1996ApJ...457..102D

THE ASTROPHYSICAL JOURNAL, 457:102–110, 1996 January 20
© 1996. The American Astronomical Society. All rights reserved. Printed in U.S.A.

# SIMULATION ANALYSIS OF Lyα FOREST SPECTRA. I. EMPIRICAL DESCRIPTION AT $z \approx 3$

ADAM DOBRZYCKI[1]
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138; adobrzycki@cfa.harvard.edu
AND
JILL BECHTOLD
Steward Observatory, University of Arizona, Tucson, AZ 85721; jbechtold@as.arizona.edu
*Received 1995 February 6; accepted 1995 July 28*

## ABSTRACT

We present moderate-resolution (~50 km s$^{-1}$ FWHM) spectra of the Lyα forest for seven quasars with redshifts ranging from 2.53 to 3.13, obtained with the Blue Spectrograph and photon-counting Reticon at the Multiple Mirror Telescope. Combined with spectra of 10 other quasars presented elsewhere, we have characterized the distribution of cloud properties in a way which was designed to minimize any subjective part of the analysis.

We used artificial absorption spectra, with the same resolution, sampling and signal-to-noise ratio as a function of wavelength as the actual data. Distributions of the physical parameters of the Lyα clouds, namely, the neutral hydrogen column density ($N$) and Doppler parameter ($b$), were approximated with $d\mathcal{N}/dN \propto N^{-\beta}$ and $d\mathcal{N}/db \propto \exp[-(b - \langle b \rangle)^2/2\,\sigma_b^2]$, respectively. We constructed a grid of simulated spectra with different input parameters. Comparison of properties of the simulated spectra with the observed spectra yielded acceptable ranges of parameters. Our technique differs from previous similar work in that we use the information contained in the distribution of the strength of the absorption in each resolution element and the distribution of separations between absorption complexes. We derive $\beta = 1.4 \pm 0.1$ for $N$ ranging from $10^{13}$ to $10^{16}$ cm$^{-2}$ and $\langle b \rangle = 30 \pm 15$ km s$^{-1}$. Most previous studies based on line lists indicated $\beta = 1.7$–$1.9$. We attribute this difference to flattening of the column density distribution for low $N$, recently confirmed by higher resolution observations. Our result for $\langle b \rangle$, though consistent with values quoted in the literature, is of lower significance, since it is less than the resolution of our spectra. We conclude by commenting on the importance of line blending in data sets of this kind.

*Subject headings:* cosmology: observations — intergalactic medium — quasars: absorption lines

## 1. INTRODUCTION

While the intrinsic properties of the so-called Lyα forest clouds are still debated, and their physical state is not really agreed upon, they have proven to be an invaluable tool in the analysis of the conditions in the early universe. Most conclusions depend on the statistical description of the lines and their evolution. The redshift evolution of the number of lines per unit redshift for lines with rest equivalent width greater than some threshold value $W_{\rm thr}$ can be written as

$$\frac{d\mathcal{N}}{dz} = \mathcal{A}_0(1 + z)^{\gamma} \, . \tag{1}$$

For a nonevolving population of clouds $\frac{1}{2} \leq \gamma \leq 1$, depending on the value of $q_0$ (Peterson 1978). Although the actual value of the index $\gamma$ for $z \gtrsim 2$ is debated, there is a consensus that it is significantly higher than unity, which is a clear indication that the clouds undergo evolution (Murdoch et al. 1986; Bajtlik, Duncan, & Ostriker 1988; Lu, Wolfe, & Turnshek 1991; Rauch et al. 1992; Frye et al. 1993; Bechtold 1994, hereafter B94). In the present study we do not concentrate on this problem, since the nature of our sample does not allow us to put significant constraints on the cloud number density evolution.

Instead, we consider the distribution of line properties at a fixed redshift, $z \approx 3$. The line equivalent widths, $W$, show a distribution which is well described by an exponential:

$$\frac{d\mathcal{N}}{dW} \propto \exp\left(-\frac{W}{W_*}\right) . \tag{2}$$

It is, however, more interesting to determine the distributions of the two physical parameters of the Lyα clouds, namely, the H I column density, $N$, and Doppler broadening parameter, $b$. In the traditional approach, this requires fitting a Voigt profile to the observed lines in data with sufficient resolution to resolve the lines (e.g., Rauch et al. 1992, and references therein). Column densities of the Lyα lines appear to follow a simple power law:

$$\frac{d\mathcal{N}}{dN} \propto N^{-\beta} , \tag{3}$$

while the Doppler parameter distribution can be approximated with a Gaussian:

$$\frac{d\mathcal{N}}{db} \propto \exp\left[-\frac{(b - \langle b \rangle)^2}{2\,\sigma_b^2}\right] . \tag{4}$$

