equal to the mean, $\sigma_i(h)^2$ is equal to $nP(x_i)$. Thus the numerator in equation (6.34) is simply the square of the difference between the expected and observed counts in each bin, while the denominator is the number of expected counts.

The actual minimization is performed for FITLYMAN by Minuit, a Fortran 77 package that was developed at the European Laboratory for Particle Physics (CERN). Minuit finds the minimum value of a user-defined multi-parameter function and analyzes the shape of the function around that minimum. The parameters in the function can be kept constant, allowed to vary freely, or allowed to take on any value within certain limits. Parameters that are variable can also be held fixed and later released at any given time during the minimization.

The Minuit version used by FITLYMAN can handle functions with up to 100 parameters, of which not more than 50 can be variable at one time (James 1994). FITLYMAN, in turn, currently allows for the simultaneous fit of up to 50 independent parameters which can be either fixed at a constant value or allowed to vary freely during the entire fitting process. Perhaps the most important feature here is that the user can specify the relations of the four line parameters between any lines in a given fit. For example, lines can be forced to share the same redshift z, so that there is only one independent parameter for their wavelengths. Or, if transitions from the same ion are defined as a multiplet group in FITLYMAN, they will be constrained to have the same z, N and b_k values; with the default b_T of zero for all lines, this results in only four independent parameters for the entire group. Another option is to require a common temperature, which forces the b_k values of the lines to scale with the square root of the ion masses (which are also stored in atompar.tbl). In general, any of the parameters can be either varied freely, assigned to the same variable, or kept fixed at a constant value. The user can simultaneously fit up to 16 lines or more, depending on the number of indendent parameters.

In addition to these user-specified conditions and the initial values for all of the independent variables, FITLYMAN needs to know the Minuit method and the boundaries of the wavelength intervals to be used in the χ^2 calculations. The information is usually entered via the menu-driven interface, but it can also be provided by tables having the same format as the output tables produced by FITLYMAN (these have the suffixes LYPAR.tbl, LYMIN.tbl and LYLIM.tbl).

6.5 Fit Results

Since there was often no continuum between two neighboring absorption lines, the spectrum was divided into 113 fit regions. Each such absorption region contained as many continuum points on each end, and as few absorption lines, as possible.

Every line or group of lines was first fit by hand using Xvoigt. The results were then used as the initial values in FITLYMAN. The output of the χ^2 minimization consisted of the line parameters λ, N and b, the error in each parameter, the reduced χ^2 of the fit, and the probability of χ^2 exceeding that value for the given number of degrees of freedom. If the result was not very good, additional components were added to the fit to improve the χ^2 value. Thus the probability for all of the final fits was greater than 96%. No additional Voigt profiles were added if they increased the χ^2 of the fit. In retrospect, however, this criterion alone was probably not enough to avoid over-fitting (see section 7.2).

In the end, a total of 382 lines were fitted within the Ly α forest between Ly β emission at $\lambda_{obs} = 4158.3$ Åand Ly α emission at $\lambda_{obs} = 4928.3$ Å, as were five Ly α lines just redward of that. Wherever possible, Ly β lines were used to constrain the much less reliable fits of saturated Ly α lines. There were also a number of metal lines present within the forest (see section 6.7), and all of them were fit simultaneously with any other transitions that could be found at the same redshift. The same was done for the two Ly α lines within the forest that had corresponding metal lines. Thus an additional 60 lines outside of the primary region of interest were included in the fits. The synthetic spectrum resulting from the line profiles of all of the fits in the Ly α forest can be seen in Appendix B.

Appendix C lists the line parameters for all of the fits according to wavelength. The errors quoted there were obtained by FITLYMAN using the Minuit command `HESSE`, which calculates them by inverting the covariance matrix. This method was used for convenience: it is the default in FITLYMAN, and the resulting error intervals are symmetric. Although the Minuit command `MINOS` has the advantage of taking both parameter correlations and non-linearities into account, it requires more computing time and yields errors of the form $x^{+\sigma_1}_{-\sigma_2}$.

6.6 Sources of Error

Fontana and Ballester (1995) describe the extensive tests performed on simulated data to assess the stability and accuracy of the results produced by FITLYMAN. The fits achieved high accuracy for unsaturated lines, but as the lines became moderately saturated the fit values showed a larger dispersion as well as the typical anti-correlation between N and b values, so that a line with known parameters might be fit using a larger N and a smaller b or vice versa. Once damping wings began to appear, however, the fits became much more precise. They worsened again considerably after the lines became so strongly saturated that the b value was practically undetermined. Though lower signal-to-noise ratios seemed to increase the possible spread of fitted values in all of these cases, no severe systematic effects could be found.

The problems in fitting saturated lines in particular were confirmed within the present spectrum. Not only were the errors for each of the line parameters much greater, but the χ^2 limit was not always useful in distinguishing between two fits that looked very similar but had very different parameters.

Fontana and Ballester mention two strategies that can improve the accuracy of the fits. The first is to take advantage of the presence of several transitions in a single system wherever possible, because their tests showed that such simultaneous fits significantly increased the precision. That is why, in the spectrum of HE 0940, Ly β lines were used to constrain Ly α fits, and why all of the transitions in the metal systems were invariably fit together as a group.

The second strategy involves the proper choice of the fit interval in which to compute χ^2. The FITLYMAN simulations showed severe systematic effects, particularly for saturated lines, when the limits of the interval were chosen inside the line. The fit intervals in this spectrum were therefore made as wide as possible in all cases, even though the high line density often made it impossible to include many continuum points on the edges of the fits.

The initial fitting with Xvoigt constituted the third strategy that was adopted in an attempt to achieve the best fit results under these conditions. Fontana and Ballester claim that the FITLYMAN algorithm is rather insensitive to the initial values but do admit that in more complex situations involving blends, "it may be necessary to provide reasonable initial values to obtain the correct answer." A few cursory fits of saturated lines in the current spectrum confirmed that different initial values could lead to different, sometimes unphysical, fit results for the same region. The hope was that a reasonable first fit with Xvoigt helped reduce the choice between local χ^2 minima in different regions of parameter space. But is not clear whether such subjective initial guesses really led to better fits.

Even more serious than the errors resulting from uncertainties in the fits themselves is the problem of line blending. As explained by Kirkman and Tytler (1997), line blending occurs when the instrumental resolution is too low to resolve two (or more) lines. They cite the example of several discrete clouds, tens of kiloparsecs apart and each with a small b value. In a spectrum of a certain resolution, these clouds would appear as a single Voigt profile and be taken for a single localized cloud with a high b value. For these reasons, it was often unclear how many components should be used to fit an absorption line in the spectrum of HE 0940. There was a clear trend in which the χ^2 decreased, but the errors in the individual parameters increased, as more components were added to a fit.

Finally, it should be mentioned that the line parameters obtained in the Voigt-profile fitting do not directly describe the actual physical situation of the absorbing gas. The assumptions implicit in the line-blending example cited above, and in the Voigt-profile decomposition in general, are probably not completely valid. As Gardner et al. (1997) have pointed out, recent theoretical work using gas dynamics, radiative cooling and photoionization have undermined the physical picture of discrete Ly α "clouds," each characterized by a single velocity dispersion b. The current theoretical models show absorbing systems that "merge continuously into a smoothly fluctuating background, often contain gas at a range of temperatures, and are usually broadened in frequency space by coherent velocity flows that do not resemble Gaussian turbulence." Nevertheless, these authors add that any spectrum can be described phenomenologically by a superposition of Voigt-profile lines. The key question at that point becomes how the Voigt-profile components and line parameters are related to the actual physical state of the absorbing gas. Theoretically, this can then be deduced by processing simulated spectra in the same way as observed spectra like the one at hand.

6.7 Heavy-Element Systems

Only a few heavy-element systems were expected to be found among the hundreds of lines in the Ly α forest. Nevertheless, an attempt was made to identify as many metal lines as possible in order to reduce the contamination of the Ly α sample.

The standard method for locating heavy-element systems is to search all available wavelengths for significant excesses of identifications at particular redshifts. A simple way to do this would be to apply an automated line-finding and -fitting program to the entire spectrum and then to look for peaks in the histogram of the redshift values. But in this case it was not judged worth the effort to fit the entire spectrum blueward of Ly β emission using the very time-consuming methods described in section 6.4.

So, in the absence of automated line-fitting programs, the search for metal lines was done by means of a simple but effective interactive MIDAS procedure written for this purpose. It displays the desired spectral region in the top half of the graphics window, then asks the user to click on an absorption line with the mouse and enter its assumed identification. When the name of a second transition is entered, the relevant spectral region is plotted in the bottom half of the graphics window along with a vertical line indicating the position at which the desired line would be expected to appear. The originally chosen location is also marked with a vertical line so that the user can judge the positions more closely, optionally select a new location more in the center or more on the wings of the reference line, and see how this would change the position for the expected transition. A mini-menu allows quick navigation between cursor selection, searches for other transitions at the same redshift, and plotting of a new spectral region. No information on relative line strengths was used in the initial search, but this procedure made it possible to check closely the positions of candidate lines and thus rule out most possibilities.

Given a Ly α line with the observed wavelength λ_{obs} in a spectrum covering the wavelength range $[\lambda_{min}, \lambda_{max}]$. The corresponding line of another transition with rest wavelength λ_m will fall within range of the spectrum only if

$$\lambda_{min} \;\leq\; \frac{\lambda_{obs}}{1215.67\ \text{Å}} \;\leq\; \lambda_{max} \tag{6.35}$$

Alternatively, the Ly α of a given metal line will appear in the spectrum only if

$$\frac{\lambda_{min}}{\lambda_m}\, 1215.67\ \text{Å} \;\leq \lambda_{obs} \leq \frac{\lambda_{max}}{\lambda_m}\, 1215.67\ \text{Å} \tag{6.36}$$

Unfortunately, the maximum of the spectral coverage at 5240 Å meant that no Ly α in the entire spectrum could be observed together with its possible Al II, Al III, Fe II or

Mg II lines, which are among the stronger and hence more commonly seen metal lines. None of the Ly α lines in the primary region of the Ly α forest could have C IV lines present in the spectrum either, although C IV lines were within range of the spectral coverage for Ly α lines with $\lambda_{obs} \leq 4142$ Å.

For the purposes of this study, metal systems were identified using the following criteria:

1. If the Ly α was in range of the spectrum, it had to be observed.
2. If the Ly α was not in range, at least one unblended doublet had to be present.
3. In either case, a simultaneous FITLYMAN fit of all transitions in the system had to yield a reasonable fit, defined as one with probability $P > 96\%$ and not too many improbable blends.

Although multiple lines of the same ion were constrained to share the same redshift, column density and b value, no other conditions were set on the metallicities or ratios of ionization species.

The general strategy in the search for metal lines was to begin with the isolated lines redward of Ly α emission, in this case the two obvious groups of C IV doublets at ~5160 Å and ~4990 Å and the small Si IV doublet at 5112, 5145 Å. Using the list of strong lines expected in quasar absorption systems by Verner et al. (1994), further members of these systems were sought using the interactive MIDAS procedure described above.

The next step was to examine all of the strongest saturated lines, both inside and outside of the Ly α forest, on the assumption that these were the most likely to show corresponding metal lines. First the C IV $\lambda\lambda 1548, 1550$ doublet was sought, if it was in range; then the Si IV $\lambda\lambda 1393, 1402$ doublet and the Si III $\lambda 1206$ and C II $\lambda 1334$ lines. These were the only metal ions found in the end, although others such as Si II and N V were also checked.

A total of five heavy-element systems were identified in this way. An additional C IV system was noticed at $z \approx 2.83$ because of the large complex of strongly saturated lines between 4644 Å and 4664 Å. The corresponding C IV absorption can be seen from about 5920 Å to about 5950 Å in the 1.52-meter spectrum of HE 0940 (Fig. 2.1). Since this feature is not resolved, however, no further information could be obtained on this system.[1]

Other extensive but more random searches for metal systems produced no additional identifications according to the criteria listed above. The absorption lines that could be identified as belonging to possible metal systems on the basis of this spectrum alone are summarized in Table 6.1.

[1]Since then, a higher-resolution spectrum than the one in Figure 2.1 has been obtained at the ESO 1.52-meter telescope. It extends up to 7500 Åand shows more details of the complex C IV system at $z \sim 2.82$. Additional metal systems have thus been identified: there is a strong, probably damped Ly α system at $z = 1.9185$ with Fe II, Si II, Al III and other metal lines, as well an Mg II system at $z = 1.06$ (D. Reimers, private communication).

ID	z	λ_{obs} (Å)	σ_λ	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b
H I 1215	2.174456	3859.09	0.02	18.30	0.06	29.7	1.0
C II 1334	2.174456	4236.41	0.02	13.21	0.11	7.9	5.7
Si IV 1393	2.174456	4424.41	0.02	13.34	0.04	22.8	3.0
Si IV 1402	2.174456	4453.03	0.02	13.34	0.04	22.8	3.0
C IV 1548	2.174587	4914.88	0.01	14.25	0.02	23.6	0.9
C IV 1550	2.174587	4923.06	0.01	14.25	0.02	23.6	0.9
H I 1215	2.220397	3914.94	0.05	14.38	0.14	54.6	7.1
C IV 1548	2.220396	4985.80	0.06	13.44	0.05	41.6	5.1
C IV 1550	2.220397	4994.09	0.06	13.44	0.05	41.6	5.1
Si III 1206	2.220513	3885.55	0.01	12.64	0.12	29.2	12.0
H I 1215	2.220513	3915.08	0.01	16.13	5.51	17.3	17.6
Si IV 1393	2.220513	4488.61	0.01	12.55	0.10	9.3	6.3
Si IV 1402	2.220513	4517.64	0.01	12.55	0.10	9.3	6.3
C IV 1548	2.220513	4985.98	0.01	13.74	0.03	8.9	1.0
C IV 1550	2.220513	4994.28	0.01	13.74	0.03	8.9	1.0
Si III 1206	2.221298	3886.50	0.02	13.13	0.15	27.9	5.0
H I 1215	2.221298	3916.04	0.02	13.86	0.34	54.2	19.3
Si IV 1393	2.221298	4489.70	0.02	12.99	0.06	29.8	5.6
Si IV 1402	2.221298	4518.74	0.02	12.99	0.06	29.8	5.6
C IV 1548	2.221298	4987.20	0.02	13.57	0.04	14.8	2.9
C IV 1550	2.221298	4995.49	0.02	13.57	0.04	14.8	2.9
Si III 1206	2.221561	3886.81	0.02	12.67	0.58	38.2	25.4
H I 1215	2.221561	3916.36	0.02	13.88	0.22	21.9	10.1
C IV 1548	2.221561	4987.60	0.02	13.21	0.13	5.5	1.3
C IV 1550	2.221561	4995.90	0.02	13.21	0.13	5.5	1.3

Table 6.1: *Heavy-element systems in the spectrum toward HE 0940–1050.*

ID	z	λ_{obs} (Å)	σ_λ	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b
Si IV 1393	2.329162	4640.04	0.01	13.71	0.04	23.3	1.4
Si IV 1402	2.329162	4670.05	0.01	13.71	0.04	23.3	1.4
C IV 1548	2.329162	5154.19	0.01	14.75	0.13	19.3	1.5
C IV 1550	2.329162	5162.76	0.01	14.75	0.13	19.3	1.5
Si IV 1393	2.330278	4641.59	0.12	13.17	0.16	25.4	5.6
Si IV 1402	2.330278	4671.61	0.12	13.17	0.16	25.4	5.6
C IV 1548	2.330278	5155.92	0.13	14.19	0.12	35.9	5.0
C IV 1550	2.330279	5164.50	0.13	14.19	0.12	35.9	5.0
Si IV 1393	2.330575	4642.01	0.03	13.36	0.11	14.3	3.2
Si IV 1402	2.330575	4672.03	0.03	13.36	0.11	14.3	3.2
C IV 1548	2.330575	5156.38	0.04	14.03	0.16	13.7	4.3
C IV 1550	2.330575	5164.96	0.04	14.03	0.16	13.7	4.3
Si IV 1393	2.331016	4642.62	0.06	12.50	0.15	12.8	6.6
Si IV 1402	2.331016	4672.65	0.06	12.50	0.15	12.8	6.6
C IV 1548	2.331016	5157.06	0.07	13.93	0.07	30.3	3.9
C IV 1550	2.331016	5165.64	0.07	13.93	0.07	30.3	3.9
H I 1025	2.667540	3761.88	0.04	15.21	0.12	63.1	5.0
H I 1215	2.667541	4458.52	0.04	15.21	0.12	63.1	5.0
Si III 1206	2.667541	4424.89	0.05	12.94	0.05	24.5	3.3
Si IV 1393	2.667541	5111.65	0.05	12.80	0.06	21.9	4.3
Si IV 1402	2.667541	5144.72	0.05	12.80	0.06	21.9	4.3
H I 1025	2.916881	4017.63	0.01	16.23	0.18	48.9	2.3
H I 1215	2.916881	4761.64	0.01	16.25	0.18	48.9	2.3
Si III 1206	2.916881	4725.72	0.01	13.07	0.06	14.3	1.5

Table 6.1: *Heavy-element systems in the spectrum toward HE 0940–1050, cont.*

Chapter 7

The Lyman-alpha Forest

Because the line of sight to high-redshift quasars covers a large range in look-back times and several orders of magnitude in the observed column densities, the properties of Ly α forest lines have been studied to see how they change with z and $N_{\rm HI}$ in the hopes of learning more about the distribution of matter and the evolution of structure in the universe. The b values have also been studied because they result directly from the particle motions and physical properties of the absorbing gas. The assumed form of these distributions will be introduced below, before turning to the data themselves.

Given a population of objects with a comoving number density n_o, so that the number density changes with redshift as $n = n_o(1+z)^3$. The number of objects $d\mathcal{N}$ that will be intersected by a line of sight at redshift z is

$$d\mathcal{N} = \sigma(z)\, n(z)\, dl \tag{7.1}$$

where σ is the geometric cross section of each object and dl is the path length in terms of proper displacement, the expression for which depends on the cosmological model (Peebles 1993). Assuming negligible space curvature in a Friedmann cosmology (with density paramer $\Omega_o > 0$ and a zero-valued cosmological constant Λ_o), the number of objects per unit redshift is then

$$\frac{d\mathcal{N}}{dz} = \frac{cn_o\,\sigma(z)}{H_o}\frac{1+z}{\sqrt{1+\Omega_o z}} \tag{7.2}$$

(Rauch 1998). Therefore, if there is no evolution in the number density n_o or cross section σ with redshift, we would expect the number of objects per unit redshift to be

$$\frac{d\mathcal{N}}{dz} \propto \begin{cases} 1+z & \text{for } \Omega_o = 0 \\ (1+z)^{1/2} & \text{for } \Omega_o = 1 \end{cases} \tag{7.3}$$

Traditionally, the redshift distribution of Ly α clouds within a given column-density range has been parametrized by investigators as

$$\frac{d\mathcal{N}}{dz} = A_o(1+z)^{\gamma}, \tag{7.4}$$

where A_o is a constant. Despite the large spread in the values for γ obtained from different samples of quasar spectra, observations show that the number of absorbers clearly increases more rapidly with redshift than expected for a non-evolving population, which would have $0.5 \leq \gamma \leq 1$.

Similarly, the column-density distribution of Ly α clouds in a given redshift range has been conventionally parametrized by a power law,

$$\frac{d\mathcal{N}}{dN_{\mathrm{HI}}} \propto N_{\mathrm{HI}}^{\;-\beta} \tag{7.5}$$

For this form of the equation, the parameter β is taken to be positive to reflect the fact that absorbers with higher column densities are increasingly rare.

7.1 Theoretical models

The usual procedure for determining the values of the parameters in equations (7.4) and (7.5) has been to perform power-law fits to data samples derived from many quasar spectra. But given the poor statistics provided by a single quasar spectrum, no attempt has been made here to do a similar determination from the Ly α forest of HE 0940 alone. Instead, it was thought better to compare the present data to theoretical models constructed from a large observational sample.

Combining equations (7.4) and (7.5) into a single distribution in both column density and redshift yields

$$\frac{\partial^2\mathcal{N}}{\partial z\,\partial N_{\mathrm{HI}}} = \mathcal{A}_o\, N_{\mathrm{HI}}^{\;-\beta}\,(1+z)^{\gamma} \tag{7.6}$$

where $\mathcal{A}_o$ is again a constant. Fardal, Giroux and Shull (1998) have recently produced global fits to this distribution for Ly α absorbers. Taking advantage of the better quality, greater redshift coverage and increased overlap in the column-density ranges between the different spectra that have become available in the past few years, the authors have combined the line lists from 7 samples of Ly α forest data with those of 2 studies of Lyman-limit systems. They developed different models for the absorption-line distribution, using maximum-likelihood methods to evaluate three-power-law fits to the data. They also show that the best single-power-law fit describes the data poorly, due to a deficit of clouds with $N_{\mathrm{HI}} \sim 10^{15}$ cm^{-2}.

Since one of the purposes of the paper by Fardal, Giroux and Shull was to explore the He II Gunn-Peterson effect, their models A1 and A2 were designed to respectively minimize and maximize the He II line and continuum opacity relative to that of hydrogen. For our purposes, however, it suffices to note that the authors assumed N_{HI} and z to be statistically independent and required that the column-density distributions be continuous. A slight exception was made for their model A3, which is the same as model A2 up to $N_{\mathrm{HI}} = 10^{17}$ cm^{-2} but uses the redshift evolution of Stengler-Larrea et al. (1995) beyond that point. This was done because Lyman-limit systems may evolve with redshift at a different rate than the Ly α forest clouds.

The distributions resulting from models A1 to A3 of Fardal, Giroux and Shull will form the basis of all further comparisons to the present data. Table 7.1 summarizes the parameter values for each of these models, and Figure 7.3 shows the shape of their distributions for the redshift range under consideration.

Model	model region	N_{min} (cm^{-2})	N_{max} (cm^{-2})	$\mathcal{A}_o$	β	γ
	1	...	1×10^{14}	9.15×10^{5}	1.40	
A1	2	1×10^{14}	1×10^{16}	2.52×10^{12}	1.86	2.58
	3	1×10^{16}	1×10^{19}	2.10×10^{2}	1.23	
	4	1×10^{19}	1×10^{22}	4.42×10^{1}	1.16	1.30
	1	...	2×10^{14}	1.43×10^{7}	1.49	
A2	2	2×10^{14}	2×10^{15}	5.41×10^{12}	1.88	2.58
	3	2×10^{15}	1×10^{19}	2.88×10^{6}	1.47	
	4	1×10^{19}	1×10^{22}	4.42×10^{1}	1.16	1.30
	1	...	2×10^{14}	1.43×10^{7}	1.49	
A3	2	2×10^{14}	2×10^{15}	5.41×10^{12}	1.88	2.58
	3	2×10^{15}	1×10^{17}	2.88×10^{6}	1.47	
	4	1×10^{17}	1×10^{22}	5.00×10^{7}	1.50	1.50

Table 7.1: *Parameter values in models A1 to A3 of Fardal, Giroux & Shull (1998).*

7.2 Doppler Parameter and Column Density

The Doppler parameters of Ly α absorption lines are studied in the hopes of deducing information about the temperature and motions of the gas in the clouds. However, as mentioned in section 6.6, none of the simplest assumptions seem to apply. Since the relation between the b value and the temperature is complicated and not yet understood, a discussion about the physical properties of the absorbing gas goes far beyond the scope of this work. Nevertheless, a part of the discussion about low b values is relevant to the data here.

Generally speaking, the new high-resolution studies with the Keck 10-meter telescope appear to be establishing that the median Doppler parameters for Ly α lines at $z \sim 3$ are roughly 30 km s^{-1} (e.g., Kim et al. 1997). The distribution can apparently be described as a Gaussian truncated below a cut-off in b to reflect the apparent dearth of lines below $b \sim 20$ (Hu et al. 1995, Kirkman & Tytler 1997).

Figure 7.1 shows a plot of the Doppler parameter vs. column density for all of the lines between 4928.3 Å and 4944 Åin the spectrum toward HE 0940-1050. Of the 19 identified metal lines, only 11 are evident because 8 of them are doublets and thus have the same N_{HI} and b values. The Ly α lines are represented as crosses, and filled squares show the members of identified metal systems. Both symbols together indicate the 5 Ly α lines that are assumed to belong to metal systems, although only the 2 lines with $N_{\text{HI}} = 15.2$ and 16.2 could actually be fitted along with their corresponding metal lines. For the purposes of this graph only, it was assumed that in the heavily blended region from 4644 Å to 4664 Å, only the three Ly α lines with the highest column densities were associated with the C IV absorption seen in the 1.52-meter spectrum.

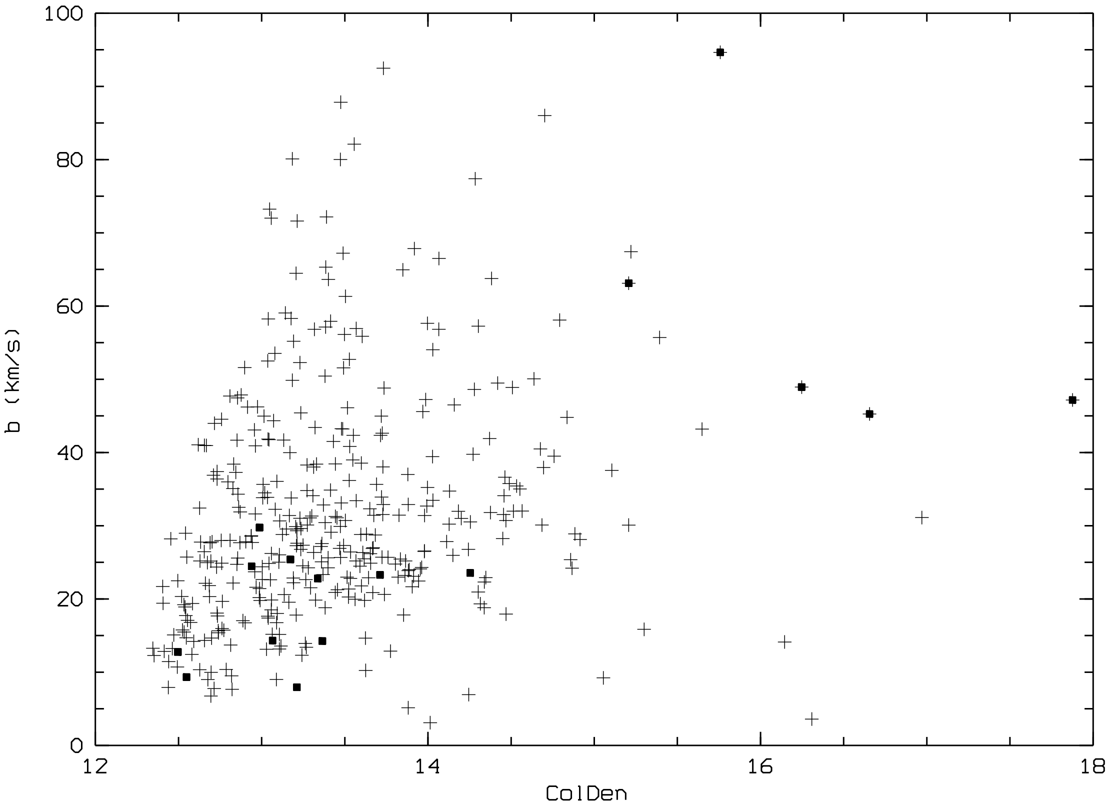


Figure 7.1: *b values vs. column density N_{HI} for all lines between Ly α and Ly β emission.*

Although there is no obvious minimum b cut-off in figure 7.1, this is probably—as is most often the case—due to unidentified metal lines. However, unless or until more metal systems can be identified, it is worth considering whether at least part of the low b values are due to problems with noise.