The most recent analysis of the column density distribution for the Lyα forest clouds was presented in a paper by Petitjean et al. (1993), who also provided a summary of previous research on this problem. They analyzed the distribution of column densities of quasar absorbers over wide range of $N$, from $\sim 10^{13}$ to $\sim 10^{22}$ cm$^{-2}$. In the range relevant to the subject of the present analysis, their data sample was a compilation of

[1] On leave from the Copernicus Astronomical Center, Bartycka 18, 00-716 Warsaw, Poland.



© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



line lists available from the literature from high-resolution (20–30 km s$^{-1}$) observations. They found, in agreement with previous studies, that for lines with $N > 10^{13.7}$ cm$^{-2}$ (the completeness limit of their sample) the distribution of column density is steep, $\beta = 1.83 \pm 0.05$. Both Petitjean et al. and Carswell et al. (1987) found an indication that there may be some flattening of the H I column density distribution for lower $N$. This flattening is seen in high-resolution spectra, as reported by Tytler (1995).

The distribution of equivalent widths (eq. [2]) can be derived from the distribution of column densities (eq. [3]) and the distribution of $b$-values (eq. [4]) through the curve of growth. In practice however, it turns out that the values of the distribution parameters quoted in the literature ($W_* = 0.25$–$0.30$ Å, e.g., Murdoch et al. 1986, B94; $\beta = 1.7$–$1.9$, e.g., Carswell et al. 1991, Petitjean et al. 1993, and references therein; $\langle b \rangle \approx 30$ km s$^{-1}$, e.g., Oke & Korycansky 1982; Carswell et al. 1991) are inconsistent with one another (Barcons & Webb 1991; Chernomordik & Ozernoy 1993). The source of the inconsistency may be a subtle problem with line blending, or the presence of clustering of weak lines (Barcons & Webb 1991). Other effects attributable to blending have been discussed by Liu & Jones (1988), Parnell & Carswell (1988), Barcons & Webb (1991), and Trevese, Giallongo, & Camurani (1992).

Proper description of the column density distribution has important astrophysical consequences. It has been shown (Bechtold 1994) that uncertainty in the parameter $\beta$ introduces bias in the determinations of the intensity of the ambient UV ionizing background, $J_\nu$. Behavior of the distribution at the low end carries important information on the structure of the diffuse matter, namely, whether it consists of a large number of low column density "cloudlets," or whether it is a smooth, featureless gas. Presently available data, when subjected to a traditional, line list based analysis, do not reach suitable column densities.

In this paper we present a new method for determining the distributions of the physical parameters of the clouds, using an automated comparison of the observed spectra with simulated ones. We show that the method is capable of unveiling effects at column densities below the resolution of the data. We obtained a large, homogeneous sample of quasar Ly$\alpha$ region spectra in order to compare them with the simulations. We concentrate here on the determination of column density and Doppler parameter distributions at $z \approx 3$. Results on an extended sample in redshift will be given in a subsequent paper.

Press, Rybicki, & Schneider (1993) and Press & Rybicki (1993) presented a similar kind of analysis. Our spectra contain qualitatively different information than theirs, however, and so applying their technique directly would not make use of all the information we have. They used low-resolution spectra of high-redshift ($z_{\rm em} = 3.14$–$4.46$) quasars from Schneider, Schmidt, & Gunn (1991). Their study was "ideologically" similar to ours in that they also did not attempt to derive the properties of individual lines. However, the spectral resolution of their data was 25 Å, compared to a typical observed equivalent width for a Ly$\alpha$ forest line of 0.5 Å or less. They analyzed depressions of the quasar continua in the resolution elements in the Ly$\alpha$ forest parts of the spectra, assuming that the continuum shortward of Ly$\alpha$ emission is a power-law extrapolation of the power law redward of Ly$\alpha$. Virtually the entire spectrum shortward of Ly$\alpha$ emission was depressed, i.e., absorbed, compared to this extrapolation in all of their objects. In fact, they excluded any regions in their spectra that were not absorbed, since they assumed these were spurious. Also, since their spectral resolution was inadequate to identify any metal line systems, these were ignored.

TABLE 1

OBSERVATIONS

| QSO | $z_{\rm em}$ | $V$ | Setup[a] | Date | Exposure (s) |
|---|---|---|---|---|---|
| 1206+119...... | 3.108 | 17.9 | 1 | 1991 May 12 | 6000 |
| 1225−017...... | 2.831 | 18.0 | 1 | 1991 May 11 | 7200 |
| 1315+472...... | 2.590 | 18.0 | 1 | 1991 May 11–12 | 13800 |
| 1607+183...... | 3.134 | 18.5 | 1 | 1991 May 12–13 | 16900 |
| 1623+269...... | 2.526 | 16.0 | 1 | 1991 May 12 | 2400 |
| 1700+643...... | 2.744 | 16.1 | 1 | 1991 May 11 | 2400 |
| 1946+770...... | 3.020 | 15.9 | 2 | 1992 Oct 24–25 | 2400 |

[a] Instrument setup: (1) MMT, Big Blue Reticon, 832 lines mm$^{-1}$, second order, Image Stacker: (2) MMT, Big Blue Reticon, 832 lines mm$^{-1}$, second order, 1″ × 3″ slit.

Their assumptions in general will tend to overestimate the amount of absorption, and in any event they are not required for the analysis of our spectra. The assumption that the continuum in the Ly$\alpha$ forest region follows the power-law energy distribution from the red part of the spectrum is questionable, since this is exactly the place at which accretion disk and free-free models for the extreme-ultraviolet (EUV) quasar energy distribution predict significant deviations from a power law, especially for high-luminosity objects (see Kuhn et al. 1995; Siemiginowska et al. 1995; and references therein). Moreover, our spectra do have regions with no significant absorption, and we can use the distribution of "gaps" in the absorption to further constrain the Ly$\alpha$ forest parameters.

This paper is organized as follows. We present the observations and define the sample used in the subsequent analysis in §§ 2 and 3. We present the flux deficit analysis of the redshift evolution of the clouds in § 4. In §§ 5 and 6 we describe the simulation software and analysis, and we discuss and summarize the results in §§ 7 and 8. In § 9 we comment on the importance of line blending in our and similar data.

## 2. OBSERVATIONS

We observed seven quasars (see Table 1) especially for this project. The spectra were obtained at the Multiple Mirror Telescope (MMT)[2] using the Blue Channel spectrograph with 832 lines mm$^{-1}$ grating in second order, 1″ × 3″ slit (Q1946+770) or Image Stacker (other quasars), "Big Blue" image tube, and photon counting Reticon (Latham 1982). A solid $CuSO_4$ order blocking filter was placed in the beam. Every 20 minutes, an exposure of a He-Ne-Ar-Hg-Cd lamp was taken for wavelength calibration. During the day, long quartz lamp exposures were obtained in order to account for instrumental pixel-to-pixel variations.

Spectra were sky subtracted, wavelength calibrated, rebinned, and combined with the use of IRAF and software written by us. For all quasars, we then determined signal-to-noise ratios and continua, using the technique described in detail in B94. The spectral resolution was determined by calculating the FWHM of the emission lines in the comparison lamp exposures.

All spectra are shown in Figure 1. Absorption lines in the spectra of all 17 quasars were extracted, and the metal-system

[2] MMT is a joint facility of the Smithsonian Institution and the University of Arizona.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D