Pettini et al. (1990) reported a correlation between the b values and the column densities, as well as a high precentage of low $b \lesssim 20$ values; however, both were later shown to be artifacts of the line-finding and fitting procedures. Rauch et al. (1992) performed simulations on echelle spectra with an average FWHM resolution of 23 km s^{-1} (not too different from that of the current spectrum at hand). After fitting simulated spectra with the same programs used on the actual spectrum, they observed the following effects:

1. Broad, shallow lines that went undetected by the line-finding programs.
2. Fits with excessively high b values arising from unresolved blends.
3. Fits with unrealistically low N and b values that were not present in the original sample; these were due to S/N effects.
4. A slight increase in other fitted b values due to blending.

The extremely high b values in the current spectrum are therefore most probably unresolved blends. In some cases, they may be due to the choice of continuum as well; broad shallow lines may be missed by line-finding programs because of noise, but they can also be "created" when the continuum fit is too high. This could be the case, for example, for the line at 4913 Å with $b = 111$.

The current data also seem to confirm that some of the fits of lines with lower column densities tend to have low b values. Among the assumed Ly α lines, 21% have a Doppler parameter of $b < 20$. It is instructive to note that when only lines with $\log N_{\rm HI} \gtrsim 13$ are considered, the percentage drops to around 10% and stays in that neighborhood for all higher cut-offs that were tried. This is shown, too, in the histogram of the b values in Figure 7.2. The form of the histogram remained basically the same for column densities above $\log N_{\rm HI} > 13$. This could be because unidentified metal lines will tend to have lower column densities than the average Ly α lines, so that many of them are removed by the cut in $N_{\rm HI}$. But there also seem to be systematic errors in the profile fitting of lines with smaller column densities. The latter argument would provide an additional reason for the introduction of the completeness limit of $\log N_{min} = 13.33$ discussed in section 7.3.

In addition to the expected errors in the b values of weaker lines, the spectrum was also checked for unrealistically low b values at higher column densities. The empirical relation from Kirkman and Tytler (1997)[1]

$$b_{min} = 14 + 4 \times \log \left[\frac{N_{\rm HI}}{10^{12.5}} \right] \text{ km s}^{-1} \tag{7.7}$$

was used only as a rough guide, since their data had a much higher resolution and signal-to-noise ratio. Nevertheless, when it was postulated that the b cut-off seen in Keck spectra among lines with $\log N_{\rm HI} > 12.5$ might only appear in the current spectrum among lines with $\log N_{\rm HI} \gtrsim 13.3$, only 12 offending lines were found below the empirical limit; and only 7 of them were beyond the range allowed by their errors.

The two lines at 4619.4 Å and 4751 Å may be unidentified Si III lines belonging to the system at $z \approx 2.83$ with unresolved C IV lines (see section 6.7). However, no acceptable fits of the heavily saturated Ly α region could be obtained under this assumption, so the original determinations were left unchanged. The interesting point

[1] Hu et al. (1997) obtain a similar, slightly lower straight-line boundary in their plot of b vs. $N_{\rm HI}$.

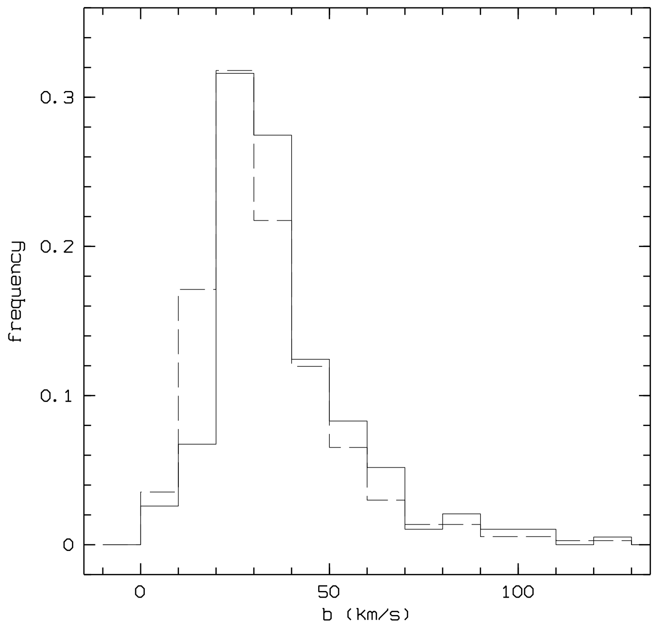


Figure 7.2: *Histogram of* b *values for the entire Ly* α *sample (dotted line) and for Ly* α *lines with* $\log N_{\mathrm{HI}} > 13.3$ *(solid line).*

here is that four of the five remaining lines—and six of the original 12 lines with suspiciously low b values—all come from the same region between 4317 Å and 4331 Å.

It is not clear whether the lines in this region could be part of a some complex metal system, or whether lines the low b values are a result of over-fitting. The standard procedure is to fit each line complex with the very fewest number of components needed to achieve a certain reduced χ^2 or probability of random deviation. Though this criterion was kept in mind while doing the fits with FITLYMAN, no specific limit had been fixed before starting the analysis; so it could be that the method was not followed strictly and that additional components were sometimes added to improve the χ^2 of the fits. Over-fitting could also have resulted from the fact that the initial fit was done by hand with Xvoigt. Because the number of fit components was determined by eye during that manual fitting process, maybe some of the initial fits had already contained too many components to begin with.

An automated line-fitting program using the smallest number of components per fit would have certainly yielded more consistent results. But it would also have been possible to mimic the steps by using FITLYMAN without Xvoigt, following a strict systematic approach in which the initial fit for every region began with a single component, regardless of how many lines it contained. Unfortunately, the fits had not been done using this more time-consuming method and it is not clear how much over-fitting there might have been or how much this would have affected the results. On the other hand, the region from 4317 Å to 4331 Å looked like perhaps it could be simply a noise problem rather than a fitting problem. In any case—whether noise, overfitting or metal lines—it was deemed safer to exclude this particular region from the final sample.

7.3 Column-Density Distribution

In addition to the line blending which affects the distribution of b values, intermediate-resolution spectra (and high-resolution spectra to a lesser extent) will also suffer from the problem of line blanketing. Line blanketing occurs when a weak line remains unnoticed because it lies within the absorption profile of a stronger line. This results in fewer recognized lines with low column densities, and leads to a fall-off on the lower end of the column-density distribution.

The correction for line blanketing—called the completeness correction by Hu et al. (1997)—cannot be done without simulations in which many synthetic spectra with the assumed b and $N_{\rm HI}$ distributions are fitted to see what percentage of lines at certain column densities will go undetected. Without the resources to do such extensive tests (though the process with FITLYMAN is described nicely by Giallongo et al. 1996), the somewhat unfortunate alternative is to simply determine a lower limit above which it is assumed that not many lines are lost to blanketing effects, and to consider only lines above this limit. The resulting sample is then said to be "complete."

The completeness limit for the data at hand was determined with the help of the models of Fardal, Giroux and Shull (Table 7.1). In order to compare actual line counts to the models, a concrete range $[z_{min}, z_{max}]$ and $[N_{min}, N_{max}]$ must be specified, so that equation (7.6) can be integrated to yield the number of absorbers $\mathcal{N}$ expected in that interval:

$$\mathcal{N} = \frac{\mathcal{A}_o}{(1-\beta)(1+\gamma)} \left[{z_{max}}^{1+\gamma} - {z_{min}}^{1+\gamma} \right] \left[{N_{max}}^{1-\beta} - {N_{min}}^{1-\beta} \right] \tag{7.8}$$

Figure 7.3 shows this comparison for column-density bins of $\Delta N_{\rm HI} = 0.33$, in which the different parameter values for the various power-law regimes had to be taken into account. Since the model parameters do not change with redshift, the entire spectrum between Ly α and Ly β emission could be used for this graph.

Even when the number counts in the data lie above those of the model predictions, the fall-off at $N_{\rm HI} \approx 10^{13}$ is clearly visible. It is also encouraging to see that this is roughly the same column density at which the b distribution assumes the more truncated shape expected from the Keck results. In order to avoid taking the lower limit of the third bin, in which the transition to significant line blanketing seems to occur, a slightly more conservative cut-off of $\log N_{min} = 13.33$ was chosen.

Figure 7.3 also illustrates why the models were scaled before being used as the basis for the Monte Carlo programs described in section 8.5. Especially when the lowest bins are ignored, the predicted counts per bin are clearly lower than the number of lines actually seen in the data. It was therefore decided to scale each model by the amount necessary to make the predictions match the number of actual counts in the column density range $13.33 \leq \log N_{\rm HI} \leq 16.0$. The upper limit was the one for which the Voigt-profile fits were judged to be reliable.

For the 115 lines observed within this column-density and redshift range, model A1 predicted 87.05 counts and models A2 and A3 predicted 84.57 counts; so in all of the calculations below, the parameter $\mathcal{A}_o$ has been increased by a multiplicative factor of 1.32 in model A1 and a factor of 1.36 in models A2 and A3. This step improved the fit between the models and the data significantly: the χ^2 for model A1 above $\log N_{\rm HI} \geq 13.33$ dropped from 21.7 to 11.1, from 19.2 to 7.8 for model A2, and from 18.0 to 7.0 for model A3. The scaled fits are shown in Figure 7.4.

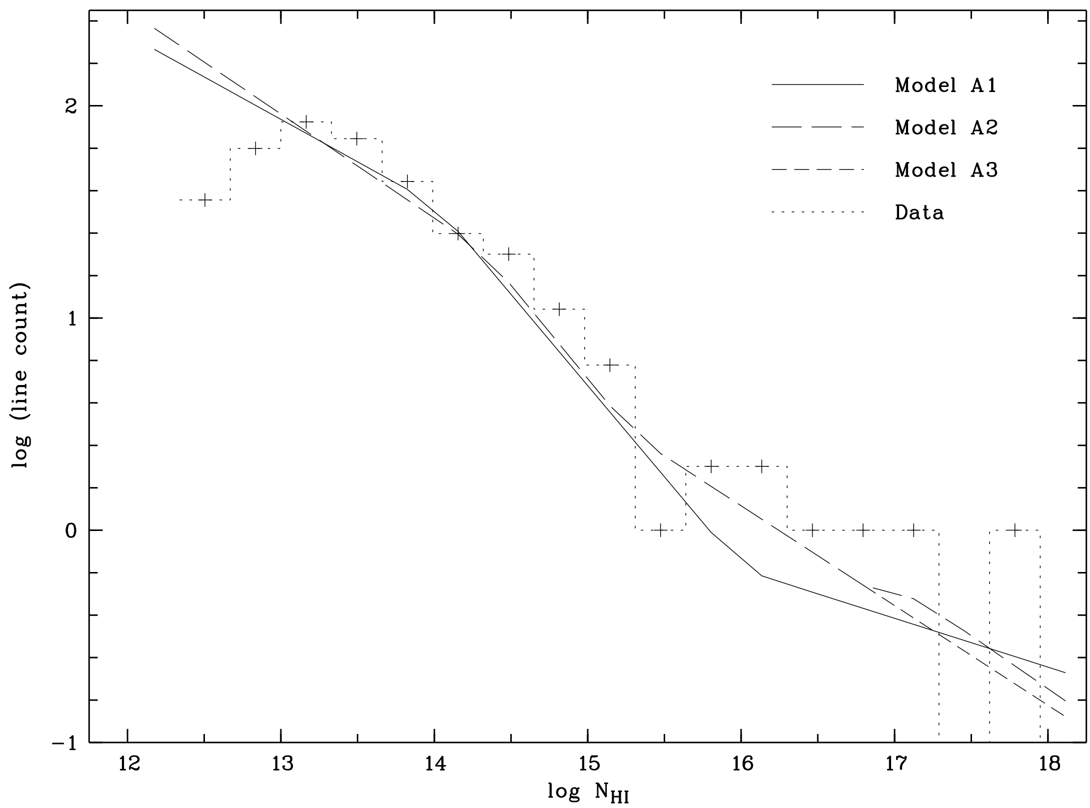


Figure 7.3: *Observed and predicted column-density distributions for Lyα lines, shown with bins of 0.33 in log $N_{\rm HI}$. Line counts are over the entire redshift interval* $2.420565 < z < 3.054$.

7.4 The Data Sample

To summarize the steps until now: 19 out of the 387 lines in Figure 7.1 had been identified as metal lines and therefore removed from consideration in all subsequent steps. The b and $N_{\rm HI}$ distributions of the remaining 368 Ly α lines were then examined and used to determine a completeness limit of $\log N_{min} = 13.33$, which is noticeably higher than the maximum detection limit of $\log N = 12.5$ for that part of the spectrum. Since lines with low column densities are much more numerous than stronger lines, this last cut-off reduced the available sample to 185 Ly α lines.

Two additional cuts were made at this point. The "noisy" region from 4317 Å to 4331 Å was excluded for reasons discussed in section 7.2. The highly saturated region described in section 6.7 was also excluded, primarily because the heavy line blanketing there made it difficult to determine the position or even the number of lines with certainty.[2] The resulting gaps of $\Delta z = 0.0115$ and $\Delta z = 0.0176$ in the z distribution were taken into account in every step of the analysis described below.

Finally, because the column densities of saturated lines show severe systematic effects in FITLYMAN when not enough continuum is present on both sides, the upper limit for the column density was chosen to be $\log N_{min} = 17.0$. But since the only line with a column density above that had already been excluded as part of the heavily saturated region, this was more of a formal limit. This cut-off did, however, have the advantage of reducing the number of theoretical Ly α forest models that had to be considered. Models A2 and A3 were identical in this range, but because model A3 incorporated the results from Lyman-limit systems and produced a better agreement to the data in section 7.3, these two equivalents will be referred to as model A3 in the discussions below.

It should be noted that in past studies, an effort was made to exclude the Ly α lines belonging to metal systems from the Ly α forest samples used to study the proximity effect. Because metal systems seemed to show different composition and clustering properties, they were assumed to belong to a different parent population of absorbers. However, recent studies with high resolution and signal-to-noise ratios have found associated C IV absorption for all of the high-redshift Ly α clouds with column densities $\log N_{\rm HI} > 15.0$ and for 60% of the clouds with column densities $14.5 < \log N_{\rm HI} < 15.0$ (Cowie et al. 1995, Tytler et al. 1995). Since the detection limit for metal absorption is still signal-to-noise-limited, it has even been argued (Lu et al. 1996 and references therein) that *all* Ly α clouds could conceivably contain metals. In any case, since the distinction seems to be weakening, the Ly α lines corresponding to metal systems identified in the spectrum of HE 0940have been kept in the Ly α sample. And since only two such lines are present, their influence on the results should be negligible anyway.

Now that the final sample of 169 Ly α lines had been selected, their redshift distributions were studied to see whether a proximity effect could indeed be detected in the spectrum of HE 0940.

[2]Slightly different limits of 4642 Å to 4663Å were used in the end, to match an existing bin boundary.

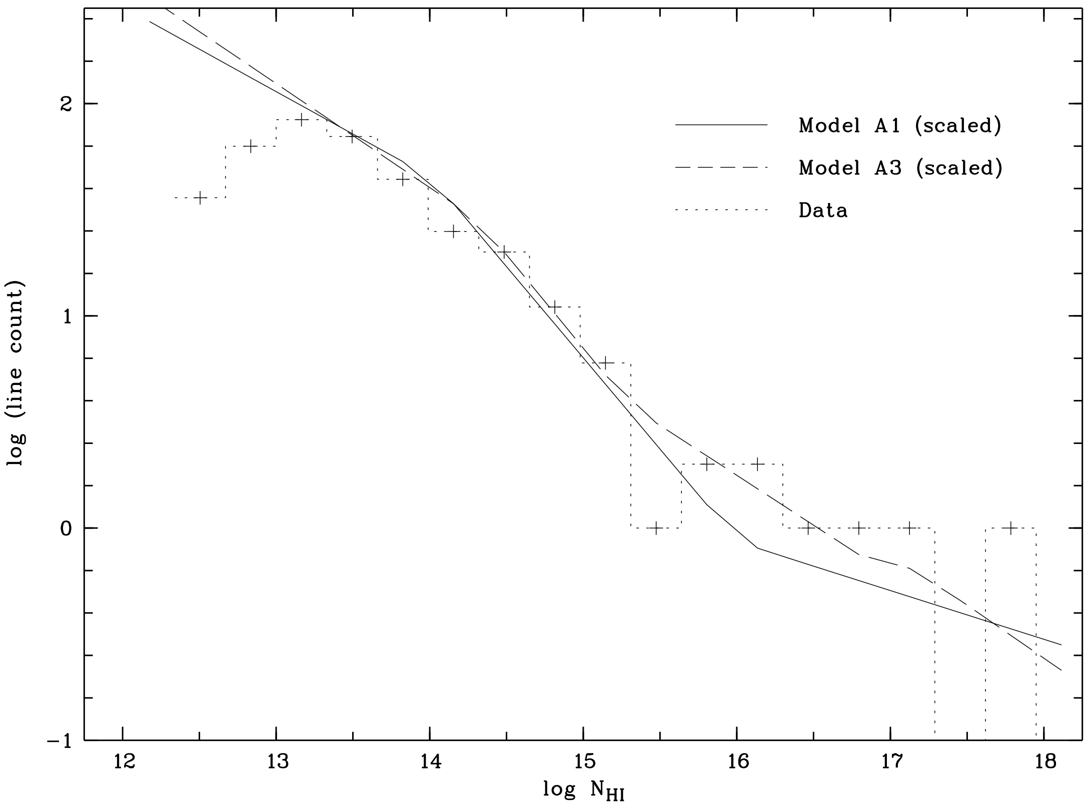


Figure 7.4: *Column-density distribution of models A1 and A3, scaled to the data over the column-density interval* $13.33 < log\ N_{\rm HI} < 16.0$ *and the redshift interval* $2.420565 < z < 3.054$*. These scaled models were the basis of the simulations in chapters 8 and 9.*

Chapter 8

The Proximity Effect

In studies of the redshift distribution of Ly α clouds described by equation (7.4), it was noticed early on (Carswell et al. 1982) that the γ values determined from individual QSOs were lower than those obtained by considering the line counts from several or many QSOs collectively. Murdoch et al. (1986) called this countervailing trend in the spectra of individual QSOs the "inverse effect" and showed that it resulted at least in part from including the region within roughly 4 Mpc of the quasar, which seemed to have a lower number density than regions in the spectrum that were further shortward of Ly α emission. They suggested that the effect was due to the radiation field of the quasar, which increased the degree of ionization of nearby hydrogen clouds and led to correspondingly lower column densities of neutral hydrogen.

Other, alternative explanations for the inverse effect were discussed (e.g., by Tytler 1987), but none found wide acceptance. For example, one suggestion was that the size of Ly α clouds in general means that clouds close to a quasar were too small to completely cover the quasar's emission-line region, so that the absorption falling within the Ly α emission line would show lower equivalent widths than similar clouds further from the quasar. But studies of common absorption lines in the images of gravitationally lensed quasars and in close QSO pairs indicate sizes on the scale of hundreds of kiloparsecs (e.g., Smette et al. 1995), while the equivalent width distribution of Ly α forest lines on the broad wings of the Ly α emission line appears to be statistically the same as that of the lines in the QSO continuum (Lu, Wolfe and Turnshek 1991).

8.1 The Photoionization Model

Bajtlik, Duncan and Ostriker (1988) were the first to present a quantitative physical model to explain the inverse effect, which they referred to as the "proximity effect." Providing a concrete model made it possible for the first time to use an independent method to infer the amount of photons coming from a metagalactic ultraviolet (UV) background at high redshifts.

First, the authors make the reasonable assumption that the Ly α clouds are highly ionized by the general background of UV photons, so that the neutral hydrogen density in a cloud will be inversely proportional to the radiation incident at frequencies greater than or equal to the Lyman limit.[1] Initially this means that in the sole presence of background radiation $J(\nu)$, which is normally given in units of erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}, the H I column density of a cloud would be

$$N_{\rm HI} = \frac{N}{\int\limits_{\nu_{LL}}^{\infty} 4\pi J(\nu)} \tag{8.1}$$

If a bright quasar is also present, with a flux density of $F_Q(\nu)$ erg cm^{-2} s^{-1} Hz^{-1}, then the cloud's column density would drop to

$$N_{\rm HI} = \frac{N}{\int\limits_{\nu_{LL}}^{\infty} [F_Q(\nu) + 4\pi J(\nu)]} \tag{8.2}$$

where N is some hypothetical column density in both cases.

What seems to be the key assumption made by Bajtlik, Duncan and Ostriker at this point is that the frequency dependence of the background radiation above the Lyman limit is the same as that of the QSO radiation. In that case, the ratio of ionization rates will be equal to the ratio

$$\omega = \frac{F_{\nu_{LL}}}{4\pi J_\nu(z)} \tag{8.3}$$

of the Lyman-limit flux densities, where $J_\nu(z)$ is the Lyman-limit background intensity and $F_{\nu_{LL}}$ is the Lyman-limit flux density from the quasar, both at the location of the cloud. The column density for a cloud in the vicinity of a bright quasar then becomes

$$N_{\rm HI} = \frac{N_o}{1+\omega}\,, \tag{8.4}$$

where N_o would have been the column density in the absence of the quasar.

8.2 Luminosity Distance and Fluxes

In order to apply the ionization model to the Monte Carlo simulations presented below, it was necessary to estimate the Lyman-limit flux density of the quasar radiation coming from HE 0940. The principal task of the next two sections will be to explain how this was done. A few definitions and elementary calculations will be introduced here before the actual determination is described in section 8.3.

In static Euclidean space, the observed bolometric flux f (energy per unit area and time) is a function of the source's bolometric luminosity L (energy per unit time) and the distance r from the source:

$$f = \frac{L}{4\pi r^2} \tag{8.5}$$

[1] The Lyman limit corresponds to ionization energy for the hydrogen 1s state, which is 911.75 Å in wavelength terms.

Although this relation does not hold in expanding or curved space, it is used to define the luminosity distance d_L, which can be thought of as the distance the source would have from the observer if equation (8.5) were valid.

$$d_L \equiv \sqrt{\frac{L}{4\pi f}} \tag{8.6}$$

Kayser, Helbig and Schramm (1997) have presented a useful summary of the variously defined cosmological distances between two objects, expressing them in terms of the two redshifts, the Hubble constant H_o, the cosmological constant Λ_o, the density parameter Ω_o and the inhomogeneity parameter η (which is the fraction of homogeneously distributed matter in the Universe, as opposed to matter distributed in "clumps" on the scale of the distances being examined). The derivations of these expressions are complicated and some require numerical integration, but in the special case of $\Lambda_o = 0$, $\eta = 1$ and $\Omega_o > 0$ the luminosity distance can be given analytically:

$$d_L(z_x, z_y) = \frac{2c}{H_o}\frac{1}{{\Omega_o}^2}\frac{1}{(1+z_x)^3}\;[\,s(z_y)\cdot t(z_x) - s(z_x)\cdot t(z_y)\,] \tag{8.7}$$

$$\text{with} \quad \begin{aligned} s(z) &= \Omega_o z - \Omega_o + 2 \\ t(z) &= \sqrt{\Omega_o z + 1} \end{aligned}$$

Now it should be noted that equation (8.6) applies to bolometric fluxes f and luminosities L, in which the energy is integrated over all wavelengths or frequencies. Spectroscopic measurements, however, can deliver only (for example) the flux f_ν per unit frequency, usually in units of erg cm^{-2} s^{-1} Hz^{-1}. The quasar flux density $F_{\nu_{LL}}$ is also such a quantity. It is therefore necessary to know how its Lyman-limit flux at the cloud can be obtained in terms of the observed quasar flux at Earth. Because no publications could be found to explain this simple derivation, it is presented here.

Given a quasar with emission redshift z_Q, whose observed spectral flux $f_{\nu_{obs}}$ in erg cm^{-2} s^{-1} Hz^{-1} is measured at the frequency ν_{obs} at Earth. Then the intrinsic QSO luminosity in erg s^{-1} Hz^{-1} at the frequency $\nu_Q = \nu_{obs}\,(1 + z_Q)$ is

$$L_{\nu_Q} = \frac{4\pi {r_Q}^2\, f_{\nu_{obs}}}{1 + z_Q} \tag{8.8}$$

where r_Q is the luminosity distance from Earth to the quasar, and the factor $1 + z_Q$ is the bandwidth correction to account for the fact that the frequency interval observed at Earth is $d\nu_{obs} = d\nu_Q/(1 + z_Q)$ (Weedman 1986).

Now suppose there is a cloud at redshift z_{cloud} with $0 \leq z_{cloud} \leq z_Q$. The local Lyman-limit flux $F_{\nu_{LL}}$ from the quasar, i.e., that seen at the frequency ν_{LL} in the cloud frame, is due to the luminosity at a higher frequency ν_Q in the quasar frame:

$$\nu_Q = \nu_{LL}\,\frac{1 + z_Q}{1 + z_{cloud}} \tag{8.9}$$

Writing equation (8.8) in the form

$$f_{\nu_{obs}} = \frac{L_{\nu_Q}}{4\pi {r_{_Q}}^2} \frac{\nu_{_Q}}{\nu_{obs}} \tag{8.10}$$

and applying it to the cloud frame, we obtain

$$F_{\nu_{LL}} = \frac{L_{\nu_Q}}{4\pi {r_{_L}}^2} \left(\frac{1+z_{_Q}}{1+z_{cloud}} \right), \tag{8.11}$$

where the luminosity distance $r_{_L}$ is now the distance from the cloud to the quasar. Combining equations (8.8) and (8.11) yields the expression

$$F_{\nu_{LL}} = \frac{f_{\nu_{obs}}}{1+z_{cloud}} \left(\frac{r_{_Q}}{r_{_L}} \right)^2 \tag{8.12}$$

where $\nu_{obs} = \nu_{_{LL}}/(1+z_{cloud})$. This, together with equation (8.7) for $r_{_Q}$ and $r_{_L}$, expresses the Lyman-limit flux at the cloud in terms of known quantities—namely the observed flux on Earth and the redshifts of the cloud and the quasar. The result is conveniently independent of the value of the Hubble parameter H_o.

8.3 Continuum Extrapolation

After deriving equation (8.12), the next step was to determine the absolute observed flux $f_{\nu_{obs}}$ from HE 0940. Because the narrow slit width used for the high-resolution echelle spectrum involves a very significant loss of light, absolute spectrophotometric data for HE 0940 had to be sought elsewhere.

Although Reimers et al. (1996) caution that the data from the Hamburg/ESO Survey should not be considered spectrophotometric, there are several reasons to believe that the exposures taken with the 1.52-meter telescope (Fig. 2.1) offer a very good estimate of the absolute spectral flux nonetheless. First of all, the observing conditions for these exposures were good. Secondly, all of the HES spectra were corrected for atmospheric extinction, and the accuracy of the flux calibration is estimated to be better than 20% for most of them. And finally, in the case of HE 0940, simulations of the B-band photometry for the 1.52-meter spectrum are in very good agreement with the measured CCD B magnitude cited in section 2.1 as well as with two separate determinations of the photographic B magnitude that were also available. For these reasons, the flux values can be considered relatively reliable down to about 4000 Å (L. Wisotzki, private communication).