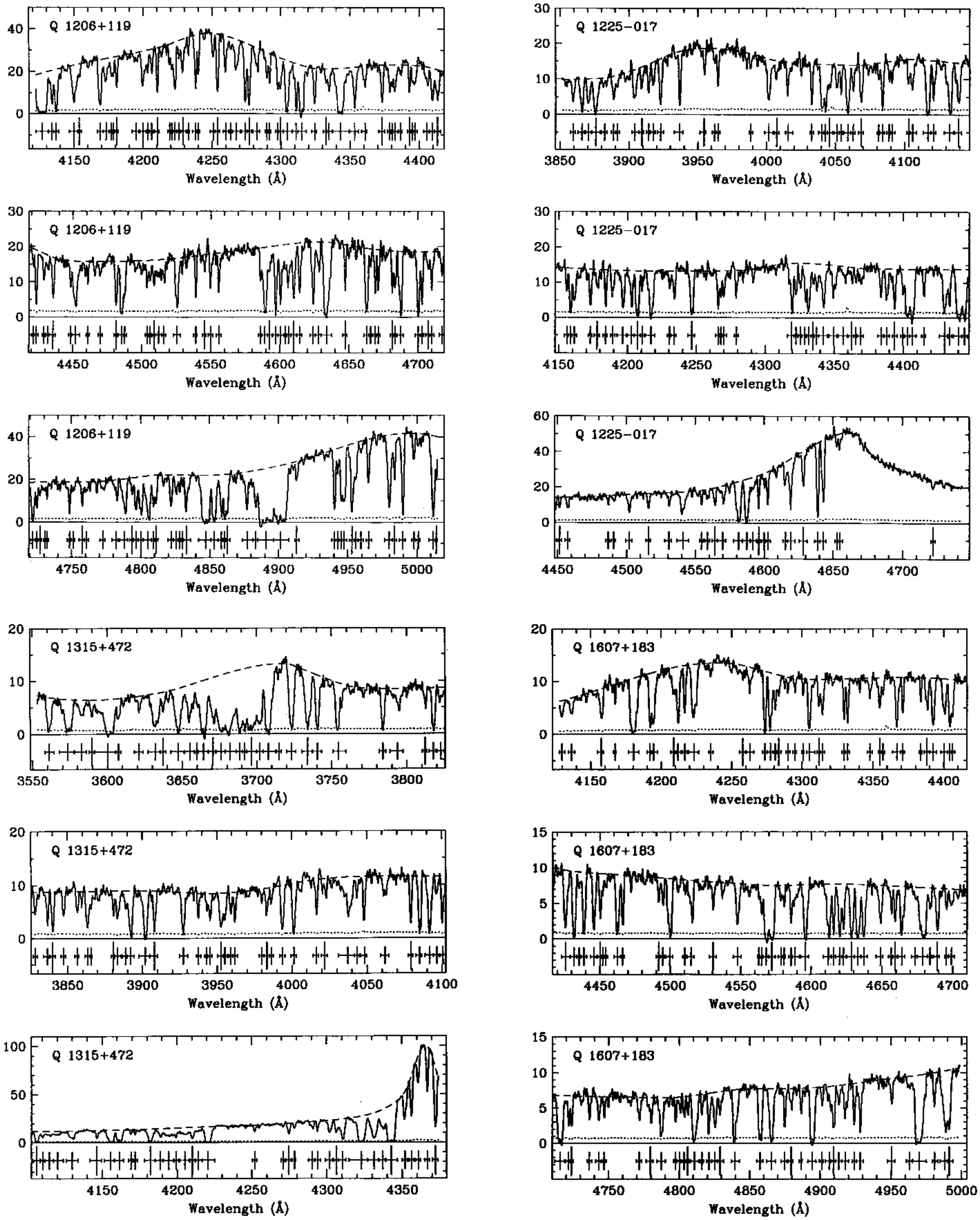
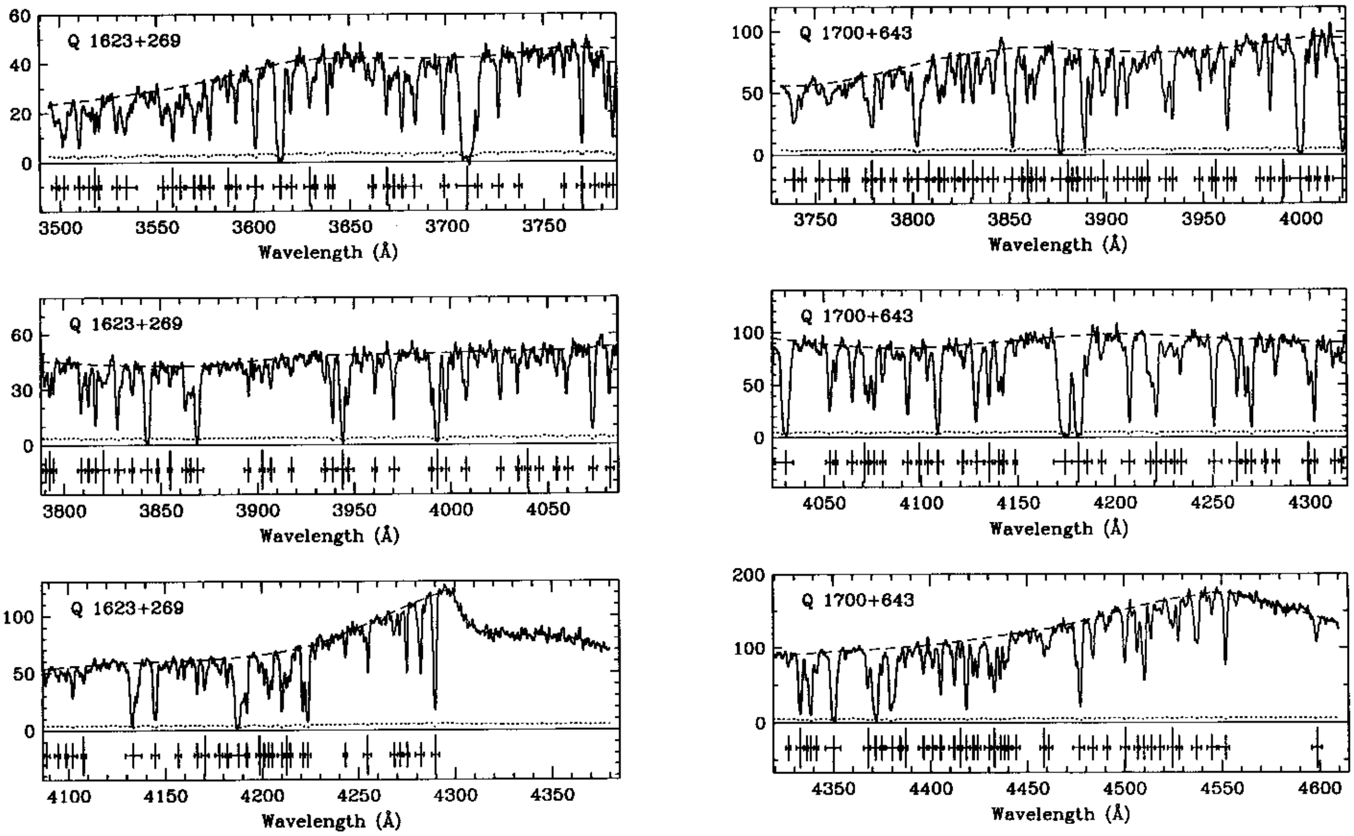



FIG. 1.—Spectra of seven quasars obtained at the MMT. Dashed line indicates the continuum, dotted line indicates 1 $\sigma$ uncertainties. Significant lines listed in Table 2 are indicated by vertical bars below each spectrum. Spectra have been smoothed with a Gaussian filter with FWHM of 2 pixels.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D





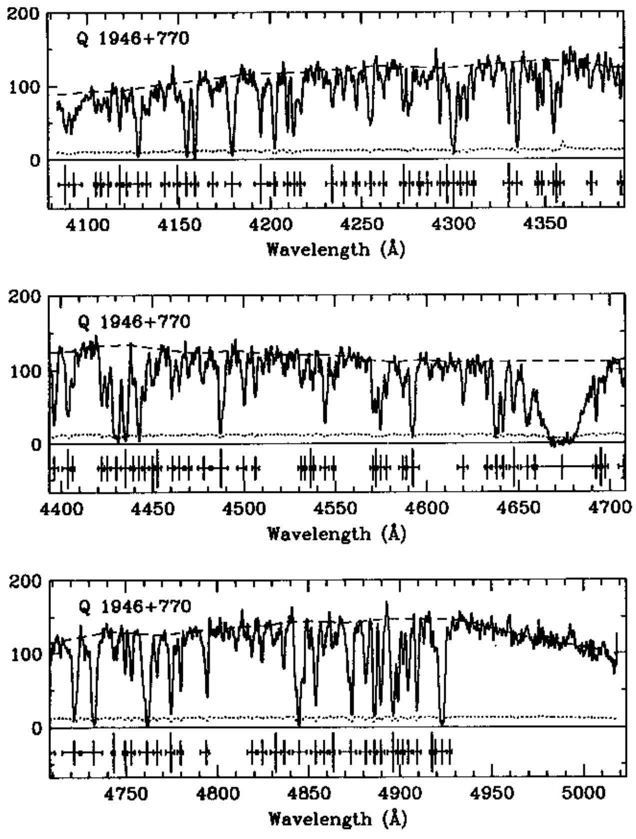




FIG. 1—*Continued*

lines were identified following the procedure outlined in B94. Table 2 lists the wavelengths, equivalent widths, and identifications of all absorption lines found in our spectra. Because of its large size, only a portion of the table is included in this version of the paper. The table will be published subsequently in its entirety on CD-ROM and can be requested from the authors.