Once this was established, the available flux f_λ in erg cm^{-2} s^{-1} Å^{-1} had to be converted into the flux f_ν in erg cm^{-2} s^{-1} Hz^{-1}. Then, in order to compute the quasar's Lyman-limit flux at a given absorber, it was necessary to determine the intrinsic quasar luminosity at correspondingly higher frequencies (this, often together with the bandwidth correction, is known as the K-correction).

The first step is easily accomplished by considering the energy emitted in the spectral interval $d\lambda$ around any given wavelength λ. If f_λ is the flux per unit wavelength there, and f_ν is the flux per unit frequency in the corresponding frequency interval $d\nu$, then $f_\lambda \, d\lambda = -f_\nu \, d\nu$ because the emitted energy is the same in both cases. The two quantities are related by

$$f_\nu = -f_\lambda \frac{d\lambda}{d\nu} = -f_\lambda \frac{d}{d\nu}\left(\frac{c}{\nu}\right) = \frac{\lambda^2}{c} f_\lambda \tag{8.13}$$

The second step requires additional work because the available spectrum does not cover the relevant wavelength region below the Lyman limit, which would have appeared at 3700 Å. The flux values were therefore extrapolated to those wavelengths by assuming that the quasar continuum can be parametrized by a power law. This common assumption usually takes the form

$$L_\nu \propto \nu^\alpha \tag{8.14}$$

for the intrinsic luminosity L_ν emitted at frequency ν, but the same applies for the flux density f_ν as well:

$$f_\nu \propto \nu^\alpha \tag{8.15}$$

In both cases, the exponent α is assumed to be constant for the quasar or at least for the spectral region under consideration. Common values for the power law in this form are $-1.0 \lesssim \alpha \lesssim -0.5$. Equation (8.13) yields the corresponding wavelength dependence:

$$f_\lambda \propto \lambda^{-(2+\alpha)} \tag{8.16}$$

This shows that the known flux density at a particular wavelength λ_{ref} suffices to calculate the flux density at any other wavelength λ_{obs}:

$$f_{\lambda_{obs}} = f_{\lambda_{ref}} \left(\frac{\lambda_{ref}}{\lambda_{obs}}\right)^{2+\alpha} \tag{8.17}$$

The problem then becomes one of determining the parameter α for the quasar at hand.

This was done by fitting the power law in equation (8.16) to the continuum in the 1.52-meter spectrum. Avoiding absorption lines as well as the broader emission lines, several segments redward of Ly α emission were chosen manually and flagged as belonging to the continuum. Since the coefficient and exponent in a power-law fit will vary drastically depending on their mutual values, some non-linear fitting methods yielded different results for different initial values or even failed to converge. The simplest alternative in this case was to take the natural logarithm

$$\ln f_\lambda = -(2+\alpha) \ln \lambda + \text{const.} \tag{8.18}$$

of both sides of equation (8.17), plot the quantity $y = \ln f_\lambda$ as a function of $x = \ln \lambda$, and fit a straight line through the 104 selected data points. The errors $\sigma_y = \sigma_f / f_\lambda$ in the data points were determined from the flux errors at each point using error propagation. The resulting linear least-squares fit and the corresponding continuum fit are shown in Figure 8.1. The slope $m = -(2+\alpha)$ of the fit yielded a power-law index of $\alpha = -1.06 \pm 0.03$. Of course the quoted error is not very meaningful, since the greater uncertainty comes from the choice of continuum points.

Given the power-law index, all that was needed was a reference wavelength and the flux density there. The flux $f_{\lambda_{ref}} = 6.18 \times 10^{-16}$ erg cm^{-2} s^{-1} Å^{-1} at $\lambda_{ref} = 6756.39$ Å was chosen because it was located away from prominent absorption or emission, in a region where the power-law fit deviated least from the measured fluxes. This allowed a determination of the emitted quasar flux, and hence the intrinsic luminosity, at any wavelength for which the power-law approximation was valid.

For example, with equations (8.8), (8.7), (8.13) and (8.17), the intrinsic Lyman-limit luminosity of HE 0940 was estimated to be

$$L_{\nu_{LL}} = \frac{4\pi {r_Q}^2}{1+z_Q}\frac{{\lambda_{obs}}^2}{c} f_{\lambda_{ref}} \left(\frac{\lambda_{ref}}{\lambda_{obs}}\right)^{2+\alpha} = 2.2 \times 10^{31} h^{-2} \text{ erg s}^{-1} \text{ Hz}^{-1} \tag{8.19}$$

for $H_o = 100h$ km s^{-1} Mpc^{-1} and $\Omega_o = 1$. In this particular case, λ_{obs} is the observed wavelength of the Lyman limit, 911.75 Å $(1+z_Q)$. Similarly, with the help of equations (8.14) and (8.9), equation (8.12) becomes

$$F_{\nu_{LL}} = \frac{L_{\nu_{LL}}}{4\pi {r_L}^2}\left(\frac{1+z_Q}{1+z_{cloud}}\right)^{\alpha} \tag{8.20}$$

where r_L is again the luminosity distance from the cloud to the quasar as given by equation (8.7). These equations are another way of expressing the relation in equation (8.12), which was used as the basis for the Monte Carlo calculations of the proximity effect.

8.4 Monte Carlo Methods

Generally speaking, Monte Carlo calculations refer to statistical methods of studying problems based on the use of random numbers (Bevington and Robinson 1992). This was the approach used to generate simulated data with random redshifts and column densities drawn from the non-uniform distributions in equations (7.4) and (7.5). Most random-number generators produce normalized uniform deviates, so that any number between 0 and 1 is equally probable. The computational problem, then, was to convert such uniformly distributed numbers into non-uniform distributions in the form of power laws.

Lyons (1986) has a very understandable description of three different Monte Carlo methods for generating non-uniform distributions, and summarizes their advantages and disadvantages. In terms of efficiency and computer time, the so-called transformation method used below is preferable to the rejection method or to weighting each

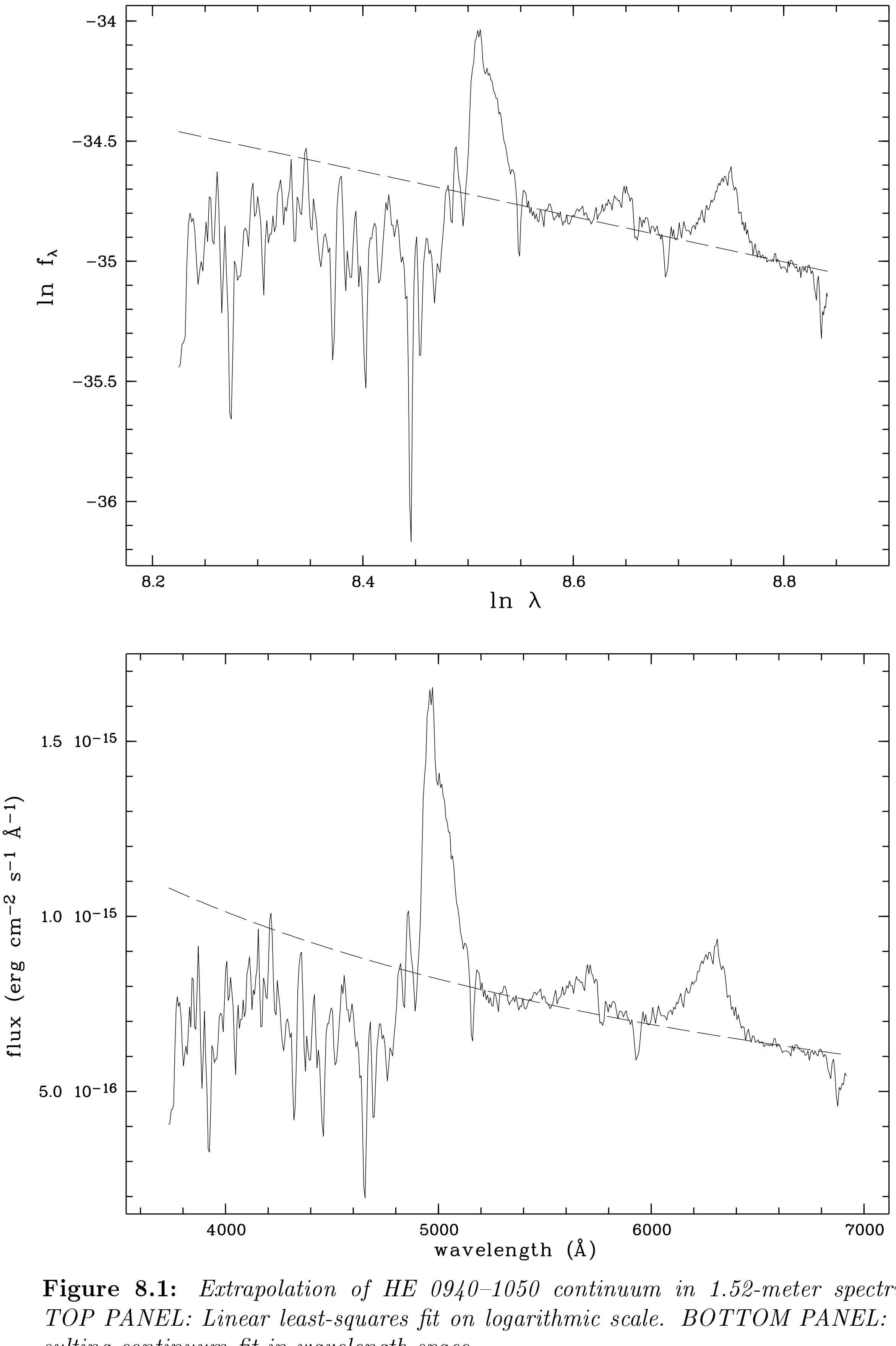


Figure 8.1: *Extrapolation of HE 0940–1050 continuum in 1.52-meter spectrum. TOP PANEL: Linear least-squares fit on logarithmic scale. BOTTOM PANEL: Resulting continuum fit in wavelength space.*

point. The transformation method can only be applied if the function involved can be integrated and then inverted, as was the case here.

Lyons explains that if we transform a variable x with a uniform distribution to $y = y(x)$, the distribution in y will generally not remain uniform. Therefore, to produce a specified non-uniform distribution $y(x)$, the problem is to find the correct transformation.

Given a random variable x, let the probability of generating a value between x and $x + dx$ be $p(x)\,dx$, whose the integral over x is normalized to 1, as is usual for probabilities. For example, for the uniform deviates $x \in (0,1)$ provided by most computer algorithms,

$$p(x)\,dx = \begin{cases} dx & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{8.21}$$

and the normalization is

$$\int_{-\infty}^{\infty} p(x)\,dx = \int_{0}^{1} p(x)\,dx = \int_{0}^{1} dx = 1 \tag{8.22}$$

Given a second random variable $y = y(x)$, with probability distribution $p(y)$. Because these are probabilities, $|p(x)\,dx| = |p(y)\,dy|$ and $f(y) = p(y)$ is a positive function whose integral is 1. Then

$$f(y) = \left| p(x)\,\frac{dx}{dy} \right| = \left| \frac{dx}{dy} \right| \tag{8.23}$$

Solving the differential equation yields the relation between x and y:

$$x = \int_{0}^{y} f(y')\,dy' =: F(y) \tag{8.24}$$

But the real goal is to obtain y as a function of x. If the function F is invertible, the transformation which takes a uniform deviate x into one distributed as $f(y)$ is

$$y(x) = F^{-1}(x) \tag{8.25}$$

Now the equations above (from Press 1992) assume that $y \in (0,1)$. If y has a different domain, namely $y \in (y_{min}, y_{max})$, then the properly normalized function is

$$F(y) = \frac{\int_{y_{min}}^{y} p(y')\,dy'}{\int_{y_{min}}^{y_{max}} p(y')\,dy'} \tag{8.26}$$

In the present case of a power-law probability distribution $p(y) = y^a$ with $a \neq -1$, equation (8.24) becomes

$$x = F(y) = \frac{\int_{y_{min}}^{y} y^a \, dy}{\int_{y_{min}}^{y_{max}} y^a \, dy} = \frac{y^{\,1+a} - y_{min}^{\,1+a}}{y_{max}^{\,1+a} - y_{min}^{\,1+a}} \tag{8.27}$$

so the desired relation is

$$y = F^{-1}(x) = \left[\, y_{min}^{\,1+a} + x \left(y_{max}^{\,1+a} - y_{min}^{\,1+a} \right) \right]^{\frac{1}{1+a}} \tag{8.28}$$

Thus, to obtain a random redshift or column-density value drawn from the correct distribution, all that was necessary was to take the random value between 0 and 1 provided by the computer and apply equation (8.28) using the substitutions below.

$$\begin{array}{llll} \text{For the } z \text{ distribution:} & y \;\rightarrow\; 1+z, & a \;\rightarrow\; \gamma \\ \text{For the } N_{\rm HI} \text{ distribution:} & y \;\rightarrow\; -\beta, & a \;\rightarrow\; N_{\rm HI} \end{array} \tag{8.29}$$

8.5 Monte Carlo Programs

The simulated data sample was composed of 10,000 simulations each for models A1 and A3 of Fardal, Giroux and Shull (1998). Each simulation was a "line list" in which each line was represented by a redshift value z and a column-density value $N_{\rm HI}$. The input to the program which produced these simulations was the model number (A1 or A3), the redshift range $[z_{min}, z_{max}]$, and the minimum and maximum allowed column densities $[N_{min}, N_{max}]$ for the lines to be created.

Normally, the Monte Carlo algorithm derived in the previous section would have enabled a computer program to do a simulation by merely calculating the total number of absorbers predicted by the model, choosing the actual total for the simulation from a Poisson distribution around the predicted total, and drawing random z and $N_{\rm HI}$ values from the appropriate distributions for all of the lines in the simulation. However, as can be seen in Table 7.1, the parameters in models A1 and A3 took on different values in different column-density regions, so these regions had to be handled separately.

The presence of distinct column-density regions required an additional three steps at the beginning of each group of 10,000 simulations:

1. Depending on the model chosen, the boundaries of the column-density regions and the parameter values $\mathcal{A}_o$, β and γ for those regions were stored in an array for easy access.[2]

[2]Note that since the models had been scaled to fit the data, the $\mathcal{A}_o$ values for model A1 were 1.32 times higher, and the $\mathcal{A}_o$ values for model A3 were 1.36 times higher, than those listed in Table 7.1.

2. If the input range $[N_{min}, N_{max}]$ covered more than one column-density region in the given model, it was broken up into smaller intervals whose boundaries matched those of the model regions.

3. For each of these smaller column-density intervals, the appropriate $\mathcal{A}_o$, β and γ values were read from the parameter array. The scaled model prediction for each column-density interval i was determined by integrating equation (7.6) to get

$$\mathcal{N}_{exp,i} = \frac{\mathcal{A}_o}{(1-\beta)(1+\gamma)} \left(z_{max}^{1+\gamma} - z_{min}^{1+\gamma} \right) \left[N_{min,i}^{1-\beta} - N_{max,i}^{1-\beta} \right] \tag{8.30}$$

   where $\mathcal{N}_{exp,i}$ is the number of absorbers with $N_{min,i} \leq N_{\rm HI} \leq N_{max,i}$ expected to be found along the entire line of sight between z_{min} and z_{max}.

Although models A1 and A3 predicted the total number of absorbers $\mathcal{N}_{exp} = \sum \mathcal{N}_{exp,i}$ that would be found in the specified redshift range $[z_{min}, z_{max}]$ and column-density range $[N_{min}, N_{max}]$, of course the *actual* number of absorbers in each simulation varied. They were determined as follows:

4. In each simulation, the algorithm looped through each of the column-density intervals one by one.

5. The actual number of absorbers $\mathcal{N}_i$ in each column-density interval i was determined by using an algorithm from Press et al. (1992) to draw a random number from a Poisson distribution around the mean $\mathcal{N}_{exp,i}$.

6. Each one of the $\mathcal{N}_i$ absorbers in the interval was assigned a redshift z and column density $N_{\rm HI}$ drawn randomly from distributions (7.4) and (7.5), respectively, with the help of equations (8.28) and (8.29).

7. After going through all of the column-density regions in the simulation, the data for the $\mathcal{N}_{total} = \sum \mathcal{N}_i$ lines in the simulation were written to disk in the form of $(z, N_{\rm HI})$ doublets.

8. Steps 4 to 6 were repeated 10,000 times.

The redshift range chosen for the simulations was $z \in (2.0, 3.068)$, and the column-density range was $\log N_{\rm HI} \in (12.5, 19.0)$. The lower column-density limit was chosen to match the lowest common detection limit in the spectrum, while the upper limit was chosen to be substantially higher than the $\log N_{\rm HI} = 17.0$ cut-off in the data sample to avoid depopulating that region when the ionizing flux from the quasar and the UVB were taken into account. The ranges in z and $N_{\rm HI}$ were also wider than those of the data in order to allow for some flexibility in the comparisons.

After the basic "unperturbed" simulations had been generated, other programs could be used to access the data pairs stored on disk and simulate the effects of ionizing radiation from a nearby quasar and the UV background.

1. For each absorber in a given simulation, the ionizing radiation from the quasar at z_{abs} was calculated using equation (8.12).

2. The UV background at $z = z_{abs}$ was either taken to be a constant or calculated using Model I of Bajtlik, Duncan and Ostriker (1988),

$$J_\nu(z) = j_\nu\,(1+z)^{3.5}\,[(1+z_{on})^{1.2} - (1+z)^{1.2}] \tag{8.31}$$

where $j_\nu = 8.17 \times 10^{-24}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1} (see section 8.6).

3. The column density of the absorber was then reduced according to equations (8.4) and (8.3). If the resulting column density dropped below the specified lower limit, the line was dropped from the sample.

4. The redshift distribution of the remaining sample was examined to see how well it compared to the actual distribution observed in the data.

The results of such comparisons are presented in the next section.

8.6 Comparisons to Models

The aim of this section is to look for evidence of a proximity effect in the spectrum of HE 0940–1050. The most effective way to do this is to plot the number of observed lines $\mathcal{N}$ as a function of the redshift, and to see if there are less lines close to the quasar redshift than are predicted by the scaled versions of models A1 and A3 of Fardal, Giroux and Shull.

As mentioned earlier, specific intervals in $N_{\rm HI}$ and z must be chosen in order to evaluate the integral in equation (7.6) and obtain concrete line counts. The choice of the largest interval in $N_{\rm HI}$ is clear, since the full data sample already has the range $13.33 \leq \log N_{\rm HI} \leq 17.0$. But since the simulations by Dobrzycki and Bechtold (1991b) indicate that the number counts of weak lines may be more strongly affected by quasar radiation, it will also be interesting to see if there is a dependence on column density. Therefore the data sample has also been divided up into a subsample of 114 "weak" lines with $13.33 \leq \log N_{\rm HI} \leq 14.0$, and a subsample of 55 "strong" lines with $14.0 \leq \log N_{\rm HI} \leq 17.0$. It was not possible to study smaller column-density intervals because of the already very poor statistics; for example, there were only 7 lines with $\log N_{\rm HI} \in [15.0, 16.0]$ and only 3 lines with $\log N_{\rm HI} \in [16.0, 17.0]$.

The comparisons between the simulations and the data are presented as plots of $\log(d\,\mathcal{N}\,/dz)$ vs. $\log(1+z)$. This form was chosen because taking the logarithm of both sides of equation (7.4) yields

$$\log \frac{d\mathcal{N}}{dz} = \gamma\ \log(1+z) + \log A_o \tag{8.32}$$

bin	from $\log(1+z)$	to $\log(1+z)$	z_{min}	z_{max}	λ_{min}	λ_{max}
1	0.534100	0.537884	2.420565	2.450514	4158.3	4194.7
2	0.537884	0.544884	2.450514	2.506580	4194.7	4262.8
3	0.544884	0.551884	2.506580	2.563557	4262.8	4332.1
4	0.551884	0.558884	2.563557	2.621460	4332.1	4402.5
5	0.558884	0.565884	2.621460	2.680304	4402.5	4474.0
6	0.565884	0.572884	2.680304	2.740105	4474.0	4546.7
7	0.572884	0.579884	2.740105	2.800876	4546.7	4620.6
8	0.579884	0.586884	2.800876	2.862636	4620.6	4695.7
9	0.586884	0.593884	2.862636	2.925398	4695.7	4772.0
10	0.593884	0.600884	2.925398	2.989181	4772.0	4849.5
11	0.600884	0.607884	2.989181	3.054000	4849.5	4928.3

Table 8.1: *Bins chosen for plots of* $\log(d\mathcal{N}/dz)$ *vs.* $\log(1+z)$.

which is a straight line in the variables $y = \log(d\mathcal{N}/dz)$ and $x = \log(1+z)$, making it easier to visually compare the data with the models. A bin size of 0.007 in $\log(1+z)$ was chosen for all the plots that will be presented here, though the redshift cut-off at the Ly β emission line makes the lowest bin only about half this size (0.003786). The bin entries were calculated by counting the number of lines $\mathcal{N}$ in each bin and dividing by the total redshift interval Δz covered by the bin, minus any gaps in redshift. The logarithm of the result was then plotted on the yaxis. Thus both the smaller bin size and the two regions excluded in section 7.4 were taken into account.

The disadvantage of this choice of representation is that the z values are not immediately evident. To facilitate comparisons between the plots here and the spectrum in Appendix B, the bin boundaries in both z and λ are listed in Table 8.1.

The calculations for the Monte Carlo simulations of models A1 and A3, both with and without ionizing radiation, were done in the same way—except that after the bin counts for the 10,000 simulations had been determined they were averaged for direct comparison to the data.

Forest models and photoionization model

In all of the figures that follow, the data points remain the same and the various models are shown as dashed or dotted lines which also indicate the size of the bins. Figure 8.2 shows the data with the scaled model predictions in the absence of quasar radiation. Even before the proximity-effect simulations are introduced, a few features seem obvious.

- The fit for the data sample as a whole, in the top plot, is generally good; except in the first and last bins, it falls almost completely within range of the (Poissonian,

$1\,\sigma$) error bars of the data. The two bins closest to the quasar seem to indicate a possible proximity effect.

- In the weak-line sample, the fit is also generally good, though the first two bins lie above, and the fifth bin lies well below, the model predictions. The strong drop in number counts in the bin closest to the quasar is even more pronounced in the weak-line sample than in the sample as a whole.

- The fit is generally not quite as good for the strongest lines as in the other two cases. Any proximity effect that might be present in the last two bins is much smaller than the deviations elsewhere, most notably in the three consecutive bins that lie clearly below the model predictions. This may be the result of the poorer statistics for this subsample; for example, the three low redshift intervals contain only 3 entries each.

Figure 8.3 shows the fits that result when the column densities in the simulations are reduced using the photoionization model and a particular model for the UVB background presented by Bajtlik, Duncan and Ostriker (1988). Their Model I was obtained by integrating a luminosity function and extrapolating to higher redshifts. When determining the integrated background radiation, it assumes that all quasars turn "on" at $z_{on} = 4$. It also assumes $\Omega_o = 1$, so this was the value adopted for all of the calculations and plots in this section as well.

The ionization model fits the data well, yielding an especially good fit for the full sample, falling within the error bars of all but the high first bin, and hitting the value in the last bin almost exactly. The fit for the weaker lines is also good, although it looks like the proximity effect for those lines is probably even stronger. The fit for the strongest lines, on the other hand, offers only a very slight improvement over the models without radiation, and this seems to be primarily because the fit happens to fall above and below the last two bin values, respectively. The χ^2 still yields an acceptable probability; but in this case the model appears to serve more as an upper limit for the possible quasar influence on the stronger lines, because any further depression in the last bin will only worsen the fit.

This was precisely the effect that was seen when the UVB at the Lyman limit was assumed to be constant with redshift at $J_{\nu_{LL}} = 10^{-21}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1} (see Figure 8.4). The comparison here yielded an even higher χ^2 than did the models without any proximity effect whatsoever. There is thus no indication of a proximity effect among the stronger lines in the current spectrum.

The χ^2 values for the fits using the unperturbed Ly α forest, the photoionization model, and the constant value of $J_{\nu_{LL}}$ are all summarized in Table 8.2. The reduced χ^2 is a factor of 10 smaller, because there are 11 bins and because scaling models A1 and A3 beforehand introduced an additional constraint.

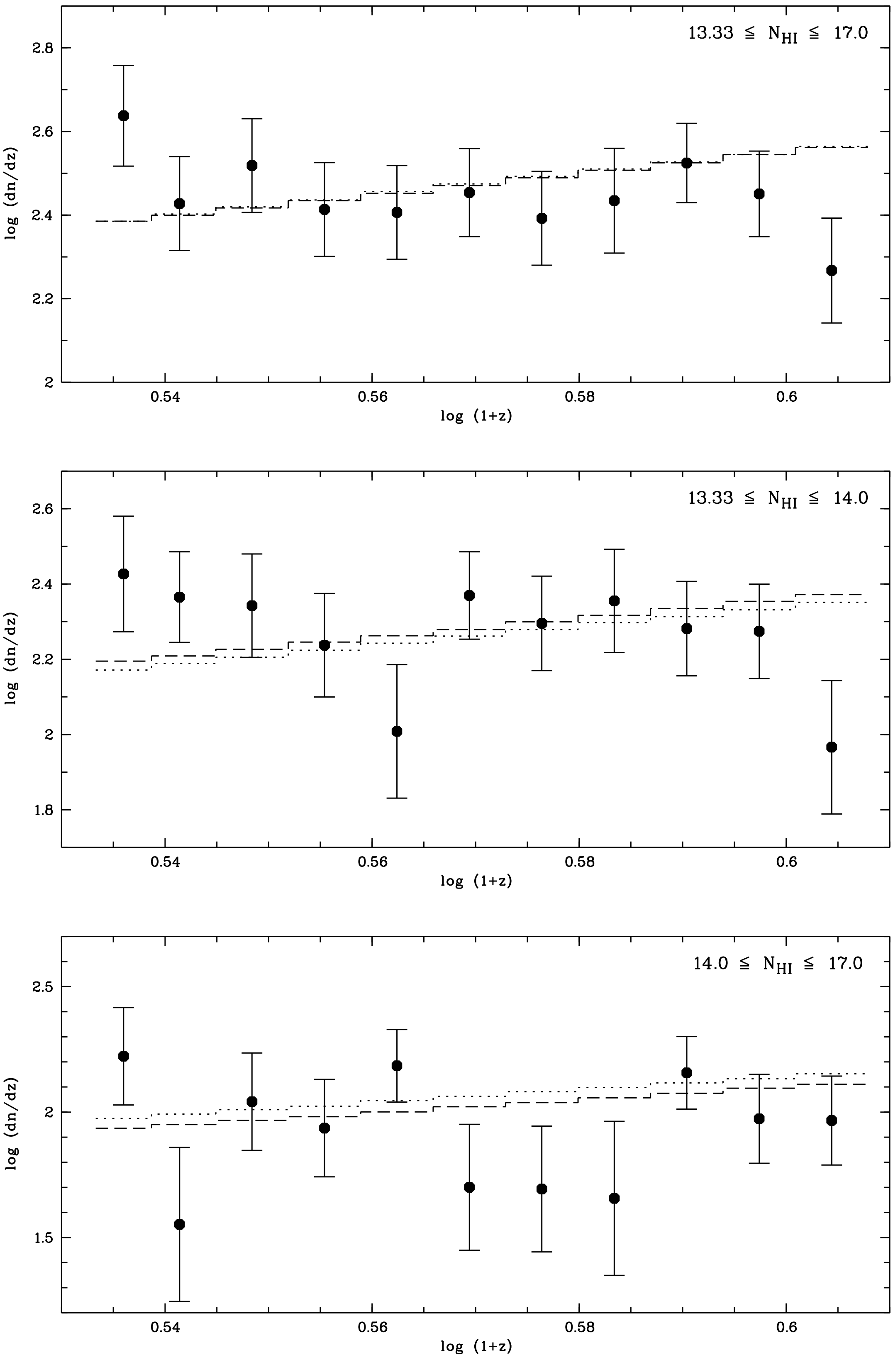


Figure 8.2: *Observed number of lines per unit redshift $d\mathcal{N}/dz$ vs. redshift z for the data, compared to models A1 (dashed line) and A3 (dotted line)* **without any effects of quasar radiation**.

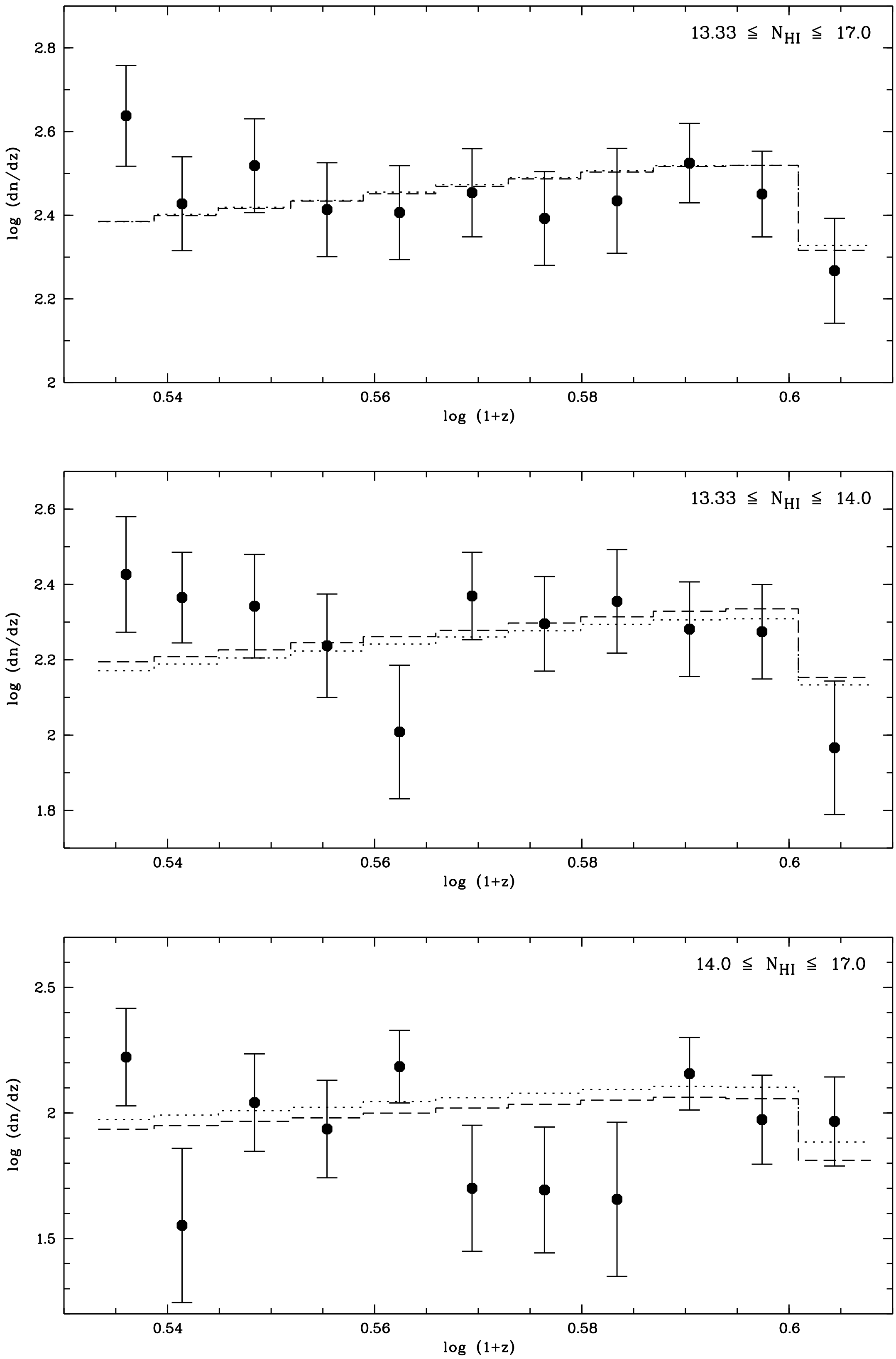


Figure 8.3: *Observed number of lines per unit redshift $d\mathcal{N}/dz$ vs. redshift z for the data, compared to models A1 (dashed line) and A3 (dotted line)* **with UVB "Model I" from Bajtlik, Duncan and Ostriker (1988)**.

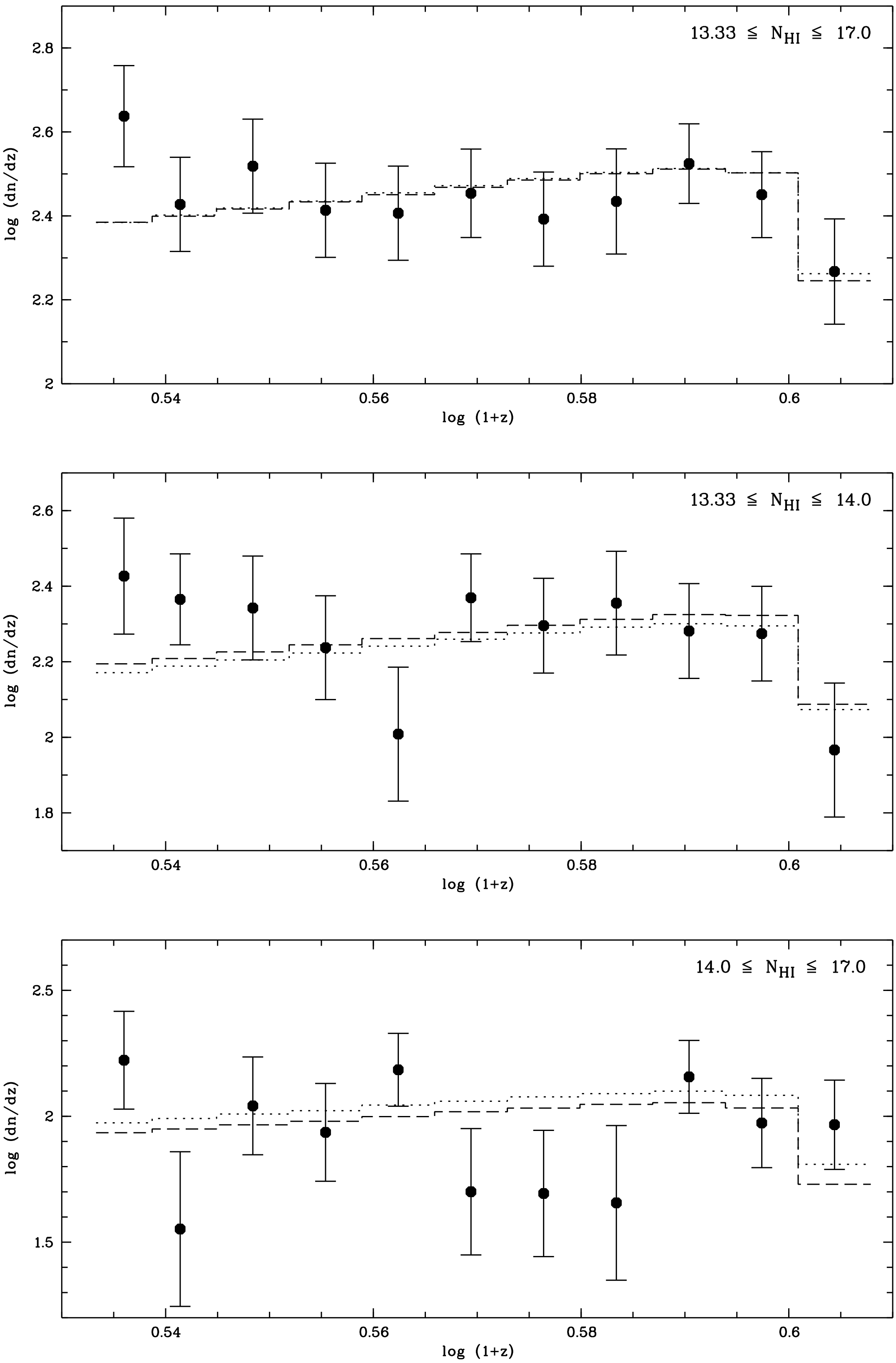


Figure 8.4: *Observed number of lines per unit redshift $d\mathcal{N}/dz$ vs. redshift z for the data, compared to models A1 (dashed line) and A3 (dotted line)* **with $J_{\nu_{LL}} = 10^{-21}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1}**.

		$N_{\rm HI} \in [13.33, 17.0]$	$N_{\rm HI} \in [13.33, 14.0]$	$N_{\rm HI} \in [14.0, 17.0]$
no radiation	A1	13.6 (19.2%)	14.0 (17.5%)	13.7 (18.9%)
	A3	13.8 (18.2%)	14.5 (15.1%)	15.0 (13.2%)
BDO model	A1	7.6 (67.1%)	9.3 (50.4%)	13.5 (19.7%)
	A3	7.7 (65.4%)	10.3 (41.6%)	13.7 (18.6%)
$J_\nu = 10^{-21}$	A1	7.2 (70.4%)	8.6 (57.6%)	14.5 (15.3%)
	A3	7.3 (70.0%)	9.7 (46.5%)	14.1 (16.8%)

Table 8.2: *χ^2 and corresponding probabilities (in parentheses) for fits of the Ly α forest with and without quasar radiation.*

Estimate of the UV Background

Finally, to get a better idea of the magnitude of the possible constraints such proximity-effect simulations might have on the UV background, its value J_ν at the Lyman limit was assumed to be merely constant over the entire redshift range from $z = 2.421$ to $z = 3.054$. This constant value was then varied over more than two orders of magnitude to see which values were most compatible with the data, which was now confined to the full sample and the subsample of weaker lines. Following Bechtold (1994), the χ^2 dependence of the fit is plotted in Figure 8.5 and the values are listed in Table 8.3. Model A1 was used for these calculations because it described the data better than model A3 in all but one of the six cases presented in Table 8.2.

Taking the lowest point as the best estimate for the metagalactic ionizing background, this yields a rough approximation of

$$\log\, J_\nu \;=\; -21.0 \pm 0.7 \qquad \text{from } 13.33 \leq N_{\rm HI} \leq 17.0 \tag{8.33}$$

$$\log\, J_\nu \;=\; -21.3 \pm 0.7 \qquad \text{from } 13.33 \leq N_{\rm HI} \leq 14.0 \tag{8.34}$$

at the Lyman limit, which is the order of magnitude that used to be most commonly cited for the proximity effect. Lu, Wolfe and Turnshek (1991) obtained the same value of $\log J_\nu = -21.0 \pm 0.5$ as the original result by Bajtlik, Duncan and Ostriker, but report $\log J_\nu \sim -22.3$ with Keck spectra and a different analysis method at the higher redshift of $z = 4.13$. Although J_ν may decrease at higher redshifts, Giallongo et al. (1996) obtain $J_\nu = (5 \pm 1) \times 10^{-22}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1} for $1.7 < z < 4.1$ and argue that earlier values of $\log J_\nu > -21$ resulted because studies with lower resolution and lower S/N than the Keck data assumed a column-density distributions in equation (7.5) that were too steep.

The HE 0940 result above also appears to confirm the assertion by Dobrzycki and Bechtold that the effect is stronger for weaker lines. As mentioned in section 8.7, the photoionization model assumes that the Ly α lines are optically thin; obviously this assumption does not generally hold. Even though smaller lines are much more numerous than saturated ones, many of these are lost in the spectra of all but the highest quality. In any case, a more careful treatment of the ionization-induced proximity effect would need to take curve-of-growth effects into account as well.

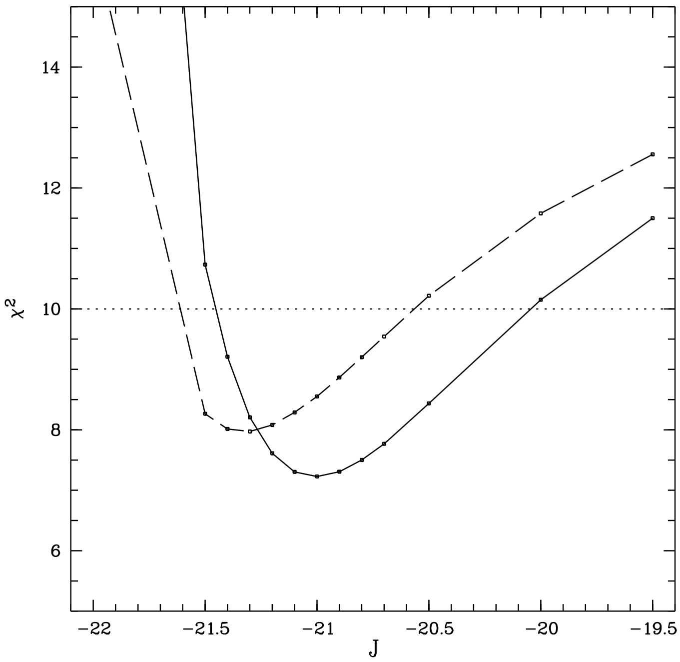


Figure 8.5: χ^2 *vs.* $\log J_{\nu_{LL}}$ *erg cm*^{2} *s*$^{-1}$ *Hz*$^{-1}$ *sr*$^{-1}$ *for the full sample of Ly*α *lines (solid line) and for weaker lines with* $13.33 \leq \log N_{\rm HI} \leq 14.0$ *(dashed line).* $J_{\nu_{LL}}$ *is assumed to be constant with redshift.*

8.7 Proximity-Effect Uncertainties

There are many sources of uncertainties and possible systematic errors involved in the determination of the metagalactic UV background by means of the proximity effect, or more specifically from the ionization model described in section 8.1. Some of the errors have to do with the difficulty of accurately determining the quasar flux density, but others have more to do with the basic assumptions of the model.

Quasar flux density

If no spectrophotometric data is available for the region at or below the observed Lyman limit, extrapolating the quasar flux density to lower wavelengths, as was done in section 8.3, could introduce significant errors. In addition to the errors in the fit itself, the uncertainties in the subjective choice of continuum points could have been further

$\log J_\nu$	χ^2 (full sample)	Q_χ (full sample)	χ^2 (weak lines)	Q_χ (weak lines)
–19.5	11.50	0.32	12.56	0.25
–20.0	10.15	0.43	11.58	0.31
–20.5	8.44	0.59	10.22	0.42
–20.7	7.77	0.65	9.54	0.48
–20.8	7.50	0.68	9.20	0.51
–20.9	7.31	0.70	8.87	0.54
–21.0	7.23	0.70	8.55	0.58
–21.1	7.30	0.70	8.29	0.60
–21.2	7.61	0.67	8.08	0.62
–21.3	8.21	0.61	7.97	0.63
–21.4	9.21	0.51	8.01	0.63
–21.5	10.73	0.38	8.26	0.60
–22.0	33.37	0.00	16.11	0.10

Table 8.3: χ^2 *and corresponding probabilities for fits of* $J_\nu = const.$

biased by the possible presence of weak unresolved emission lines. It appears possible to avoid many of these errors by means of more complicated weighting procedures—developed to measure the H I Gunn-Peterson effect—which use correlations in adjacent pixel fluxes to identify the continuum redward of, and even within, the Ly α forest (Fang et al. 1998). But by far the greatest systematic error in the estimate of the quasar's flux density would result if it turns out that the assumed power law does not describe the quasar continuum well. Though this assumption is widely used, no systematic study or quantitative measure of its validity could be found in the recent literature.

Other authors (Bajtlik, Duncan and Ostriker 1988; Lu, Wolfe and Turnshek 1991) have used the empirical fitting formula (Tytler 1987)

$$\log F_\nu (1 + z_Q)^{-1} = -0.4(m_v - k_v) - 20.09 \tag{8.35}$$

to estimate the observed Lyman-limit flux density F_ν from the quasar's visual magnitude with K-corrections from Evans and Hart (1977). However, in addition to pointing out that these K-corrections do not extend beyond $z = 2.5$, Bechtold (1994) has shown that the resulting F_ν values differ from both the power-law extrapolations and the directly observed Lyman-limit fluxes by more than the error of $\lesssim 40\%$ quoted by Tytler. Instead, this should probably be considered a lower limit for the Lyman-limit flux estimate in general.

As emphasized in section 8.1, one important assumption that is made at the very beginning is that the spectral shape of the ionizing background is the same that of individual quasars. Only then would the ratio of the ionization rates from the UV background and the UV quasar flux equal the ratio of their Lyman-limit flux densities.

But the redshift-dependence of UVB models by Haardt and Madau (1996) and Fardal, Giroux and Shull (1998) exhibits what the latter call "bumps and wiggles" in the background that make this assumption unlikely. Nevertheless, Bechtold (1995) argues that the difference has little effect for "reasonable" spectral energy distributions and estimates the uncertainty introduced to be $\sim 25\%$.

Basic assumptions

The comprehensive list by Bechtold (1995) also includes other systematic effects, which are reiterated only briefly here.

- If the shape of the column-density distribution is flatter than assumed, then J_ν will decrease significantly (Chernomordik and Ozernoy 1993).
- In the unlikely case that the clouds are not as highly ionized as believed, more UV photons from the quasar are needed to produce the observed proximity effect and the actual J_ν is lower than its estimate.
- Similarly, if the absorption lines are saturated rather than optically thin, then the reduction in N_{HI} and thus the ratio of quasar flux to UVB was actually greater than assumed; so the actual J_ν will be lower than its estimate (Chernomordik and Ozernoy 1993).
- If quasars vary on time scales comparable to the $\sim 10^4$ years needed for clouds to reach ionization equilibrium (Bajtlik, Duncan and Ostriker 1988), the observed flux may differ appreciably from the average of the flux responsible for the ionization of the cloud.
- If amplification by gravitational lensing makes quasars appear brighter, then J_ν will also be overestimated.
- If quasars appear dimmer because of dust from intervening galaxies, then the J_ν will also be underestimated.
- If quasar redshifts are systematically higher than assumed, then the actual J_ν has been overestimated (Espey 1993).

Redshift determinations

This last point bears some further discussion because it may radically change the results presented in the previous section. Studies of quasar emission lines (e.g., Gaskell 1982) have shown that the high-ionization lines such as C IV are systematically blueshifted with respect to the lower-ionization lines such as O I and Mg II by an average of roughly 1000 km s^{-1}. Since the C IV "line" at 6280 Å in the 1.52-meter spectrum (Figure 2.1) determined the quasar redshift of $z_{em} = 3.054$ that was used throughout this study, it is quite possible that systematic effects have led to a serious overestimate in the ionizing UV background.

The six absorption lines above 4928.3 Å in the spectrum toward HE 0940 lend further support to this suspicion. The five lines that fall between 4928.3 and 4944 Å are included in Appendices B and C, while the sixth can be seen at 4959 Å in Appendix A. All efforts to identify these lines as metals failed, so assuming they are Ly α lines and using them to estimate the quasar's emission redshift would lead to values of $z = 3.0668$ or $z = 3.0792$, depending on whether the redmost of the group of five or the single line is chosen. The corresponding velocity shifts of 950 km s^{-1} or 1900 km s^{-1} fall well within the range of known values (e.g., Espey et al. 1989).

If the quasar's emission redshift is in fact higher, then the quasar is farther away from the affected Ly α clouds and its sphere of influence is larger than had been assumed. Since a smaller QSO flux produced the observed proximity effect, the influence of the UV background must have been proportionally smaller as well. While the calculations of the preceding section would have to be redone with a higher QSO redshift to determine the exact change in the UVB estimate, it is possible to get an idea of the size of this effect by looking at studies that include a larger sample of quasars. In one such study, Espey (1993) prudently uses each quasar's previous emission-line redshift as the upper limit of the range in which to re-evaluate the proximity effect for a higher QSO redshift. This is important because there is always the danger that the entire proximity effect can be created as a mere artifact of choosing an inaccurately high redshift.

As expected, the results of Espey's study show that an underestimate of the QSO redshift leads to an underestimate of its sphere of influence and an overestimate of the UV background intensity. In his reanalysis of the sample used by Lu et al. (1991), this discrepancy amounted to up to a factor of 6 or 7 for assumed velocity shifts of up to 2500 km s^{-1}. Judging from these results, the UV background in the previous section would have been overestimated by a factor of about 3 to 5, again depending on which of the Ly α lines mentioned above is chosen to be indicative of the "true" emission redshift of HE 0940.

General reflections

As can be seen from the discussion here and the list above, the systematic effects are many; some are difficult to quantify, and several are potentially very large. It is therefore not surprising that, until now, the proximity effect has provided estimates of the UV ionizing background to only within an order of magnitude (which is also the about the size of the greatest discrepancies in current comparisons to theoretical models).

In addition to all of these considerations, the most basic assumption underlying proximity-effect studies of this sort is that the redshift distribution of Ly α absorbers is the same everywhere except when radiation from a nearby quasar leads to greater ionization of the hydrogen and thus to fewer Ly α clouds in that vicinity. But such a decrease need not necessarily stem from enhanced ionization. Some other environmental factors around the quasar may be responsible for the effect instead, for example if quasars exist in high-density environments where velocity dispersions are high (Fernández-Soto et al. 1995) or if neutral gas absorption is suppressed in a high-pressure environment due, e.g., to a nearby supercluster (Rauch et al. 1992, Rauch 1998). Until such alternative explanations can be evaluated or until the large uncertainties listed above can be reduced significantly, it will be difficult to establish whether the proximity effect does indeed measure the UV background intensity. But even as there seems to be a growing awareness that local conditions may play just as large a role, the estimates of the UV background based on the proximity effect seem to be approaching those based on theoretical models of the integrated metagalactic background; the comparison, and the concept of the proximity effect itself, may indeed prove useful.

Chapter 9

Investigation of a Possible Void

After proposing the ionization model for the proximity effect, Bajtlik, Duncan and Ostriker (1988) examine the suggestion by B. Paczyński that an isotropically emitting foreground quasar could also deplete the Ly α clouds in the nearby lines of sight to other quasars. The possibility of such a two-quasar proximity effect has led to some searches for voids in the Ly α forest (e.g., Fernández-Soto et al. 1995), but until now only two Mpc-sized voids have been found in the literature (see section 9.1).

With two quasars, additional factors come into play. For example, if the foreground quasar radiates anisotropically, a void might still be observed in the other spectrum, although generally not at the same redshift as that of the foreground quasar. But the anisotropy would also mean that we might fail to see a void even if a sufficiently bright quasar is seen close enough to the line of sight. Alternatively, we might see a void in the vicinity of a quasar that appears too faint to have caused the void; or in the most extreme situation, we may observe no quasar at all. The same effects could also appear if the foreground quasar varies on time scales shorter than the light-travel time to the void or the equilibriation time scales for the Ly α clouds. This could make it difficult to use the two-quasar proximity effect as an estimate of the ionizing UV background at a void.

On the other hand, these very difficulties would make interesting research topics in themselves. With large enough statistics it could be possible to determine precisely whether the quasar radiation is "beamed" or obscured in certain directions, as in some models of the active galactic nuclei—and, if so, what shape or opening angles the radiation should have.

9.1 Estimate of Void Size

Looking at the spectrum of HE 0940, there is a conspicuous region of approximately 30 Å that seems to be relatively free of lines. One possible cause is of couse that there is another quasar along the line of sight, causing the void by means of photoionization.

If we consider that the possible void could extend the entire distance between the line at 4336.4 Å and the line at 4367.4 Å, this corresponds to a redshift interval of

$$\Delta z = 2.592551 - 2.567062 = 0.025489 \tag{9.1}$$

Converting this to a luminosity distance using equation (8.7) and the parametrization for the Hubble parameter $H_o = 100h$ km s^{-1} Mpc^{-1} yields an estimated void size of $3.2h^{-1}$ Mpc for $\Omega_o = 1$, or a size of $4.9h^{-1}$ Mpc for $\Omega_o = 0.2$. Figure 9.1 illustrates better how the luminosity distance corresponding to this particular redshift interval depends on the density parameter.

The estimate above must be taken as an upper limit for the size of the possible void, because there is at least one small line visible within this range. Further small lines were identified upon closer inspection (see section 9.2), so if the region does turn out to be statistically significant it would technically be what Kovner and Rees (1989) call a "clearing," to distinguish strongly underdense regions from those containing no lines at all. In the following, however, the term "void candidate" or "void" will continue to be used loosely to refer to the 30 Å region under consideration.

Even if any such lines were ignored completely, the resulting void would still be smaller than two others found in the literature. Dobrzycki and Bechtold (1991a,b) report a void in the spectrum of Q0302–003 between $\lambda \approx 5050$ Å and $\lambda \approx 5100$ Å, which is a redshift interval of $\Delta z \approx 0.04$ at $z \approx 3.17$. According to equation (8.7), this corresponds to a size of approximately $3.5h^{-1}$ Mpc and $5.6h^{-1}$ Mpc for $\Omega_o = 1$ and $\Omega_o = 0.2$, respectively.[1] Srianand (1997) also reports a void of about 50 Å in the spectrum of Tol 1038–2712, with $\Delta z = 0.0423$ centered at $z = 2.18635$. But since it is at a lower redshift, the estimated size is larger: $7.1h^{-1}$ Mpc and $10.6h^{-1}$ Mpc for $\Omega_o = 1$ and $\Omega_o = 0.2$, respectively.

9.2 Presence of Small Lines

Since there was the small but obvious absorption line at 4354 Å, the entire 30 Å region was studied more closely to determine the significance level of any other possible lines. Using the method described in section 6.3, the $5\,\sigma$ detection limit was determined to correspond to $\log N_{min} = 12.5$. The line at 4354 Å has a column density of $\log N_{\rm HI} = 13.0$ and—not surprisingly—was thus significant to the $15\,\sigma$ level. Two additional lines above this detection limit were also found at 4340 Å and 4358 Å. The errors in position and column density were twice that of the first line, and the errors in b were four and five times higher (see Table 9.1). However, the fitted column densities of $\log N_{\rm HI} = 12.9$ and $\log N_{\rm HI} = 12.7$, respectively, correspond to $12\,\sigma$ and $8\,\sigma$ features; so they were not dismissed.

[1] Dobrzycki and Bechtold quote proper sizes of $4.5h^{-1}$ and $11.25h^{-1}$ Mpc for the given values of Ω_o. However, in the text above the sizes were given in terms of luminosity distance rather than proper distance for convenience and for consistency, so that the three cases could be compared directly.