## 3. SAMPLE

The sample used in the analysis is a subset of the sample used in B94. The idea was to create a large, very homogeneous sample consisting of high signal-to-noise ratio spectra, all with similar resolution. The B94 sample had an additional 10 spectra fulfilling these criteria. In total, we have spectra of 17 quasars, with emission redshifts ranging from 2.526 to 3.384. Sixteen of them were obtained with the MMT, and one (Q1548+092) was obtained with the 5 m Hale telescope on Mount Palomar. The sample is comparable in size to the sample of Bajtlik et al. (1988), but it is by far more homogeneous.

In all cases, regions of spectra containing known metal system absorption lines were excluded from the analysis. In order to remove any influence from the proximity effect on the results, we considered only regions with redshifts smaller than $z_{em} - 0.15$. Regions shortward of the Ly$\beta$ emission line were also excluded. The sample is described in Table 3.

## 4. REDSHIFT EVOLUTION OF CLOUDS: FLUX DEFICIT ANALYSIS

The number density evolution of clouds versus redshift cannot be constrained well with the data presented here by itself, owing to the sample's limited redshift range. However,

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



TABLE 2

LINE LISTS[a]

| Number | $\lambda_{obs}$ | $W_\lambda$ (Å) | Identification | Possible Identification |
|---|---|---|---|---|
| Q1206+119 | | | | |
| 1 ........... | 5014.04 | 0.580 ± 0.026 | | |
| 2 ........... | 5011.70 | 2.557 ± 0.049 | Si IV λ1402, z = 2.5725 | |
| 3 ........... | 5000.92 | 0.235 ± 0.043 | | |
| 4 ........... | 4997.76 | 0.542 ± 0.046 | N V λ1242, z = 3.0222 | |
| 5 ........... | 4989.64 | 1.886 ± 0.050 | | |
| 6 ........... | 4983.55 | 0.882 ± 0.043 | N V λ1238, z = 3.0222 | |
| 7 ........... | 4979.85 | 2.441 ± 0.058 | Si IV λ1393, z = 2.5725 | |
| 8 ........... | 4964.89 | 1.038 ± 0.055 | | |
| 9 ........... | 4959.41 | 0.294 ± 0.055 | | |
| 10 ........... | 4955.72 | 0.997 ± 0.045 | N V λ1242, z = 2.9871 | |

[a] Table 2 is published in its entirety in computer-readable form in the AAS CD-ROM Series, Vol. 6.

we can combine our data with the literature and consider the flux deficit between Lyβ and Lyα emission lines, defined as

$$D_A \equiv \left\langle 1 - \frac{f_o}{f_c} \right\rangle , \quad (5)$$

where $f_o$ and $f_c$ are the observed and continuum fluxes, respectively, as a function of redshift (Oke & Korycansky 1982; Jenkins & Ostriker 1991; Zuo & Lu 1993). Zuo & Lu (1993, hereafter ZL) calculated $D_A$ for a large set of quasar spectra and used it to estimate $\gamma$. Their sample included nine of the quasars from our sample. We calculated $D_A$ for all our 17 quasars following the procedure outlined in ZL, in all cases excluding regions containing known metal systems. These values are listed in Table 3.

We added the values of $D_A$ from our quasars to their sample (replacing these quasars that we had in common) and calculated $\gamma$ for three cases: all quasars, ZL sample only, and our sample only. The results are shown on Figure 2. This plot shows $\ln[-\ln(1 - D_A)]$ versus $1 + z_c \equiv 1 + (z_{min} + z_{max})/2$ (see Table 3); as shown by ZL, these two quantities follow a linear relation with $1 + \gamma$ as the proportionality coefficient. In general, quasars in our sample do follow the overall trend described in ZL. Adding our quasars to the ZL sample changes $\gamma$ only very slightly, and the difference is not statistically significant. For all quasars $\gamma_{all} = 2.33 \pm 0.29$, while for the ZL sample only (with our values of $D_A$ replacing theirs for objects that we have in common) $\gamma_{ZL} = 2.51 \pm 0.25$.