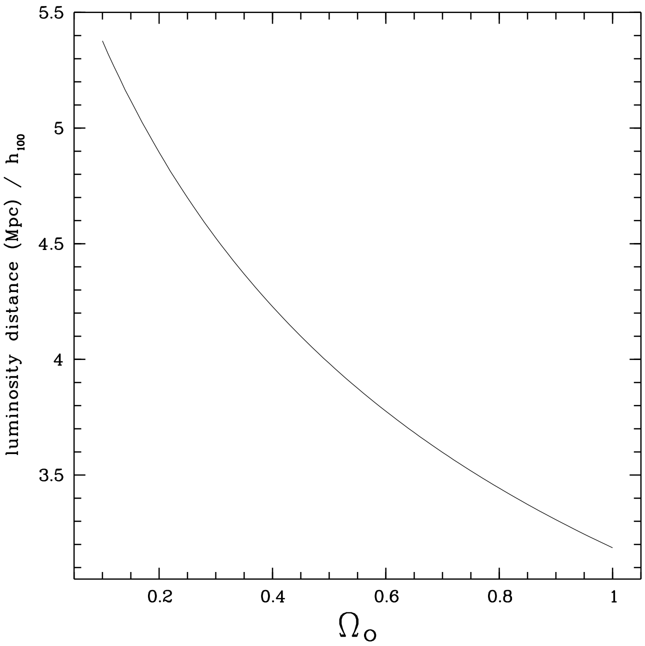


Figure 9.1: *Void size vs. density parameter* Ω_o

λ	σ_λ	z	$\log N_{\mathrm{HI}}$	$\sigma_{\log N}$	b	σ_b
4336.37	0.17	2.567062	13.41	0.15	34.9	11.4
4339.71	0.11	2.569806	12.87	0.20	27.7	27.9
4354.12	0.05	2.581664	12.99	0.08	20.2	5.0
4357.84	0.17	2.584723	12.69	0.19	25.2	19.4
4367.36	0.05	2.592551	13.21	0.07	27.6	4.7
4369.89	0.04	2.594638	16.97	0.70	31.1	3.6

Table 9.1: *Lines in and bounding the void candidate.*

Dobrzycki and Bechtold (1991b) also found lines in their simulations of quasar-induced voids in absorption spectra. They generated artificial spectra from the usual redshift and column-density distributions and a Gaussian Doppler-parameter distribution with $\langle b \rangle = 35$ km s^{-1} and $\sigma_b = 7$ km s^{-1}. The interesting result is that even when the foreground QSO was very strong, significant lines remained even in the part of the spectrum where the quasar's influence was greatest. The authors therefore concluded that the first signature of a two-quasar proximity effect is a lack of weak lines, rather than a lack of lines in general. This makes sense when one considers that lines on the linear curve of growth are already so close to the detection limit that they will quickly fall below it if their column densities decrease by a small amount; while strong lines will tend to remain saturated.

It is therefore interesting to note that in the present void candidate only relatively small ($N_{\rm HI} \leq 10^{13}$) lines are visible. One explanation could be that the region represents merely a region of low density in the large-scale distribution of matter (or at least of neutral hydrogen), rather than the sphere of influence of a foreground quasar.

9.3 Significance of the Void

Probably the best way to estimate the significance of a possible void in a spectrum is to create simulated spectra with the same resolution, signal-to-noise ratio, and sampling as their observed spectrum, and to analyze these spectra with the same automated routines used on the actual data. This is what done by Dobrzycki and Bechtold (1991b), who then adjusted the total number of lines and the power-law exponent γ in their simulated spectra until they obtained the same distribution of gaps as in the actual spectrum. The best-fit parameters were then used to generate 50,000 simulations, which showed only 11 gaps that were as large or larger than the observed gap. The authors thus concluded that the probability of finding such a void is $(2.2 \pm 0.7) \times 10^{-4}$.

A much simpler approach was taken by Srianand (1997), where lists of z values were generated for the observed redshift window. The distribution was assumed to be Poissonian because of the limited redshift interval under consideration. The approach taken here used the distributions in equation (7.6) instead, since it took advantage of the two Monte Carlo sets of 10,000 simulations that had already been created as described in section 8.5. The procedure is basically the one described by Rauch et al. (1992). The significance of the void was evaluated by first determining the distribution of maximum redshift intervals containing 0, 1, 2 and 3 lines. The redshift range that was considered was the same as in the spectrum of HE 0940–1050, spanning from $z = 2.4206$ to $z = 3.054$. In this first estimate, the column-density cut-off was chosen to match the detection limit of $\log N_{min} = 12.5$ that had been determined in section 9.2.

The lines in each of the $2 \times 10,000$ were first sorted according to redshift. A simple computer algorithm went through the list, searching for the maximum redshift interval between two consecutive lines as well as the largest redshift intervals containing 1, 2 and 3 lines. The largest gap sizes were recorded for each case. After all of the simulations had been processed in this way, a histogram was made of the maximum redshift intervals. Figure 9.2 shows the distributions of these maximum redshift intervals, plotted with a bin size of $\Delta z = 0.008$. It is reassuring to see that they have the same form as the analytical expression for the a priori probability distribution derived by Ostriker, Bajtlik and Duncan (1988).

The probability of such redshift gaps occurring by chance was then estimated by counting how many gaps of equal or greater size appeared in the 10,000 simulations for each model. The fact that the observed spectrum had the two small absorption lines

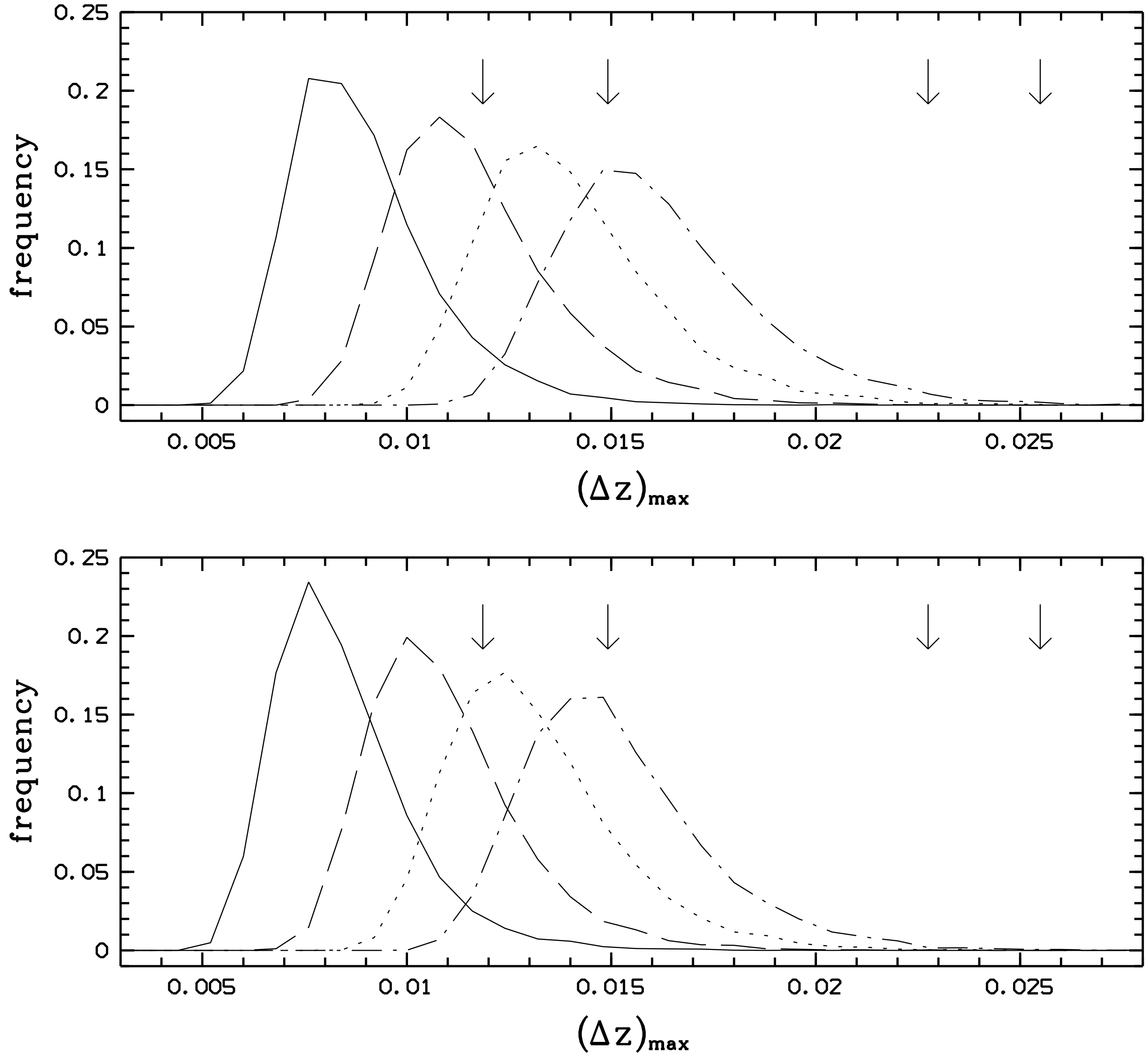


Figure 9.2: *Distribution of maximum redshift intervals containing 0, 1, 2 and 3 lines (curves from left to right), for* $\log N_{\rm HI} \geq 12.5$. *TOP: Model A1. BOTTOM: Model A3. The arrows indicate the sizes of the observed gaps toward HE 0940.*

Redshift interval	Number of lines	Probability (model A1)	Probability (model A3)
0.011858	0	6.40 %	3.62 %
0.014917	1	6.96 %	3.47 %
0.022745	2	0.32 %	0.12 %
0.025489	3	0.29 %	0.11 %

Table 9.2: *Probability of a larger redshift interval with 0, 1, 2 and 3 lines occurring by chance in 10,000 simulations for model A1 and A3, respectively.*

right in the middle of the void increased this probability greatly, since gaps roughly half of the full size of the void are relatively common: between 3% and 6%, judging from the results for models A3 and A1, respectively (see Table 9.3). However, some impression of the full void's significance can be gained by considering the z interval with two or three lines, where the probability of obtaining these two types of intervals by chance drops to 0.3% for model A1 and 0.1% for model A3. If Poisson statistics are used to determine the errors, the combined results from both models show that the probability of finding a void of this size with three lines is $(2.0 \pm 0.3) \times 10^{-3}$.

To be fair, it could be argued that a much higher cut-off of $\log N_{min} = 13.33$ would be more appropriate, given the incompleteness of the data at low column densities. In that case, the three small lines within the region would fall below this limit, as would the line on the right boundary. The saturated line to its right would thus become the new boundary of the possible void (now without any lines at all), so this time the computer algorithm searched for slightly larger redshift intervals of $\Delta z \geq 0.027576$ containing 0 lines. The resulting distributions are shown in Figure 9.3. In the 10,000 simulations for each model, the algorithm found 590 and 548 gaps of equal or greater size. The corresponding probability of finding such a gap by chance is between 5% and 6%, so from this point of view the void candidate appears insignificant, the mere result of statistical fluctuations.

The difference between the two estimates with different column-density cut-offs could be a reflection of the fact that such attempts to estimate the statistical significance of a void are extremely sensitive to very slight changes in the estimated line density (Dobrzycki and Bechtold 1991b). Scaling models A1 and A3 to match the line counts in the spectrum itself was in fact a form of estimating the line density. Since the void was included in this estimate, perhaps a new determination that excludes this part of the spectrum would yield a higher line density and thus a higher significance for the void—though this is debatable given the relatively small fraction of the spectrum that it represents.

The uncertainty seems to arise because of the difference between the detection limit and the estimated completeness limit in this spectrum.[2] While the second estimate of the void significance seems to indicate that the stronger lines show no particularly

[2] Unfortunately, no reference to either of these limits was found in the publications cited here.

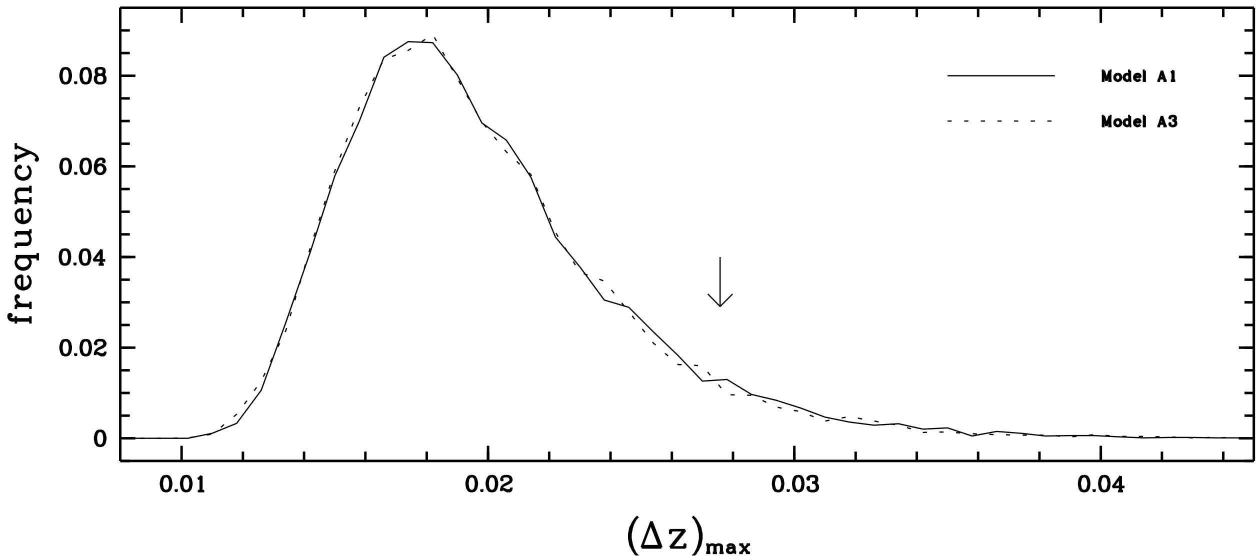


Figure 9.3: *Distribution of maximum redshift intervals between lines with* $\log N_{\rm HI} \geq$ **13.33***, obtained from simulations of model A1 (dashed line) and A3 (solid line). The arrow indicates the size of the observed gap toward HE 0940.*

large gaps, the first estimate shows an apparent underdensity of weaker lines in the region around 4350 Å which is very clearly *not* due to any possible line blanketing.

9.4 Perspectives for a Foreground Quasar

Although the existence of a void has not been fully established, it is still interesting to consider what conditions would be necessary for identifying a foreground quasar responsible for causing it. The idea is plausible because the number density of weaker lines appears to be more strongly affected by quasar radiation than that of stronger lines (section 8.6), and the difference in the aforementioned significance levels for the void candidate could result from precisely such an effect.

If a foreground quasar does indeed exist, a few necessary conditions for observing it come to mind:

- First, the lifetime of the foreground quasar must be at least as great as the light-travel time to (the farthest part of) the void plus the time needed to ionize the Ly α clouds.
- Similarly, if a quasar begins to fade immediately after ionizing the clouds, it must survive at least as long as their recombination time in order to be observed together with the void; and even then it may not be recognized as having been bright enough to cause the void.

- Finally, if the quasar is radiating anisotropically, the energy must be emitted not only in the direction of the void but also towards Earth, so that it is not obscured from our line of sight.

So, assuming that such a quasar is causing an underdensity of lines in the spectrum of HE 0940, how bright or how close to the line of sight must it be? Combining equations (8.3) and (8.20) yields the condition for the foreground quasar's intrinsic Lyman-limit luminosity and its distance r_L from the void:

$$\left(\frac{L_{\nu_{LL}}}{4\pi\, {r_L}^2}\right) \quad \geq \quad 4\pi\, J_{\nu_{LL}} \cdot \omega_{min} \tag{9.2}$$

where ω_{min} indicates the ratio of QSO to UVB Lyman-limit flux densities at which the quasar's sphere of influence is considered to end. Inserting the proper units and writing $J_{\nu_{LL}} = J_{-21} \times 10^{-21}$ erg cm^{-2} s^{-1} Hz^{-1} sr^{-1} shows that the foreground quasar's sphere of influence extends out to a luminosity distance of

$$r_{max} \quad = \quad \left[\frac{1}{1.50\, J_{-21} \cdot \omega_{min}} \left(\frac{L_{\nu_{LL}}}{10^{30}\ \mathrm{erg\ s^{-1}\ Hz^{-1}}}\right)\right]^{1/2} \quad \mathrm{Mpc} \tag{9.3}$$

Bajtlik, Duncan and Ostriker (1988) consider a minimum ratio of $\omega_{min} = 0.1$ and conclude that the influence of a bright quasar ($L_{\nu_{LL}} \approx 3 \times 10^{31}$ erg s^{-1} Hz^{-1}) can affect a line of sight as far as 1° away on the sky. However, it makes more sense to argue that the column density of a typical Ly α forest line ($N_{\mathrm{HI}} \sim 10^{14}$ cm^{-2}) must be reduced by a factor of about 10 to make it undetectable (Dobrzycki and Bechtold 1991b). More reasonable values of $1 \leq \omega_{min} \leq 10$ thus correspond to distances of a few megaparsecs at most.

A few more calculations will show how these distances are related to angular separations on the sky. By definition of the angular size distance d_θ, the angle θ subtended on the sky by the proper length d_P at a distance is d_θ from Earth is $\theta = d_P/d_\theta$, where

$$d_\theta(z_x, z_y) = d_L(z_x, z_y) \cdot \left(\frac{1+z}{1+z_Q}\right)^2 \tag{9.4}$$

The angular size distance from Earth to the quasar is therefore

$$r_\theta \quad = \quad d_\theta(0, z_Q) \quad = \quad \frac{r_Q}{(1+z_Q)^2} \quad = \quad \frac{2c}{H_o}\left[\frac{\sqrt{1+z_Q} - 1}{(1+z_Q)^{3/2}}\right] \tag{9.5}$$

for $\Omega_o = 1$ (Kayser et al. 1997). Assuming that the possible foreground quasar has the same redshift $z_Q \approx 2.580$ as the void candidate, this is a distance of $r_\theta = 790h^{-1}$ Mpc. Because the distance to the foreground quasar is much greater than the distance between the quasar and the void, the luminosity distance r_{max} is basically the same as the proper distance d_P and the angle subtended on the sky becomes

$$\theta \approx \frac{r_{max}}{r_\theta}. \tag{9.6}$$

Equations (9.3), (9.5) and (9.6) can be combined to obtain the direct relation between ω, $L_{\nu_{LL}}$ and θ:

$$L_{\nu_{LL}} = c_1 \ \omega_{min} \, J_{-21} \ \frac{\theta^2 \, {r_{_Q}}^2}{(1+z_{_Q})^4} \tag{9.7}$$

where $c_1 = 1.50 \times 10^{30}$ erg s^{-1} Hz^{-1} Mpc^{-2}. Figure 9.4 makes use of this relation to show the angular extent on the sky of the possible quasar's sphere of influence for different values of ω. Logically and as seen on the scale of the y-axis, such a zone of enhanced ionization is larger for lower estimates of the UVB Lyman-limit intensity. However, roughly speaking, even a bright source would probably have to be within $10'$ (minutes of arc) on the sky, while a faint but closeby UV source could produce the same effect.

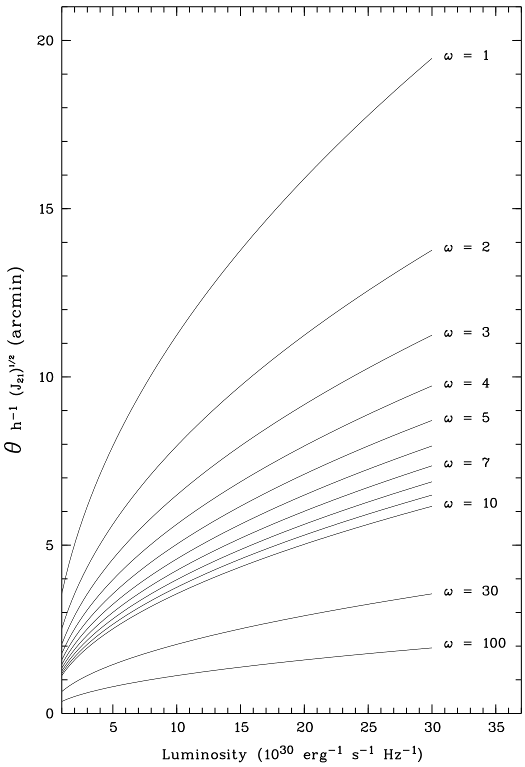


Figure 9.4: *Angular extent of enhanced ionization around a hypothetical foreground quasar with* $z_{_Q} \approx 2.580$.

In order to calculate the range of B magnitudes for the possible foreground quasar, the equations derived in sections 8.2 and 8.3 were first used to obtain the flux observed at $\lambda_{obs} = 4400$ Å for a given Lyman-limit luminosity $L_{\nu_{LL}}$:

$$f_{\lambda_{obs}} = \frac{c}{\lambda_{obs}} \frac{1+z_{_Q}}{4\pi {r_{_Q}}^2} \left[\frac{\lambda_{LL}(1+z_{_Q})}{\lambda_{obs}} \right]^{\alpha} L_{\nu_{LL}} \tag{9.8}$$

Combining this with equation (9.7) and the standard units resulted in the condition

$$f_{\lambda_{obs}} = c_2 \, \frac{{\lambda_{LL}}^{\alpha}}{{\lambda_{obs}}^{2+\alpha}} \, (1 + z_Q)^{\alpha-3} \;\; \omega_{min} \, J_{-21} \, \theta^2 \tag{9.9}$$

where $c_2 = 0.0376$ erg cm^{-2} s^{-1} Å. The B magnitude could then be obtained with the help of the following formula (Weedman 1986):

$$B = -2.5 \log f_{4400} - 20.42 \tag{9.10}$$

where f_{4400} is the flux density in erg cm^{-2} s^{-1} Å^{-1} at $\lambda_{obs} = 4400$ Å.

Assuming a spectral index of $\alpha = -1$ for convenience, Figure 9.5 shows the B magnitude as a function of the angular separation θ on the sky for different values of ω and J_{-21}. For separations of only a few minutes of arc, the magnitude of the foreground quasar is not strongly constrained and even weak sources could ionize the hydrogen strongly along the line of sight to quasar HE 0940–1050.

Obviously, then, it would be worth searching for the quasar that might have caused the void candidate described above (and not only because it would lend credence to the assertion that the feature is probably significant). Of the two confirmed voids in the literature, one foreground quasar has been found at the same redshift as the void, while the other has been found at a higher redshift. As Srianand (1997) has noted, the detection of a quasar with a slightly lower redshift than the observed void would lend support to anisotropic models of quasar radiation. And in the very worst case, even a negative result of the search for a foreground quasar could provide an additional data point for further studies of quasar anisotropy, variability, and lifetimes.

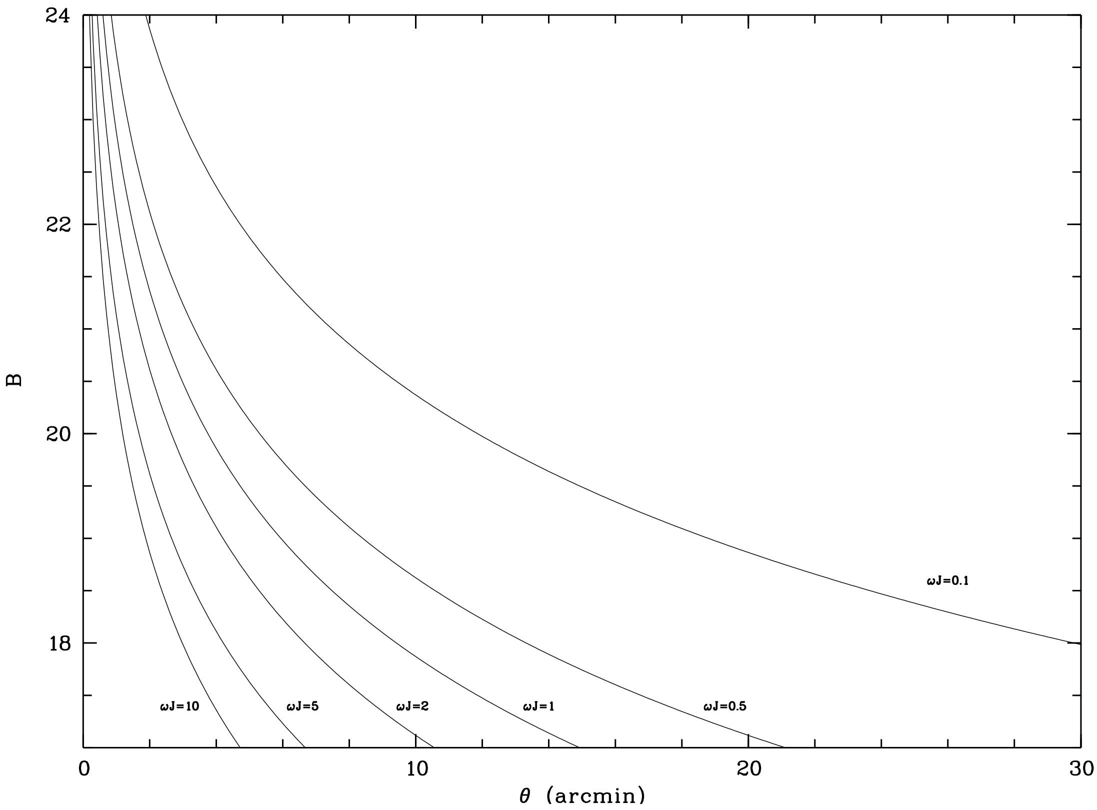


Figure 9.5: *B magnitude vs. angular separation from line of sight to HE 0940–1050 for a foreground quasar with* $z_Q \approx 2.580$.

Chapter 10

Conclusions

The study of the Ly α forest in the spectrum of the $z_{em} = 3.054$ quasar HE 0940–1050 has shown a number of interesting results. After the long reduction process needed to obtain a normalized spectrum for the echelle data, Voigt-profile fits yielded the redshift, column density and Doppler parameter for all of the Ly α lines lines between Ly α and Ly β emission above a $5\,\sigma$ detection limit of $N_{\rm HI} \approx 10^{12.5}$ cm^{-2}. During this process, six heavy-element systems were identified and the metal lines excluded from further consideration. The effects of limited S/N were taken into account, and the column-density distribution was examined in order to determine a completeness limit of $N_{\rm HI} = 10^{13.3}$ cm^{-2} for the final sample that was used to study the proximity effect.

Despite the poor statistics offered by a single quasar spectrum, the redshift distribution of the data shows that the Ly α forest models of Fardal, Giroux and Shull (1998) provide a good description for the full sample except in the very lowest and highest bins. The counts in the latter bin closest to the quasar can thus be explained in terms of the proximity effect, but the former appears to be a statistical overdensity of lines. Neither of these effects disappear when the bin boundaries of locations are chosen differently.

Comparison of the data to Monte Carlo simulations of the Ly α forest in the presence of a quasar seem to confirm the assertion by Bechtold and Dobrzycki (1991b) that the ionizing quasar radiation should appear to affect the weaker lines more drastically than the stronger ones. The drop in the number density is indeed greater for the lines with $\log N_{\rm HI} \leq 13.33$ than for the lines in the sample as a whole, while the stronger lines with $\log N_{\rm HI} \geq 14.0$ show no appreciable proximity effect at all. In fact, the fit of the Ly α forest models is also poor for these column densities, and this may be an indication that the statistics are simply too low here to draw any useful conclusions.

Model I of the intensity of the metagalactic UV background presented by Bajtlik, Duncan and Ostriker (1988) yielded acceptable χ^2 fits to the data, as did the simple assumption that the Lyman-limit flux J_ν is constant with redshift; so in order to obtain a better order-of-magnitude estimate of the UVB by means of the proximity effect, J_ν was varied to find the values that were most compatible with both the full sample and the subsample of weaker lines in the spectrum of HE 0940–1050. This simple test yielded $\log J_\nu = -21.0 \pm 0.7$ and $\log J_\nu = -21.3 \pm 0.7$ for the two samples, respectively. If—as may well be the case—the emission redshift of HE 0940 has been underestimated by 1000–2000 km s^{-1}, then the results for the Lyman-limit background intensity should be reduced by a factor of about 3 to 5.

Finally, a possible void of $3.2h^{-1}$ Mpc (for $\Omega_o = 1$) or $4.9h^{-1}$ Mpc (for $\Omega_o = 0.2$) was identified at $z \approx 2.580$ in the spectrum. Its statistical significance was evaluated by means of Monte Carlo methods. If the column-density cut-off is taken to be equal to the detection limit within the void, the estimated probability of obtaining such a ~ 30 Å redshift interval containing only three lines lies at $(2.0 \pm 0.3) \times 10^{-3}$. If only stronger lines are considered, then the probability of finding a gap of comparable size between two absorption lines rises to 5 or 6%. However, it is not clear whether such a high cut-off is warranted in this case, since line blanketing is not a factor within the void. Therefore it is still worthwhile to undertake a search for the possible foreground quasar, which would offer further evidence that the photoionization model is indeed the correct explanation for the proximity effect. A simple calculation shows that such a quasar would probably be found within a few minutes of arc from the line of sight to quasar HE 0940–1050.

References

[1] Bajtlick, S., Duncan, R. C., Ostriker, J. P. 1988. Quasar ionization of Lyman-alpha clouds: The proximity effect, a probe of the ultraviolet background at high redshift. *Ap. J.* **327**, 570.

[2] Barcons, X., and Fabian, A. C. 1987. Cosmological implications of the anomalies in the redshift distribution of QSO Lyman α absorption systems. *Mon. Not. R. Astr. Soc.* **229**, 157.

[3] Bechtold, J. 1993. The first QSOs and their influence on the intergalactic medium. In *The Environment and Evolution of Galaxies*, eds. J. M. Shull and H. A. Thronson, Jr. (Dordrecht: Kluwer Academic Publishers), 559.

[4] Bechtold, J. 1994. The Lyman-alpha forest near 34 quasi-stellar objects with $z > 2.6$. *Ap. J. S.* **91**, 1.

[5] Bechtold, J. 1995. The Proximity Effect. In *QSO Absorption Lines*, Proc. ESO Workshop held November 21–24, 1994, ed. G. Meylan (Berlin: Springer-Verlag), 299.

[6] Bevington, P. R., and Robinson, D. K. 1992. *Data Reduction and Error Analysis for the Physical Sciences*. 2nd ed. New York: McGraw-Hill, Inc.

[7] Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A., and Tytler, D. 1982. Observations of the spectra Q0122–380 and Q1101–264. *Mon. Not. R. Astr. Soc.* **198**, 110.

[8] Caulet, A. The evolution of low-ionization QSO absorption systems. *Ap. J.* **340**, 90.

[9] Chernomordik, V. V., and Ozernoy, L. M. 1993. The quasar proximity effect in an equivalent-width-limited sample of the Lyman-alpha forest. *Ap. J.* **404**, L5.

[10] Cowie, L. L., Songaila, A., Kim, T.-S., and Hu, E. M. 1995. The metallicity and internal structure of the Lyman-alpha forest clouds. *Astron. J.* **109**, 1522.

[11] Crotts, A. P. S. 1989. Spatial structure in the Lyman-alpha forest. *Ap. J.* **336**, 550.

[12] Dobrzycki, A., and Bechtold, J. 1991a. A ~ 10 Mpc void in the Lyα forest at $z = 3.17$. *Ap. J.* **377**, L69.

[13] Dobrzycki, A., and Bechtold, J. 1991b. The influence of quasars on the Lyman-α forest. In *The Space Distribution of Quasars*, ASP Conference Series, vol. 21, ed. D. Crampton, 272.

[14] Dobrzycki, A., and Bechtold, J. 1996. Simulation analysis of Lyα forest spectra. I. Empirical description at $z \approx 3$. *Ap. J.* **457**, 102.

[15] D'Odorico, S., Ghigo, M., Ponz, D. 1987. An Atlas of the Thorium-Argon Spectrum for the ESO Echelle Spectrograph in the $\lambda\lambda$ 3400–9000 Å Region. ESO Scientific Report No. 6, March 1987. Garching bei München: European Southern Observatory.

[16] Duncan, R. C., Ostriker, J. P., and Bajtlik, S. 1989. Voids in the Lyα forest. *Ap. J.* **345**, 39.

[17] Espey, B. R., Carswell, R. F., Bailey, J. A., Smith, M. G., and Ward, M. J. 1989. Hα emission lines in high-redshift quasars. *Ap. J.* **342**, 666.

[18] Espey, B. R. 1993. The influence of QSO emission-line velocity shifts on estimates of the intergalactic background intensity. *Ap. J.* **411**, L59.

[19] European Southern Observatory. 1995. ESO-MIDAS User Guide. Volume A: System. Reference number MID-MAN-ESO-11000-0002.

[20] European Southern Observatory. 1995. ESO-MIDAS User Guide. Volume B: Data Reduction. Reference number MID-MAN-ESO-11000-0003.

[21] Evans, A., and Hart, D. 1977. *Astron. Astrophys.* **58**, 241.

[22] Fang, Y., Fan, X., Tytler, D., and Crotts, A. P. S. 1998. Remeasurement of the H I Gunn-Peterson Effect toward QSO PKS 1937–101 with Keck observations. *Ap. J.* **497**, 67.

[23] Fardal, M. A., Giroux, M. L., and Shull, J.M. 1998. The high-redshift He II Gunn-Peterson effect: Implications and future prospects. *Astron. J.* **115** (6), 2206.

[24] Fernández-Soto, A., Barcons, X., Carballo, R., and Webb, J. K. 1995. The proximity effect on the Lyman α forest due to a foreground QSO. *Mon. Not. R. Astr. Soc.* **277**, 235.

[25] Fontana, A., and Ballester, P. 1995. FITLYMAN: A Midas Tool for the Analysis of Absorption Spectra. *ESO Messenger* **80**, 37.

[26] Gardner, J. P., Katz, N., Hernquist, L., and Weinberg, D. H. 1997. The population of damped Lyman-alpha and Lyman limit systems in the cold dark matter model. *Ap. J.* **484**, 31.

[27] Gaskell, C. M. 1982. A redshift difference between high and low ionization emission-line regions in QSOs—Evidence for radial motions. *Ap. J.* **263**, 79.

[28] Giallongo, E., Cristiani, S., D'Odorico, S., Fontana, A., and Savaglio, S. 1996. The proximity effect, the UV background and the statistics of the Lyman-alpha lines at high resolution. *Ap. J.* **466**, 46.

[29] Giallongo, E., Cristiani, S., D'Odorico, S., Fontana, A., and Savaglio, S. 1997. The UV background and the Ly α clouds at high redshift. In *Structure and Evolution of the Intergalactic Medium from QSO Absorption Lines*, Proc. 13th IAP Astrophysics Colloquium. Paris: Editions Frontières.

[30] Haardt, F., and Madau, P. 1996. Radiative transfer in a clumpy universe. II. The ultraviolet extragalactic background. *Ap. J.* **461**, 20.

[31] Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A., and Rauch, M. 1995. The distribution of column densities and b values in the Lyman-alpha forest. *Ap. J.* **110**, 1526.

[32] James, F. 1991. Minuit – Function minimization and error analysis. Program Library Long Writeup D506. Geneva: CERN.

[33] Kayser, R., Helbig, P., and Schramm, T. 1997. A general and practical method for calculating cosmological distances. *Astron. Astrophys.* **318** (3), 680.

[34] Kim, T.-S., Hu, E. M., Cowie, L. L., and Songaila, A. 1997. The redshift evolution of the Lyα forest. *Astron. J.* **114** (1), 1.

[35] Kirkman, D., and Tytler, D. 1997. Intrinsic properties of the $\langle z \rangle = 2.7$ Lyman alpha forest from Keck spectra of QSO HS 1946+7658. *Ap. J.* **484**, 672

[36] Kitchin, C. R. 1991. *Astrophysical Techniques.* 2nd ed. Bristol: IOP Publishing Ltd.

[37] Kitchin, C. R. 1995. *Optical Astronomical Spectroscopy.* Bristol: IOP Publishing Ltd.

[38] Kovner, I., and Rees, M. J. 1989. Clearings in Lyα forests. *Ap. J.* **345**, 52.

[39] Lang, K. R. 1980. *Astrophysical Formulae.* 2nd ed. Berlin: Springer-Verlag.

[40] Lu, L., Wolfe, A. M., and Turnshek, D. A. 1991. The redshift distribution of Lyα clouds and the proximity effect. *Ap. J.* **367**, 19.

[41] Lu, L., Sargent, W. L. W., Womble, D. S., and Takada-Hidai, M. 1996. The Lyα Forest at $z \sim 4$: Keck HIRES Observations of Q000–26. *Ap. J.* **472**, 509.

[42] Lynds, R. 1971. The absorption-line spectrum of 4C 05.34. *Ap. J.* **164**, L73.

[43] Lyons, L. 1986. *Statistics for Nuclear and Particle Physics.* Cambridge: Cambridge University Press.

[44] Madau, P., and Shull, J. M. 1996. Cosmic metal production and the contribution of QSO absorption systems to the ionizing background. *Ap. J.* **457**, 551.

[45] Mar, D. 1994. Xvoigt User's Guide. Version 3.0. School of Physics, University of Sydney.

[46] McLean, I. S. 1997. *Electronic Imaging in Astronomy: Detectors and Instrumentation.* Chichester: John Wiley & Sons.

[47] Murdoch, H. S., Hunstead, R. W., Pettini, M., and Blades, J. C. 1986. Absorption spectrum of the $z = 3.78$ QSO 2000–330. II. The redshift and equivalent width distributions of primordial hydrogen clouds. *Ap. J.* **309**, 19.

[48] Ostriker, J. P., Bajtlik, S., Duncan, R. C. 1988. Clustering and voids in the Ly α forest. *Ap. J.* **327**, L35.

[49] Pasquini, P. 1993. CASPEC Update to the Operating Manual, September 1993. European Southern Observatory.

[50] Peebles, P. J. E. 1993. *Principles of Physical Cosmology*. Princeton: Princeton University Press.

[51] Pettini, M., Hunstead, R. W., Smith, L., and Mar, D. P. 1990. The Lyman α forest at 6 km s^{-1} resolution. *Mon. Not. R. Astr. Soc.* **246**, 545.

[52] Press, W. P., Teukolsky, S. A., Vetterling, W. T., and Flannery ,B. P. 1992. *Numerical Recipes in C. The Art of Scientific Computing.* 2nd ed. Cambridge: Cambridge University Press.

[53] Randich, S. and Pasquini, L. CASPEC Operating Manual, December 1996. 3p6-MAN-0633-0002. European Southern Observatory.

[54] Rauch, M., Carswell, R. F., Chaffee, F. H., Foltz, C. B., Webb, J. K., Weymann, R. J., Bechtold, J., and Green, R. F. 1992. The Lyman forest of 0014+813. *Ap. J.* **390**, 387.

[55] Rauch, M., Carswell, R. F., Webb, J. K., and Weymann, R. J. 1993. A reanalysis of the spectrum of QSO 2206 – 199. *Mon. Not. R. Astr. Soc.* **260**, 589.

[56] Rauch, M. 1998. The Lyman alpha forest in the spectra of QSOs. Preprint astro/ph-9806286. *Ann. Rev. Astron. Astrophys.* **36** (in press).

[57] Reimers, D., Köhler, T., and Wisotzki, L. 1996. The Hamburg/ESO survey for bright QSOs. II. Follow-up spectroscopy of 160 quasars and Seyferts. *Astron. Astrophys. Suppl. Ser.* **115**, 235.

[58] Reimers, D., and Wisotzki, L. 1997. The Hamburg/ESO survey. *ESO Messenger* **88**, 14.

[59] Smette, A., Robertson, J. G., Shaver, P. A., Reimers, D., Wisotzki, L., and Köhler, T. 1995. The gravitational lens candidate HE 1104–1805 and the size of absorption systems. *Astron. Astrophys. Suppl.* **113**, 199.

[60] Spitzer, L. 1978. *Physical Processes in the Interstellar Medium*. New York: Wiley & Sons.

[61] Srianand, R. 1997. Radiation-induced void in the spectrum of Tol 1038–2712. *Ap. J.* **478**, 511.

[62] Stengler-Larrea, E. A. et al. 1995. The Hubble Space Telescope quasar absorption line key project. V. Redshift evolution of Lyman limit absorption in the spectra of a large sample of quasars. *Ap. J.* **444**, 64.

[63] Tytler, D. 1987. The redshift distribution of QSO Lyman-alpha absorption systems. *Ap. J.* **321**, 69.

[64] Tytler, D., Fan, X.-M., Burles, S., Cottrell, L., Davis, C., Kirkman, D, and Zuo, L. 1995. Ionization and abundances of intergalactic gas. In *QSO Absorption Lines*, Proc. ESO Workshop held November 21–24, 1994, ed. G. Meylan (Berlin: Springer-Verlag).

[65] Tytler, D. 1997. An overview and introduction to quasar absorption in cosmology. In *Structure and Evolution of the Intergalactic Medium from QSO Absorption Lines*, Proc. 13th IAP Astrophysics Colloquium. Paris: Editions Frontières.

[66] Unsöld, A., and Baschek, B. 1991. *Der neue Kosmos*. 5th ed. Berlin: Springer-Verlag.

[67] Verner, D. A., Barthel, P. D., and Tytler, D. 1994. Atomic data for absorption lines from the ground level at wavelengths greater than 228 Å. *Astron. Astrophys. Suppl. Ser.* **108**, 287.

[68] Weedman, D. W. *Quasar Astronomy*. 1986. Cambridge: Cambridge University Press.

[69] Weymann, R. J., Carswell, R. F., and Smith, M. J. 1981. Absorption lines in the spectra of quasi-stellar objects. *Ann. Rev. Astron. Astrophys.* **19**, 41.

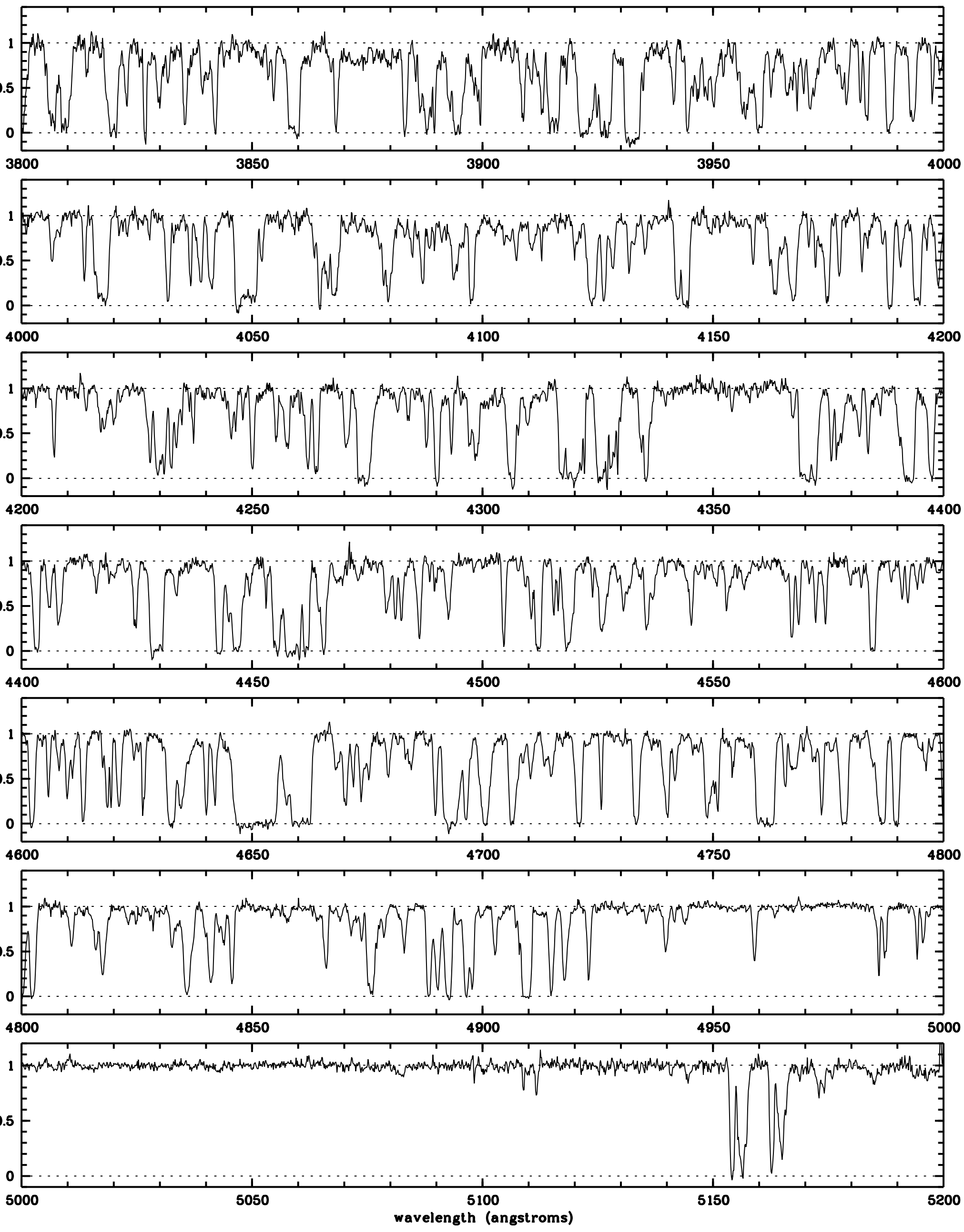
[70] Wisotzki, L. 1994. CCD-Datenreduktion und Bildverarbeitung mit MIDAS. Astronomisches Praktikum an der Hamburger Sternwarte 1994/95. Laboratory Manual.

[71] Wisotzki, L., Köhler, T., Groote, D., and Reimers, D. 1996. The Hamburg/ESO survey for bright QSOs. I. Survey design and candidate selection procedure. *Astron. Astrophys. Suppl. Ser.* **115**, 227.

Parts of this research made use of the article and abstract services of NASA's Astrophysics Data System.

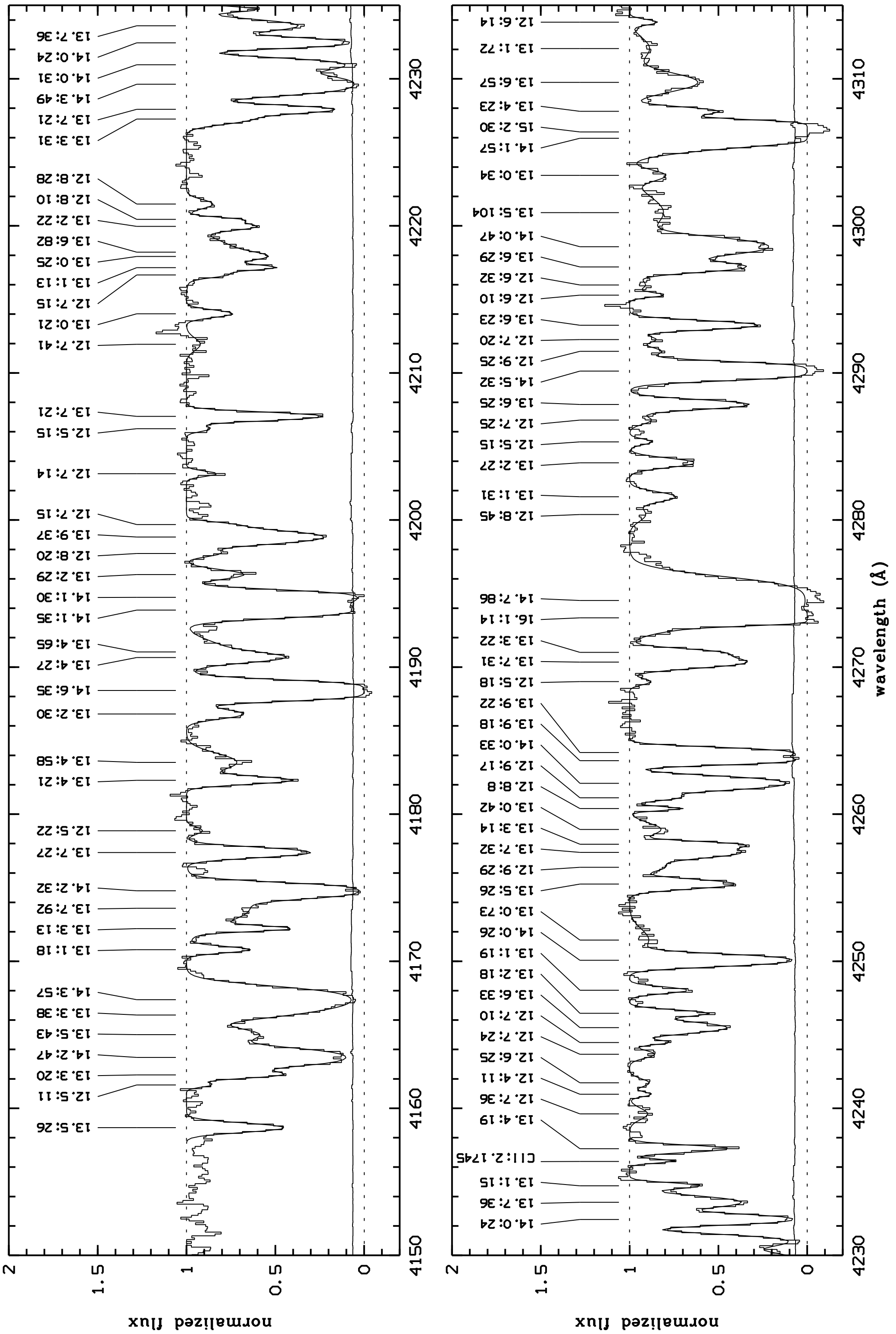
Appendix A

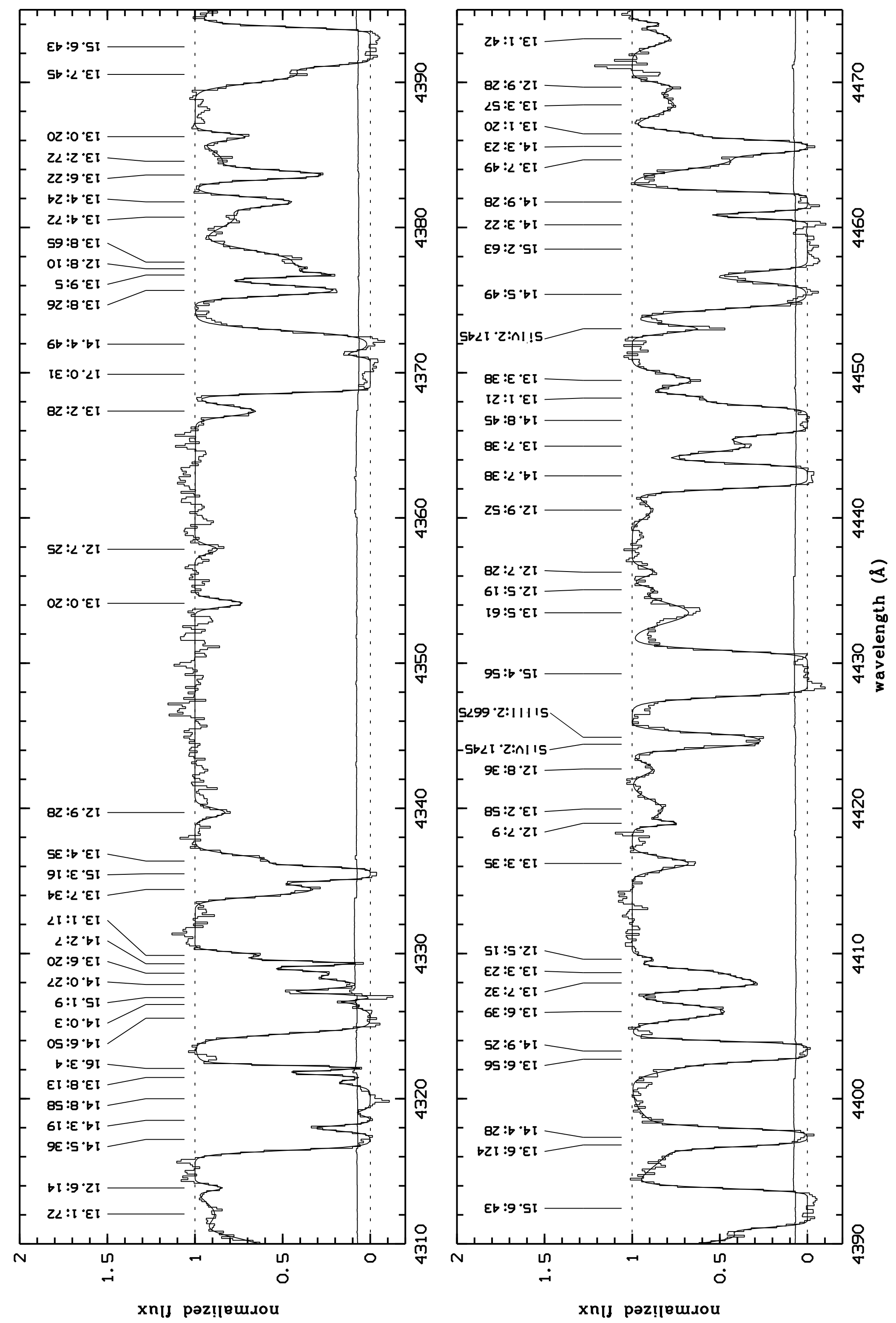
Normalized spectrum

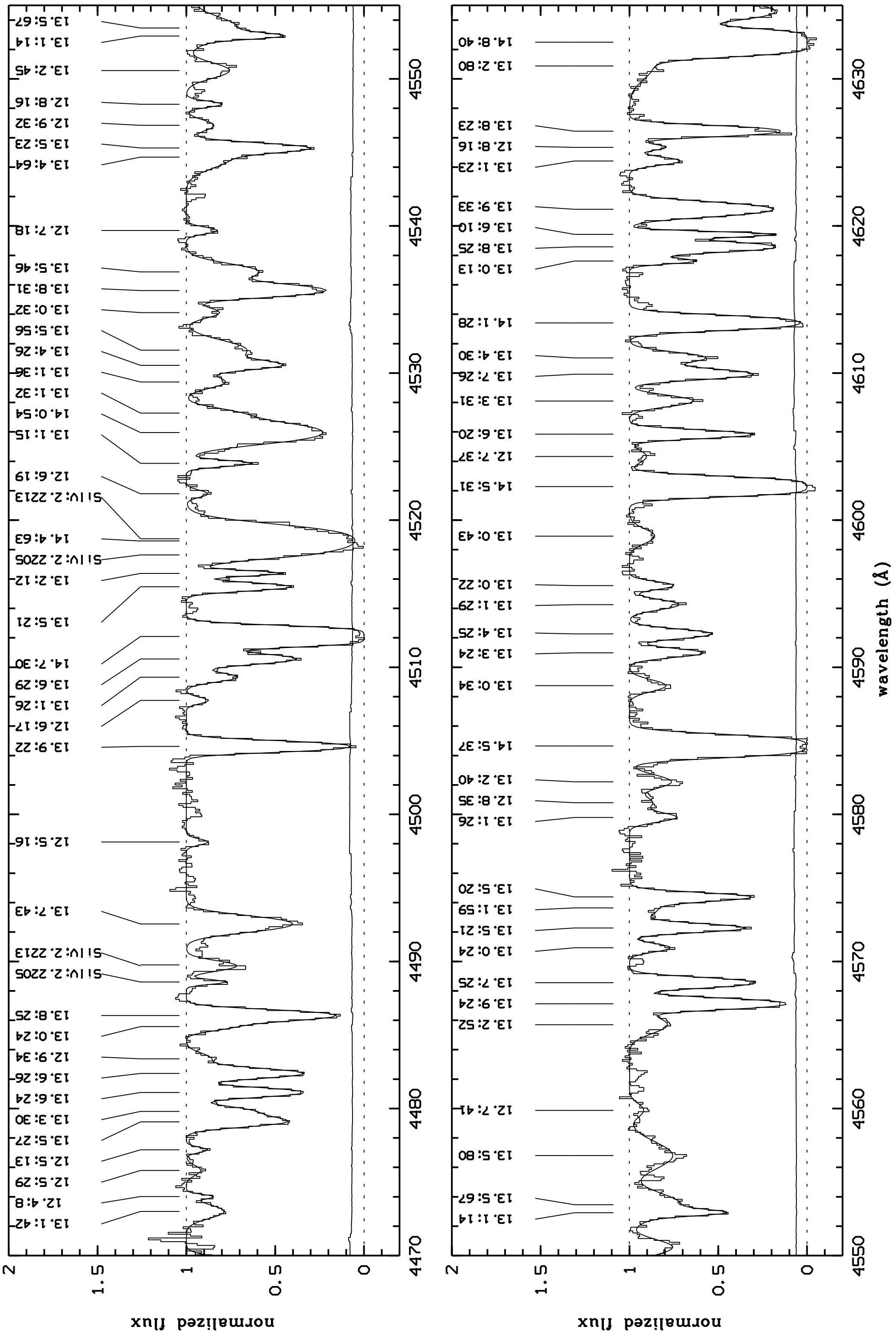


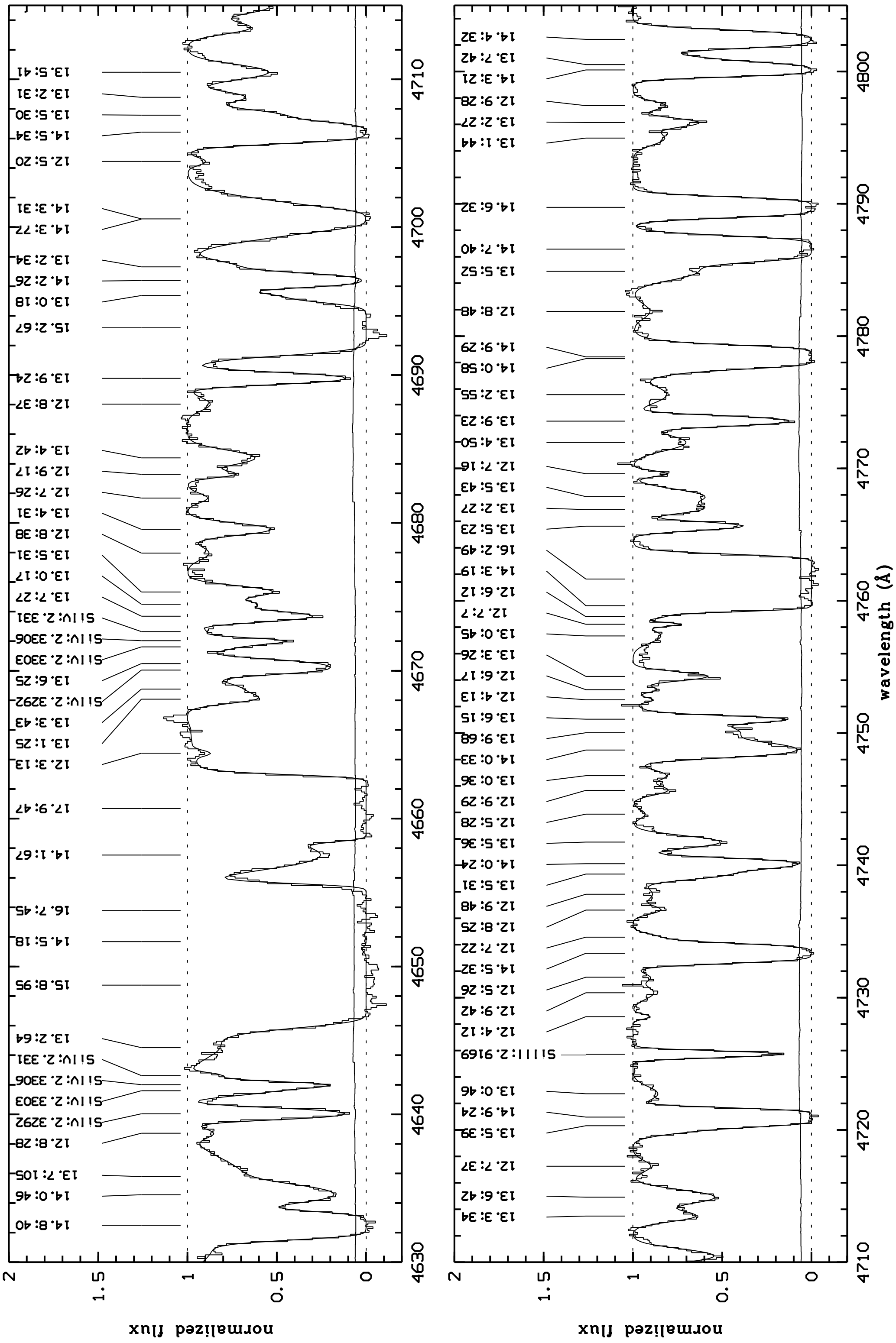
Appendix B

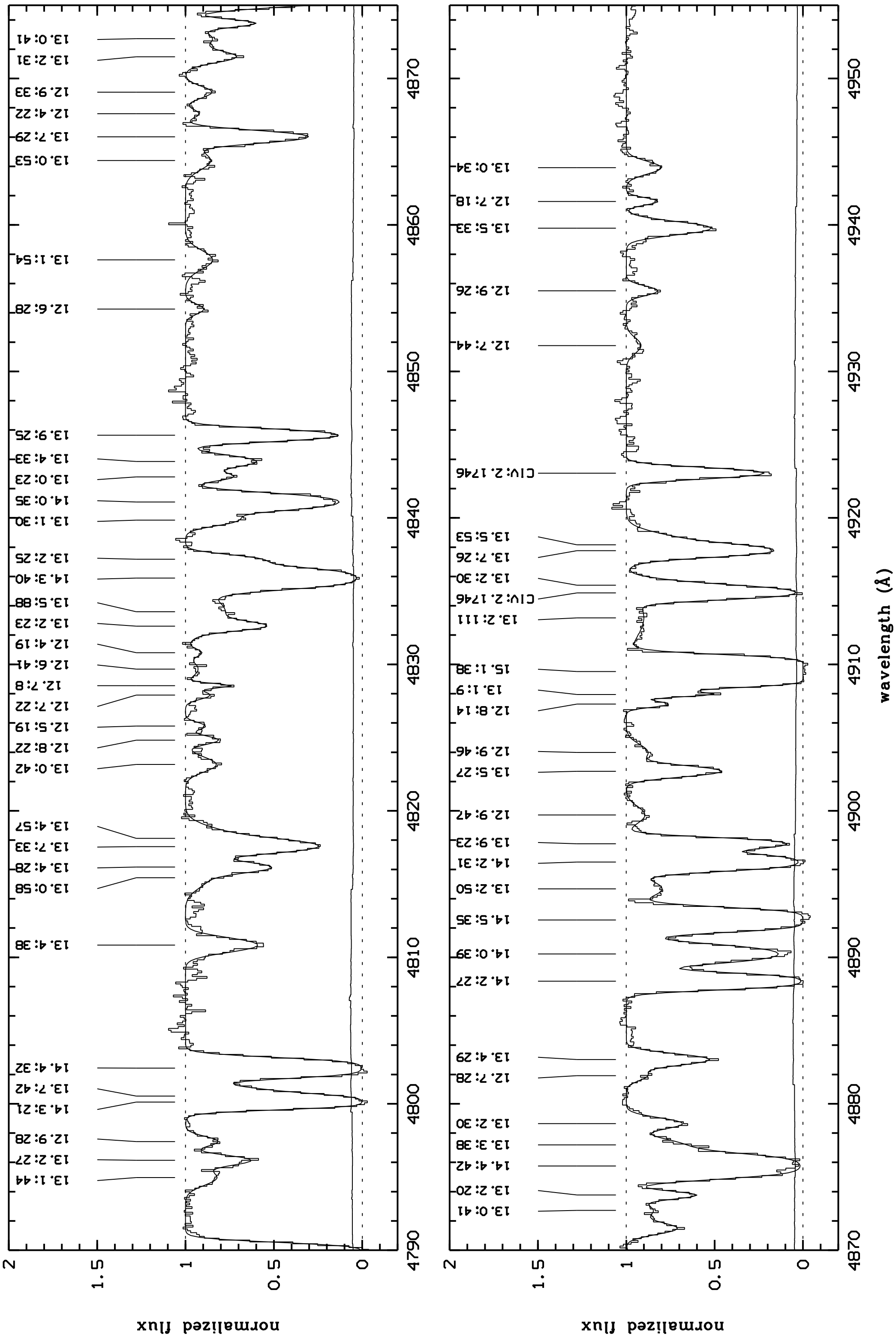
Line Fits


normalized flux
wavelength (Å)
2
1.5
1
0.5
0
4150
4160
4170
4180
4190
4200
4210
4220
4230
4240
4250
4260
4270
4280
4290
4300
4310
13.7:36
14.0:24
14.0:31
14.3:49
13.7:21
13.3:31
12.8:28
12.8:10
13.2:22
13.6:82
13.0:25
13.1:13
12.7:15
13.0:21
12.7:41
13.7:21
12.5:15
12.7:14
12.7:15
13.9:37
12.8:20
13.2:29
14.1:30
14.1:35
13.4:65
13.4:27
14.6:35
13.2:30
13.4:58
13.4:21
12.5:22
13.7:27
14.2:32
13.7:92
13.3:13
13.1:18
14.3:57
13.3:38
13.5:43
14.2:47
13.3:20
12.5:11
13.5:26
12.6:14
13.1:72
13.6:57
13.4:23
15.2:30
14.1:57
13.0:34
13.5:104
14.0:47
13.6:29
12.6:32
12.6:10
13.6:23
12.7:20
12.9:25
14.5:32
13.6:25
12.7:25
12.5:15
13.2:27
13.1:31
12.8:45
14.7:86
16.1:14
13.3:22
13.7:31
12.5:18
13.9:22
13.9:18
14.0:33
12.9:17
12.8:8
13.0:42
13.3:14
13.7:32
12.9:29
13.5:26
13.0:73
14.0:26
13.1:19
13.2:18
13.6:33
12.7:10
12.7:24
12.6:25
12.4:11
12.7:36
13.4:19
CII:2.1745
13.1:15
13.7:36
14.0:24


normalized flux
wavelength (Å)
2
1.5
1
0.5
0
4310
4320
4330
4340
4350
4360
4370
4380
4390
4400
4410
4420
4430
4440
4450
4460
4470
13.1:72
12.6:14
14.5:36
14.3:19
14.8:58
13.8:13
16.3:4
14.6:50
14.0:3
15.1:9
14.0:27
13.6:20
14.2:7
13.1:17
13.7:34
15.3:16
13.4:35
12.9:28
13.0:20
12.7:25
13.2:28
17.0:31
14.4:49
13.8:26
13.9:5
12.8:10
13.8:65
13.4:72
13.4:24
13.6:22
13.2:72
13.0:20
13.7:45
15.6:43
15.6:43
13.6:124
14.4:28
13.6:56
14.9:25
13.6:39
13.7:32
13.3:23
12.5:15
13.3:35
12.7:9
13.2:58
12.8:36
SiIV:2.1745
SiIII:2.6675
15.4:56
13.5:61
12.5:19
12.7:28
12.9:52
14.7:38
13.7:38
14.8:45
13.1:21
13.3:38
SiIV:2.1745
14.5:49
15.2:63
14.3:22
14.9:28
13.7:49
14.3:23
13.1:20
13.3:57
12.9:28
13.1:42


normalized flux
normalized flux
wavelength (Å)


normalized flux
wavelength (Å)


normalized flux
wavelength (Å)

Appendix C

Line Lists

C.1 Fitted lines below Ly β emission

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
1	3761.88	0.04	2.667540	15.21	0.12	63.1	5.0	H I 1025
2	3763.29	0.10	2.668916	14.34	0.20	22.3	5.6	H I 1025
3	3764.61	0.02	2.670202	14.91	0.18	28.1	2.3	H I 1025
4	3767.06	0.23	2.672595	13.74	0.19	48.8	18.3	H I 1025
5	3767.84	0.04	2.673354	14.35	0.11	22.9	3.8	H I 1025
6	3768.57	0.09	2.674065	13.06	0.18	19.9	9.2	H I 1025
7	3859.09	0.02	2.174456	18.30	0.06	29.7	1.0	H I 1215
8	3885.55	0.01	2.220513	12.64	0.12	29.2	12.0	Si III 1206
9	3886.50	0.02	2.221298	13.13	0.15	27.9	5.0	Si III 1206
10	3886.81	0.02	2.221561	12.67	0.58	38.2	25.4	Si III 1206
11	3914.94	0.05	2.220397	14.38	0.14	54.6	7.1	H I 1215
12	3915.08	0.01	2.220513	16.13	5.51	17.3	17.6	H I 1215
13	3916.04	0.02	2.221298	13.86	0.34	54.2	19.3	H I 1215
14	3916.36	0.02	2.221561	13.88	0.22	21.9	10.1	H I 1215
15	3922.38	0.04	2.824017	15.76	0.08	94.6	4.5	H I 1025
16	3924.86	0.04	2.826435	14.47	0.08	17.9	4.6	H I 1025
17	3926.63	0.03	2.828157	16.66	0.45	45.3	4.9	H I 1025
18	3929.80	0.06	2.831252	14.07	0.03	66.5	6.3	H I 1025
19	3932.45	0.01	2.833839	17.88	0.15	47.2	1.2	H I 1025
20	3955.31	0.06	2.253607	13.19	0.11	21.1	6.5	H I 1215
21	3955.53	0.12	2.856337	12.84	0.12	37.3	11.5	H I 1025
22	3956.35	0.06	2.254457	13.84	0.07	33.6	7.9	H I 1215
23	3957.00	0.08	2.254998	13.48	0.32	10.8	9.4	H I 1215
24	3957.00	0.01	2.857773	13.89	0.04	24.0	1.7	H I 1025
25	3957.49	0.10	2.255396	13.61	0.33	24.4	22.9	H I 1215
26	3958.22	0.18	2.255997	13.47	0.23	27.4	11.9	H I 1215
27	3959.89	0.02	2.860584	15.22	0.04	67.4	1.8	H I 1025
28	3961.72	0.07	2.862375	13.04	0.20	17.7	8.9	H I 1025
29	3962.58	0.03	2.863210	14.15	0.06	26.0	4.2	H I 1025
30	3963.03	0.20	2.259956	13.28	0.12	61.3	16.1	H I 1215
31	3963.36	0.31	2.863969	13.18	0.36	33.8	21.3	H I 1025
32	3966.08	0.02	2.866626	14.28	0.06	77.4	4.9	H I 1025
33	3966.10	0.04	2.866642	14.25	0.09	30.5	4.5	H I 1025
34	3982.77	0.41	2.882890	13.55	0.47	39.0	17.5	H I 1025
35	3983.32	0.02	2.883430	14.87	0.10	24.2	1.8	H I 1025
36	4015.94	0.04	2.915227	14.34	0.06	18.8	2.5	H I 1025
37	4017.63	0.01	2.916881	16.25	0.18	48.9	2.3	H I 1025
38	4031.69	0.06	2.930590	14.00	0.21	57.6	9.2	H I 1025
39	4031.81	0.02	2.930699	14.88	0.06	28.9	2.1	H I 1025
40	4142.40	0.01	3.038523	15.11	0.05	37.6	0.9	H I 1025
41	4144.08	0.01	2.408885	17.51	0.12	19.4	0.6	H I 1215

C.2 All lines between Ly α and Ly β emission

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
1	4158.68	0.02	2.420897	13.47	0.03	25.7	2.4	H I 1215
2	4161.59	0.09	2.423286	12.49	0.22	10.7	12.3	H I 1215
3	4162.27	0.04	2.423852	13.32	0.09	19.9	5.2	H I 1215
4	4163.46	0.03	2.424826	14.16	0.03	46.5	4.9	H I 1215
5	4165.01	0.09	2.426099	13.49	0.13	43.2	15.9	H I 1215
6	4166.34	0.27	2.427199	13.31	0.49	38.0	23.3	H I 1215
7	4167.38	0.10	2.428051	14.30	0.06	57.3	5.8	H I 1215
8	4170.79	0.04	2.430855	13.09	0.07	18.0	4.0	H I 1215
9	4172.22	0.02	2.432031	13.27	0.08	13.4	3.6	H I 1215
10	4173.60	0.23	2.433167	13.73	0.09	92.5	17.6	H I 1215
11	4174.79	0.02	2.434148	14.18	0.06	32.0	2.9	H I 1215
12	4177.39	0.02	2.436286	13.67	0.03	26.9	1.8	H I 1215
13	4178.87	0.14	2.437504	12.50	0.20	22.5	14.5	H I 1215
14	4182.30	0.02	2.440321	13.44	0.05	20.9	2.6	H I 1215
15	4183.52	0.09	2.441330	13.41	0.06	57.9	9.4	H I 1215
16	4186.81	0.04	2.444036	13.21	0.05	29.5	5.2	H I 1215
17	4188.41	0.01	2.445348	14.55	0.09	35.0	2.2	H I 1215
18	4190.65	0.06	2.447191	13.36	0.21	27.2	8.3	H I 1215
19	4191.03	0.26	2.447504	13.39	0.20	65.3	14.2	H I 1215
20	4193.87	0.11	2.449846	14.13	0.12	34.7	5.3	H I 1215
21	4194.74	0.10	2.450562	14.13	0.13	30.2	5.0	H I 1215
22	4196.29	0.05	2.451833	13.21	0.06	29.3	5.2	H I 1215
23	4197.72	0.08	2.453013	12.76	0.16	19.7	10.1	H I 1215
24	4198.84	0.04	2.453931	13.88	0.04	37.0	4.6	H I 1215
25	4199.69	0.12	2.454633	12.74	0.34	15.4	12.1	H I 1215
26	4203.15	0.07	2.457474	12.66	0.14	14.3	8.4	H I 1215
27	4206.21	0.10	2.459991	12.55	0.19	14.7	10.6	H I 1215
28	4207.04	0.02	2.460679	13.67	0.03	20.9	1.9	H I 1215
29	4211.95	0.21	2.464711	12.66	0.19	41.0	21.9	H I 1215
30	4214.03	0.04	2.466423	12.99	0.07	21.4	4.4	H I 1215
31	4216.67	0.16	2.468593	12.70	0.53	14.7	19.1	H I 1215
32	4217.16	0.06	2.469002	13.11	0.20	13.2	7.3	H I 1215
33	4217.92	0.10	2.469623	13.04	0.55	24.9	17.7	H I 1215
34	4218.21	0.53	2.469864	13.56	0.26	82.1	25.4	H I 1215
35	4219.97	0.09	2.471312	13.19	0.19	22.2	11.2	H I 1215
36	4220.45	0.09	2.471704	12.79	0.34	10.4	8.4	H I 1215
37	4221.48	0.09	2.472557	12.81	0.11	28.0	9.6	H I 1215
38	4227.27	0.19	2.477315	13.30	0.21	31.3	12.9	H I 1215
39	4227.92	0.04	2.477854	13.74	0.09	20.6	3.4	H I 1215
40	4229.63	0.06	2.479257	14.28	0.04	48.6	4.6	H I 1215
41	4230.95	0.06	2.480347	13.98	0.06	31.4	4.5	H I 1215
42	4232.44	0.02	2.481572	13.96	0.05	24.3	2.4	H I 1215
43	4233.61	0.04	2.482529	13.69	0.04	35.7	5.2	H I 1215
44	4234.73	0.04	2.483455	13.11	0.08	15.2	4.9	H I 1215
45	4236.41	0.02	2.174456	13.21	0.11	7.9	5.7	C II 1334

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
46	4237.25	0.02	2.485528	13.38	0.04	18.8	2.4	H I 1215
47	4239.62	0.16	2.487478	12.73	0.15	36.4	12.7	H I 1215
48	4240.95	0.10	2.488573	12.44	0.22	11.4	11.4	H I 1215
49	4241.73	0.14	2.489211	12.63	0.19	25.2	15.9	H I 1215
50	4243.67	0.10	2.490811	12.73	0.14	24.4	10.2	H I 1215
51	4244.48	0.05	2.491470	12.70	0.13	10.0	7.1	H I 1215
52	4245.48	0.04	2.492296	13.57	0.04	33.4	4.2	H I 1215
53	4246.46	0.03	2.493105	13.21	0.07	17.8	4.1	H I 1215
54	4248.02	0.04	2.494389	13.06	0.07	18.5	5.2	H I 1215
55	4250.06	0.02	2.496066	13.98	0.04	26.5	1.9	H I 1215
56	4251.44	0.35	2.497200	13.05	0.16	73.2	27.3	H I 1215
57	4255.25	0.03	2.500333	13.54	0.04	26.4	2.8	H I 1215
58	4256.38	0.19	2.501266	12.94	0.27	28.6	19.2	H I 1215
59	4257.40	0.10	2.502101	13.65	0.13	32.3	10.6	H I 1215
60	4257.95	0.06	2.502556	13.26	0.26	13.9	5.7	H I 1215
61	4258.96	0.12	2.503383	13.04	0.11	41.8	15.1	H I 1215
62	4260.38	0.03	2.504551	12.82	0.11	7.7	4.7	H I 1215
63	4261.11	0.07	2.505153	12.90	0.15	16.7	8.1	H I 1215
64	4262.10	0.03	2.505964	13.99	0.04	32.7	2.7	H I 1215
65	4263.63	0.10	2.507228	13.85	0.20	17.8	5.2	H I 1215
66	4264.19	0.11	2.507684	13.91	0.17	21.7	5.6	H I 1215
67	4269.01	0.12	2.511652	12.54	0.20	17.7	11.5	H I 1215
68	4270.33	0.13	2.512742	13.67	0.13	31.4	8.0	H I 1215
69	4271.00	0.15	2.513290	13.29	0.31	21.5	8.3	H I 1215
70	4273.35	0.13	2.515222	16.15	2.12	14.1	5.7	H I 1215
71	4274.52	0.05	2.516187	14.70	0.05	86.0	3.5	H I 1215
72	4280.37	0.31	2.520995	12.76	0.27	44.5	34.4	H I 1215
73	4281.58	0.09	2.521990	13.11	0.11	30.7	8.6	H I 1215
74	4283.89	0.04	2.523890	13.25	0.05	27.3	4.0	H I 1215
75	4285.32	0.10	2.525065	12.53	0.20	15.5	12.6	H I 1215
76	4286.79	0.16	2.526281	12.67	0.20	25.0	16.5	H I 1215
77	4287.85	0.03	2.527152	13.62	0.04	25.5	2.8	H I 1215
78	4290.13	0.02	2.529025	14.45	0.10	31.5	2.6	H I 1215
79	4291.47	0.12	2.530127	12.85	0.18	24.7	13.8	H I 1215
80	4292.27	0.13	2.530787	12.69	0.25	20.3	17.4	H I 1215
81	4293.24	0.02	2.531585	13.64	0.04	22.9	2.5	H I 1215
82	4295.29	0.07	2.533273	12.63	0.24	10.3	9.6	H I 1215
83	4295.98	0.30	2.533841	12.63	0.33	32.4	30.5	H I 1215
84	4297.21	0.04	2.534845	13.63	0.06	28.8	4.6	H I 1215
85	4298.58	0.04	2.535978	13.99	0.04	47.2	4.7	H I 1215
86	4300.90	0.24	2.537886	13.47	0.10	104.4	29.5	H I 1215
87	4303.44	0.10	2.539972	13.01	0.12	33.8	9.9	H I 1215
88	4305.96	0.42	2.542043	14.07	0.44	56.8	14.1	H I 1215
89	4306.38	0.07	2.542395	15.21	0.94	30.1	11.4	H I 1215
90	4307.79	0.06	2.543550	13.37	0.10	23.3	6.1	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
91	4309.76	0.07	2.545170	13.57	0.05	56.9	8.0	H I 1215
92	4312.06	0.29	2.547062	13.06	0.15	72.0	29.2	H I 1215
93	4313.85	0.08	2.548538	12.59	0.17	14.2	8.4	H I 1215
94	4317.20	0.04	2.551290	14.49	0.12	35.7	3.9	H I 1215
95	4318.51	0.10	2.552371	14.32	0.53	19.3	10.9	H I 1215
96	4320.00	0.13	2.553597	14.79	0.19	58.1	20.9	H I 1215
97	4321.46	0.05	2.554799	13.77	0.20	12.9	5.7	H I 1215
98	4322.08	0.01	2.555308	16.31	0.32	3.6	0.6	H I 1215
99	4325.54	0.07	2.558152	14.64	0.09	50.1	3.5	H I 1215
100	4326.48	0.29	2.558929	14.01	1.91	3.1	11.9	H I 1215
101	4326.97	0.03	2.559327	15.06	0.74	9.2	1.6	H I 1215
102	4327.86	0.04	2.560064	13.98	0.06	26.6	5.2	H I 1215
103	4328.64	0.05	2.560703	13.62	0.09	19.8	5.1	H I 1215
104	4329.29	0.03	2.561241	14.25	0.51	6.9	1.8	H I 1215
105	4329.87	0.05	2.561713	13.09	0.08	16.8	5.2	H I 1215
106	4334.41	0.05	2.565452	13.72	0.04	33.9	4.3	H I 1215
107	4335.49	0.03	2.566342	15.30	2.54	15.9	11.2	H I 1215
108	4336.37	0.17	2.567062	13.41	0.15	34.9	11.4	H I 1215
109	4339.71	0.11	2.569806	12.87	0.20	27.7	27.9	H I 1215
110	4354.12	0.05	2.581664	12.99	0.08	20.2	5.0	H I 1215
111	4357.84	0.17	2.584723	12.69	0.19	25.2	19.4	H I 1215
112	4367.36	0.05	2.592551	13.21	0.07	27.6	4.7	H I 1215
113	4369.89	0.04	2.594638	16.97	0.70	31.1	3.6	H I 1215
114	4371.97	0.06	2.596345	14.42	0.05	49.5	4.1	H I 1215
115	4375.67	0.02	2.599391	13.76	0.03	25.7	2.5	H I 1215
116	4376.73	0.03	2.600261	13.88	1.95	5.1	5.0	H I 1215
117	4377.16	0.08	2.600616	12.82	0.32	9.5	11.7	H I 1215
118	4377.62	0.12	2.600991	13.85	0.06	65.0	7.7	H I 1215
119	4380.72	0.20	2.603541	13.39	0.10	72.2	15.6	H I 1215
120	4381.77	0.03	2.604405	13.40	0.08	24.3	3.9	H I 1215
121	4383.62	0.02	2.605932	13.60	0.07	21.8	3.2	H I 1215
122	4384.56	0.38	2.606703	13.21	0.20	71.6	31.6	H I 1215
123	4386.26	0.06	2.608103	12.99	0.10	19.8	6.0	H I 1215
124	4390.54	0.11	2.611625	13.72	0.08	45.0	7.5	H I 1215
125	4392.45	0.03	2.613191	15.65	0.39	43.2	5.2	H I 1215
126	4396.81	0.22	2.616778	13.62	0.08	124.5	21.9	H I 1215
127	4397.34	0.02	2.617218	14.45	0.19	28.3	3.8	H I 1215
128	4402.71	0.34	2.621635	13.60	0.29	55.9	12.0	H I 1215
129	4403.28	0.03	2.622101	14.86	0.41	25.4	5.3	H I 1215
130	4406.00	0.03	2.624337	13.60	0.03	38.6	3.0	H I 1215
131	4407.97	0.10	2.625962	13.73	0.11	31.5	6.4	H I 1215
132	4408.68	0.15	2.626542	13.27	0.30	22.6	9.3	H I 1215
133	4409.61	0.11	2.627306	12.47	0.24	15.1	18.0	H I 1215
134	4416.20	0.05	2.632730	13.27	0.05	34.8	5.0	H I 1215
135	4418.96	0.05	2.635003	12.68	0.18	9.0	8.5	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
136	4419.95	0.16	2.635813	13.18	0.09	58.3	14.9	H I 1215
137	4422.71	0.16	2.638085	12.80	0.14	36.0	15.1	H I 1215
138	4424.41	0.02	2.174456	13.34	0.04	22.8	3.0	Si IV 1393
139	4424.89	0.05	2.667541	12.94	0.05	24.5	3.3	Si III 1206
140	4429.27	0.02	2.643480	15.39	0.19	55.7	3.8	H I 1215
141	4433.46	0.08	2.646928	13.50	0.06	61.3	13.6	H I 1215
142	4435.05	0.15	2.648237	12.53	0.28	19.2	16.1	H I 1215
143	4436.26	0.13	2.649233	12.70	0.18	27.8	17.7	H I 1215
144	4440.54	0.20	2.652748	12.90	0.13	51.6	20.5	H I 1215
145	4442.89	0.02	2.654686	14.70	0.12	37.9	2.8	H I 1215
146	4444.94	0.08	2.656374	13.73	0.07	38.0	6.7	H I 1215
147	4446.71	0.03	2.657828	14.84	0.27	44.8	9.4	H I 1215
148	4448.26	0.10	2.659101	13.14	0.20	20.6	8.9	H I 1215
149	4449.48	0.06	2.660103	13.33	0.05	38.4	6.8	H I 1215
150	4453.03	0.02	2.174456	13.34	0.04	22.8	3.0	Si IV 1402
151	4455.40	0.03	2.664971	14.51	0.06	48.9	3.0	H I 1215
152	4458.52	0.04	2.667541	15.21	0.12	63.1	5.0	H I 1215
153	4460.19	0.12	2.668916	14.34	0.20	22.3	5.6	H I 1215
154	4461.75	0.02	2.670201	14.91	0.18	28.1	2.3	H I 1215
155	4464.66	0.27	2.672595	13.74	0.19	48.8	18.3	H I 1215
156	4465.59	0.04	2.673354	14.35	0.11	22.9	3.8	H I 1215
157	4466.45	0.11	2.674065	13.06	0.18	19.9	9.2	H I 1215
158	4468.45	0.22	2.675707	13.32	0.13	56.8	19.2	H I 1215
159	4469.67	0.15	2.676714	12.94	0.29	27.7	15.8	H I 1215
160	4473.00	0.09	2.679454	13.13	0.08	41.7	9.6	H I 1215
161	4474.02	0.07	2.680292	12.44	0.20	7.9	8.9	H I 1215
162	4475.79	0.18	2.681747	12.54	0.22	29.0	21.6	H I 1215
163	4477.19	0.09	2.682898	12.46	0.19	13.2	12.1	H I 1215
164	4479.09	0.14	2.684465	13.49	0.20	27.3	7.3	H I 1215
165	4479.81	0.26	2.685052	13.27	0.33	30.1	16.0	H I 1215
166	4481.09	0.02	2.686110	13.59	0.03	24.5	2.5	H I 1215
167	4482.38	0.03	2.687169	13.61	0.04	26.3	3.2	H I 1215
168	4483.38	0.22	2.687988	12.86	0.22	34.3	19.9	H I 1215
169	4485.57	0.18	2.689795	13.00	0.27	24.4	12.5	H I 1215
170	4486.33	0.04	2.690415	13.83	0.05	25.4	3.0	H I 1215
171	4488.61	0.01	2.220513	12.55	0.10	9.3	6.3	Si IV 1393
172	4489.70	0.02	2.221298	12.99	0.06	29.8	5.6	Si IV 1393
173	4492.55	0.03	2.695531	13.73	0.03	42.6	2.9	H I 1215
174	4498.12	0.12	2.700117	12.52	0.21	15.8	12.7	H I 1215
175	4504.62	0.02	2.705461	13.94	0.06	22.5	1.7	H I 1215
176	4507.74	0.10	2.708033	12.55	0.16	17.1	7.9	H I 1215
177	4509.32	0.06	2.709331	13.11	0.07	26.0	5.4	H I 1215
178	4510.56	0.03	2.710352	13.60	0.04	28.8	3.5	H I 1215
179	4512.09	0.02	2.711609	14.69	0.19	30.1	3.1	H I 1215
180	4515.47	0.02	2.714392	13.46	0.04	21.3	2.4	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
181	4516.38	0.02	2.715138	13.24	0.05	12.3	2.6	H I 1215
182	4517.64	0.01	2.220513	12.55	0.10	9.3	6.3	Si IV 1402
183	4518.59	0.02	2.716955	14.38	0.03	63.8	2.2	H I 1215
184	4518.74	0.02	2.221298	12.99	0.06	29.8	5.6	Si IV 1402
185	4521.80	0.10	2.719594	12.58	0.19	19.4	17.2	H I 1215
186	4523.86	0.03	2.721288	13.06	0.06	15.2	3.7	H I 1215
187	4525.96	0.05	2.723016	14.03	0.03	54.0	4.4	H I 1215
188	4527.28	0.15	2.724100	13.08	0.22	32.2	12.8	H I 1215
189	4529.39	0.12	2.725840	13.09	0.11	36.1	12.3	H I 1215
190	4530.53	0.05	2.726773	13.40	0.15	25.6	7.2	H I 1215
191	4531.55	0.20	2.727618	13.50	0.12	56.1	13.9	H I 1215
192	4534.11	0.10	2.729720	12.96	0.10	31.6	8.6	H I 1215
193	4535.61	0.04	2.730956	13.83	0.04	31.4	3.0	H I 1215
194	4536.88	0.10	2.731999	13.52	0.07	46.1	8.9	H I 1215
195	4539.69	0.07	2.734314	12.73	0.12	18.1	6.4	H I 1215
196	4544.68	0.30	2.738417	13.40	0.17	63.6	15.1	H I 1215
197	4545.31	0.03	2.738932	13.53	0.11	22.8	4.3	H I 1215
198	4546.84	0.12	2.740193	12.87	0.13	31.9	10.8	H I 1215
199	4548.28	0.05	2.741377	12.76	0.10	16.0	7.8	H I 1215
200	4550.57	0.07	2.743260	13.24	0.05	45.4	7.5	H I 1215
201	4552.92	0.03	2.745193	13.12	0.10	13.6	4.5	H I 1215
202	4553.46	0.11	2.745636	13.49	0.05	67.2	8.4	H I 1215
203	4556.81	0.10	2.748391	13.47	0.04	80.0	10.6	H I 1215
204	4559.88	0.21	2.750917	12.67	0.16	41.0	18.6	H I 1215
205	4565.70	0.10	2.755708	13.23	0.07	52.3	10.5	H I 1215
206	4567.12	0.02	2.756878	13.88	0.04	23.9	1.7	H I 1215
207	4568.55	0.02	2.758050	13.66	0.03	24.9	1.8	H I 1215
208	4570.94	0.07	2.760013	12.96	0.09	23.7	6.9	H I 1215
209	4572.27	0.03	2.761112	13.52	0.05	21.4	3.0	H I 1215
210	4573.63	0.45	2.762233	13.14	0.21	59.1	27.2	H I 1215
211	4574.39	0.03	2.762852	13.52	0.08	20.3	3.7	H I 1215
212	4579.79	0.09	2.767298	13.06	0.12	26.2	7.5	H I 1215
213	4580.79	0.24	2.768119	12.83	0.23	35.1	19.5	H I 1215
214	4582.21	0.08	2.769290	13.17	0.10	40.0	16.3	H I 1215
215	4584.64	0.02	2.771289	14.46	0.07	36.6	2.2	H I 1215
216	4588.75	0.08	2.774668	13.03	0.08	33.9	8.6	H I 1215
217	4590.98	0.03	2.776499	13.28	0.04	24.3	3.1	H I 1215
218	4592.27	0.03	2.777559	13.36	0.04	25.1	2.7	H I 1215
219	4594.25	0.05	2.779193	13.12	0.06	28.8	5.0	H I 1215
220	4595.55	0.05	2.780265	12.97	0.07	21.6	4.8	H I 1215
221	4598.91	0.12	2.783026	12.96	0.09	43.1	10.5	H I 1215
222	4602.28	0.01	2.785799	14.47	0.09	30.7	1.9	H I 1215
223	4604.33	0.18	2.787483	12.71	0.15	36.9	15.8	H I 1215
224	4605.84	0.02	2.788728	13.56	0.03	20.0	1.8	H I 1215
225	4608.10	0.05	2.790586	13.29	0.05	31.1	5.0	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
226	4609.93	0.03	2.792086	13.65	0.04	26.2	3.1	H I 1215
227	4611.04	0.05	2.792999	13.38	0.06	30.4	6.1	H I 1215
228	4613.41	0.02	2.794954	14.11	0.05	27.8	1.8	H I 1215
229	4617.62	0.04	2.798412	13.03	0.07	13.1	4.8	H I 1215
230	4618.58	0.02	2.799205	13.80	0.03	24.8	2.7	H I 1215
231	4619.43	0.02	2.799905	13.63	0.10	10.2	1.9	H I 1215
232	4621.13	0.02	2.801305	13.88	0.03	32.9	1.6	H I 1215
233	4624.41	0.05	2.804004	13.05	0.06	22.6	4.7	H I 1215
234	4625.35	0.06	2.804771	12.77	0.10	15.7	5.9	H I 1215
235	4626.45	0.01	2.805676	13.82	0.03	23.0	1.5	H I 1215
236	4630.87	0.76	2.809318	13.18	0.32	80.1	48.3	H I 1215
237	4632.50	0.02	2.810652	14.76	0.15	39.5	4.1	H I 1215
238	4634.56	0.04	2.812350	13.97	0.13	45.6	6.8	H I 1215
239	4635.79	0.81	2.813359	13.72	0.28	105.1	42.1	H I 1215
240	4638.73	0.12	2.815778	12.76	0.16	28.0	10.0	H I 1215
241	4640.04	0.01	2.329162	13.71	0.04	23.3	1.4	Si IV 1393
242	4641.59	0.12	2.330278	13.17	0.16	25.4	5.6	Si IV 1393
243	4642.01	0.03	2.330575	13.36	0.11	14.3	3.2	Si IV 1393
244	4642.62	0.06	2.331016	12.50	0.15	12.8	6.6	Si IV 1393
245	4644.51	0.28	2.820536	13.21	0.14	64.5	22.1	H I 1215
246	4648.74	0.04	2.824017	15.76	0.08	94.6	4.5	H I 1215
247	4651.68	0.04	2.826435	14.47	0.08	17.9	4.6	H I 1215
248	4653.78	0.04	2.828157	16.66	0.45	45.3	4.9	H I 1215
249	4657.54	0.07	2.831252	14.07	0.03	66.5	6.3	H I 1215
250	4660.68	0.02	2.833838	17.88	0.15	47.2	1.2	H I 1215
251	4664.42	0.11	2.836914	12.35	0.21	13.3	10.6	H I 1215
252	4668.07	0.11	2.839919	13.10	0.42	25.0	9.8	H I 1215
253	4668.76	0.32	2.840480	13.32	0.26	43.4	17.4	H I 1215
254	4670.05	0.01	2.329162	13.71	0.04	23.3	1.4	Si IV 1402
255	4670.48	0.04	2.841898	13.57	0.05	25.2	3.2	H I 1215
256	4671.61	0.12	2.330278	13.17	0.16	25.4	5.6	Si IV 1402
257	4672.03	0.03	2.330575	13.36	0.11	14.3	3.2	Si IV 1402
258	4672.65	0.06	2.331016	12.50	0.15	12.8	6.6	Si IV 1402
259	4673.69	0.04	2.844534	13.67	0.04	27.0	3.2	H I 1215
260	4674.50	0.08	2.845205	13.04	0.19	17.4	8.1	H I 1215
261	4675.33	0.06	2.845888	13.45	0.06	31.1	5.5	H I 1215
262	4677.95	0.14	2.848046	12.83	0.12	38.4	12.3	H I 1215
263	4679.57	0.03	2.849373	13.44	0.04	31.3	3.3	H I 1215
264	4681.68	0.15	2.851109	12.65	0.19	26.5	20.5	H I 1215
265	4683.29	0.07	2.852431	12.89	0.16	17.0	9.7	H I 1215
266	4684.39	0.08	2.853342	13.43	0.06	41.5	6.8	H I 1215
267	4688.03	0.14	2.856337	12.84	0.12	37.3	11.5	H I 1215
268	4689.78	0.02	2.857773	13.89	0.04	24.0	1.7	H I 1215
269	4693.20	0.03	2.860584	15.22	0.04	67.4	1.8	H I 1215
270	4695.37	0.09	2.862375	13.04	0.20	17.7	8.9	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
271	4696.39	0.04	2.863210	14.15	0.06	26.0	4.2	H I 1215
272	4697.31	0.36	2.863969	13.18	0.36	33.8	21.3	H I 1215
273	4700.54	0.03	2.866626	14.28	0.06	77.4	4.9	H I 1215
274	4700.56	0.05	2.866642	14.25	0.09	30.5	4.5	H I 1215
275	4704.45	0.12	2.869837	12.52	0.17	20.3	11.5	H I 1215
276	4706.43	0.05	2.871469	14.46	0.08	34.1	3.0	H I 1215
277	4707.58	0.15	2.872414	13.47	0.19	29.9	11.8	H I 1215
278	4708.78	0.07	2.873405	13.23	0.08	31.0	6.6	H I 1215
279	4710.48	0.04	2.874801	13.53	0.03	40.8	3.7	H I 1215
280	4713.49	0.06	2.877276	13.31	0.06	34.1	5.2	H I 1215
281	4714.92	0.05	2.878452	13.55	0.04	42.3	4.7	H I 1215
282	4717.26	0.17	2.880378	12.73	0.16	37.4	21.2	H I 1215
283	4720.31	0.49	2.882890	13.55	0.47	39.0	17.5	H I 1215
284	4720.97	0.03	2.883430	14.87	0.10	24.2	1.8	H I 1215
285	4722.73	0.15	2.884880	12.98	0.10	46.2	12.8	H I 1215
286	4725.72	0.01	2.916881	13.07	0.06	14.3	1.5	Si III 1206
287	4728.56	0.11	2.889675	12.35	0.23	12.3	13.8	H I 1215
288	4730.37	0.21	2.891163	12.85	0.18	41.7	23.5	H I 1215
289	4731.54	0.17	2.892126	12.55	0.18	25.7	12.8	H I 1215
290	4733.34	0.02	2.893610	14.51	0.15	32.0	3.6	H I 1215
291	4734.57	0.22	2.894616	12.66	0.33	22.1	18.0	H I 1215
292	4736.61	0.14	2.896298	12.76	0.28	24.9	15.3	H I 1215
293	4737.81	0.27	2.897286	12.88	0.30	47.9	41.1	H I 1215
294	4739.34	0.17	2.898542	13.50	0.20	30.7	10.5	H I 1215
295	4740.12	0.05	2.899185	13.95	0.08	24.1	3.1	H I 1215
296	4741.76	0.03	2.900530	13.53	0.03	36.2	3.4	H I 1215
297	4743.87	0.20	2.902267	12.45	0.22	28.2	21.7	H I 1215
298	4745.66	0.15	2.903742	12.94	0.19	28.6	13.8	H I 1215
299	4746.78	0.17	2.904662	13.01	0.16	35.7	16.2	H I 1215
300	4748.72	0.05	2.906256	14.03	0.08	33.5	3.9	H I 1215
301	4750.02	0.23	2.907327	13.92	0.09	67.8	11.4	H I 1215
302	4751.04	0.02	2.908167	13.62	0.07	14.7	2.5	H I 1215
303	4752.51	0.10	2.909377	12.41	0.18	12.8	9.0	H I 1215
304	4753.28	0.09	2.910005	12.57	0.17	16.8	10.9	H I 1215
305	4754.28	0.03	2.910834	13.31	0.04	26.3	3.5	H I 1215
306	4757.33	0.16	2.913343	13.01	0.15	45.0	20.5	H I 1215
307	4758.22	0.04	2.914068	12.69	0.16	6.8	8.4	H I 1215
308	4758.80	0.10	2.914548	12.58	0.18	12.4	7.9	H I 1215
309	4759.62	0.04	2.915227	14.34	0.06	18.8	2.5	H I 1215
310	4761.64	0.01	2.916881	16.25	0.18	48.9	2.3	H I 1215
311	4765.64	0.02	2.920175	13.51	0.04	23.0	2.2	H I 1215
312	4766.89	0.18	2.921201	13.21	0.29	27.3	11.7	H I 1215
313	4767.87	0.23	2.922010	13.48	0.16	43.2	13.2	H I 1215
314	4769.60	0.06	2.923434	12.74	0.11	15.7	5.8	H I 1215
315	4771.96	0.08	2.925373	13.38	0.05	50.4	6.9	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
316	4773.58	0.02	2.926706	13.86	0.04	23.2	1.7	H I 1215
317	4775.58	0.11	2.928351	13.19	0.07	55.2	11.7	H I 1215
318	4778.30	0.07	2.930590	14.00	0.21	57.6	9.2	H I 1215
319	4778.43	0.02	2.930699	14.88	0.06	28.9	2.1	H I 1215
320	4781.87	0.18	2.933525	12.81	0.12	47.7	16.5	H I 1215
321	4784.89	0.18	2.936013	13.49	0.10	51.6	10.0	H I 1215
322	4786.58	0.03	2.937400	14.68	0.09	40.5	2.9	H I 1215
323	4789.74	0.01	2.939998	14.57	0.09	32.0	1.7	H I 1215
324	4794.97	0.17	2.944299	13.07	0.13	44.3	15.9	H I 1215
325	4796.14	0.06	2.945268	13.24	0.09	26.8	5.1	H I 1215
326	4797.42	0.07	2.946318	12.90	0.08	27.8	7.1	H I 1215
327	4800.12	0.04	2.948535	14.30	0.24	21.0	5.8	H I 1215
328	4800.52	0.26	2.948870	13.71	0.28	42.3	7.7	H I 1215
329	4802.43	0.02	2.950442	14.38	0.08	31.8	2.2	H I 1215
330	4810.83	0.04	2.957352	13.44	0.03	38.4	4.4	H I 1215
331	4815.43	1.03	2.961134	13.04	0.61	58.2	47.1	H I 1215
332	4816.16	0.05	2.961731	13.36	0.26	27.6	7.4	H I 1215
333	4817.56	0.05	2.962883	13.73	0.31	32.9	8.0	H I 1215
334	4818.12	1.02	2.963344	13.38	0.73	57.1	29.5	H I 1215
335	4823.17	0.09	2.967498	13.04	0.07	41.8	9.0	H I 1215
336	4824.82	0.06	2.968860	12.83	0.11	22.2	10.9	H I 1215
337	4825.79	0.11	2.969651	12.54	0.15	18.9	9.2	H I 1215
338	4827.90	0.12	2.971390	12.68	0.17	21.8	13.2	H I 1215
339	4828.54	0.04	2.971919	12.71	0.13	7.8	7.1	H I 1215
340	4829.67	0.25	2.972850	12.62	0.23	41.1	28.9	H I 1215
341	4830.79	0.15	2.973771	12.41	0.28	19.4	14.1	H I 1215
342	4832.61	0.04	2.975262	13.19	0.09	22.9	4.6	H I 1215
343	4833.59	0.25	2.976074	13.48	0.07	87.8	15.6	H I 1215
344	4835.89	0.03	2.977966	14.27	0.03	39.8	2.9	H I 1215
345	4837.17	0.07	2.979019	13.25	0.10	24.5	4.7	H I 1215
346	4839.84	0.10	2.981215	13.15	0.11	29.6	8.3	H I 1215
347	4841.09	0.03	2.982238	14.00	0.03	35.2	2.4	H I 1215
348	4842.80	0.08	2.983644	13.00	0.11	22.7	6.1	H I 1215
349	4843.85	0.05	2.984509	13.37	0.05	32.8	5.7	H I 1215
350	4845.64	0.02	2.985983	13.87	0.03	25.2	1.4	H I 1215
351	4854.25	0.13	2.993069	12.63	0.15	27.8	14.4	H I 1215
352	4857.62	0.11	2.995834	13.08	0.07	53.5	11.9	H I 1215
353	4864.39	0.15	3.001410	13.04	0.10	52.5	15.8	H I 1215
354	4866.02	0.02	3.002744	13.68	0.02	28.7	1.8	H I 1215
355	4867.59	0.13	3.004042	12.41	0.16	21.7	9.6	H I 1215
356	4869.07	0.08	3.005256	12.86	0.07	32.5	6.7	H I 1215
357	4871.48	0.08	3.007242	13.17	0.09	31.4	6.3	H I 1215
358	4872.72	0.17	3.008260	12.96	0.22	40.9	24.3	H I 1215
359	4873.77	0.04	3.009125	13.16	0.07	19.6	3.0	H I 1215
360	4875.75	0.03	3.010753	14.37	0.04	41.9	2.4	H I 1215

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
361	4877.19	0.14	3.011936	13.27	0.15	38.3	12.4	H I 1215
362	4878.65	0.05	3.013133	13.21	0.05	29.9	5.0	H I 1215
363	4881.92	0.17	3.015825	12.69	0.18	27.7	12.9	H I 1215
364	4883.02	0.04	3.016728	13.42	0.04	29.1	3.6	H I 1215
365	4888.37	0.01	3.021131	14.24	0.06	26.8	1.5	H I 1215
366	4890.23	0.02	3.022662	14.03	0.02	39.4	1.8	H I 1215
367	4892.55	0.01	3.024571	14.53	0.07	35.4	1.6	H I 1215
368	4894.67	0.10	3.026315	13.18	0.08	49.9	12.0	H I 1215
369	4896.50	0.02	3.027823	14.20	0.04	31.0	2.4	H I 1215
370	4897.75	0.02	3.028847	13.91	0.03	23.1	1.8	H I 1215
371	4899.72	0.12	3.030469	12.86	0.08	47.5	12.6	H I 1215
372	4902.70	0.02	3.032920	13.47	0.03	26.9	1.8	H I 1215
373	4903.99	0.14	3.033977	12.92	0.10	46.2	13.3	H I 1215
374	4907.29	0.04	3.036694	12.81	0.06	13.7	3.8	H I 1215
375	4907.94	0.02	3.037231	13.09	0.04	9.0	2.5	H I 1215
376	4909.51	0.01	3.038523	15.11	0.05	37.6	0.9	H I 1215
377	4913.18	0.43	3.041539	13.19	0.13	111.3	38.1	H I 1215
378	4914.88	0.01	2.174587	14.25	0.02	23.6	0.9	C IV 1548
379	4915.40	0.14	3.043367	13.24	0.17	29.6	8.1	H I 1215
380	4917.76	0.03	3.045306	13.72	0.07	25.7	2.8	H I 1215
381	4918.16	0.13	3.045638	13.53	0.12	52.7	5.1	H I 1215
382	4923.06	0.01	2.174587	14.25	0.02	23.6	0.9	C IV 1550

C.3 Fitted lines above Ly α emission

	λ_{obs} (Å)	σ_λ	z	log N (cm^{-2})	σ_N	b (km s^{-1})	σ_b	ID
1	4931.76	0.15	3.056827	12.72	0.10	44.0	10.7	H I 1215
2	4935.51	0.05	3.059906	12.86	0.06	25.6	5.5	H I 1215
3	4939.78	0.02	3.063420	13.48	0.02	33.1	2.0	H I 1215
4	4941.61	0.04	3.064924	12.73	0.06	17.7	3.7	H I 1215
5	4943.91	0.04	3.066818	13.02	0.04	34.5	3.3	H I 1215
6	4985.80	0.06	2.220396	13.44	0.05	41.6	5.1	C IV 1548
7	4985.98	0.01	2.220513	13.74	0.03	8.9	1.0	C IV 1548
8	4987.20	0.02	2.221298	13.57	0.04	14.8	2.9	C IV 1548
9	4987.60	0.02	2.221561	13.21	0.13	5.5	1.3	C IV 1548
10	4994.09	0.06	2.220397	13.44	0.05	41.6	5.1	C IV 1550
11	4994.28	0.01	2.220513	13.74	0.03	8.9	1.0	C IV 1550
12	4995.49	0.02	2.221298	13.57	0.04	14.8	2.9	C IV 1550
13	4995.90	0.02	2.221561	13.21	0.13	5.5	1.3	C IV 1550
14	4996.75	0.07	3.110281	12.62	0.08	23.6	6.3	H I 1215
15	5111.65	0.05	2.667541	12.80	0.06	21.9	4.3	Si IV 1393
16	5144.72	0.05	2.667541	12.80	0.06	21.9	4.3	Si IV 1402
17	5154.19	0.01	2.329162	14.75	0.13	19.3	1.5	C IV 1548
18	5155.92	0.13	2.330278	14.19	0.12	35.9	5.0	C IV 1548
19	5156.38	0.04	2.330575	14.03	0.16	13.7	4.3	C IV 1548
20	5157.06	0.07	2.331016	13.93	0.07	30.3	3.9	C IV 1548
21	5162.76	0.01	2.329162	14.75	0.13	19.3	1.5	C IV 1550
22	5164.50	0.13	2.330279	14.19	0.12	35.9	5.0	C IV 1550
23	5164.96	0.04	2.330575	14.03	0.16	13.7	4.3	C IV 1550
24	5165.64	0.07	2.331016	13.93	0.07	30.3	3.9	C IV 1550

Acknowledgments

It is a pleasure to acknowledge the help of those who contributed to the completion of this work.

Many and sincere thanks go to Prof. Dieter Reimers for the very worthwhile research topic, and for his admirable willingness to allow this work to be attempted under difficult and unusual circumstances.

I am also very grateful to Dr. Susanne Köhler, who provided thoughtful and amazingly patient guidance, as well as continued advice even while on leave.

To Ana de la Varga Villagra, warm thanks for the endless discussions, tireless technical support and general solidarity.

Sebastián López and Dr. Lutz Wisotzki kindly provided data, materials and occasional advice that was always appreciated. The entire group in “the Villa” created a pleasant working atmosphere that I was sorry not to be able to enjoy more often.

Special thanks as well to Irmgard and Jens Flick and Dorothea Samtleben for their logistical support in Hamburg. Last but not least, thanks to my husband Ulrich and to my parents for their continual love and encouragement.