Note that some of our quasars, especially Q1315+472, show somewhat more absorption than other quasars in the sample. The most obvious explanation for this would be if for some reason the continuum for these objects has been overestimated. We reexamined the continuum fits for these quasars, and we

TABLE 3

THE SAMPLE

| QSO | $z_{em}$ | $z_{min}$[a] | $z_{max}$[b] | $D_A$[c] | Setup[d] | Reference[e] |
|---|---|---|---|---|---|---|
| 0014+813... | 3.384 | 2.748 | 3.234 | 0.211 | 1 | 1 |
| 0114−089... | 3.205 | 2.680 | 3.055 | 0.190 | 1 | 1 |
| 0256−000... | 3.374 | 2.691 | 3.224 | 0.225 | 1 | 1 |
| 0302−003... | 3.283 | 2.679 | 3.133 | 0.216 | 1 | 1 |
| 0636+680... | 3.174 | 2.522 | 3.024 | 0.236 | 1 | 1 |
| 0831+128... | 2.739 | 2.155 | 2.589 | 0.136 | 1 | 1 |
| 0913+072... | 2.784 | 2.193 | 2.603 | 0.140 | 1 | 1 |
| 1206+119... | 3.108 | 2.466 | 2.958 | 0.221 | 1 | 2 |
| 1215+333... | 2.606 | 2.043 | 2.456 | 0.137 | 1 | 1 |
| 1225−017... | 2.831 | 2.232 | 2.681 | 0.199 | 1 | 2 |
| 1315+472... | 2.590 | 2.029 | 2.440 | 0.229 | 1 | 2 |
| 1334−005... | 2.842 | 2.242 | 2.692 | 0.136 | 1 | 1 |
| 1548+092... | 2.748 | 2.406 | 2.598 | 0.153 | 2 | 1 |
| 1607+183... | 3.134 | 2.488 | 2.984 | 0.217 | 1 | 2 |
| 1623+269... | 2.526 | 1.987 | 2.376 | 0.150 | 1 | 2 |
| 1700+643... | 2.744 | 2.159 | 2.594 | 0.215 | 1 | 2 |
| 1946+770... | 3.020 | 2.392 | 2.825 | 0.169 | 3 | 2 |

[a] Redshift corresponding to the position of the Lyβ emission line in the spectrum in all cases but one. In the case of Q1548+092, the lower limit is determined by $S/N > 5$.

[b] $z_{em} - 0.15$ in all but two cases. For Q0913+072 and Q1946+770, regions with damped Lyα systems are excluded.

[c] Flux deficit $D_A$, as defined in eq. (5).

[d] Instrument setup: (1) MMT, Big Blue Reticon, 832 lines mm$^{-1}$, second order, Image Stacker; (2) Mount Palomar 5 m telescope, 2D-Frutti Photon Counter, 1200 lines mm$^{-1}$, second order, 1″ slit; (3) MMT, Big Blue Reticon, 832 lines mm$^{-1}$, second order, 1″ × 3″ slit.

[e] REFERENCES.—(1) Bechtold 1994; (2) this paper.

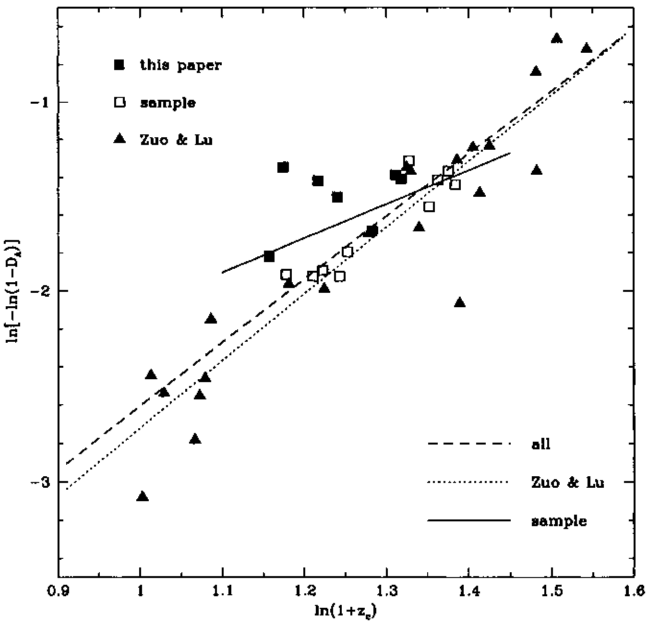



FIG. 2.—$\ln[-\ln(1 - D_A)]$ as a function of $\ln(1 + z_c)$ for quasars in ZL sample (*filled triangles*), our sample (*filled and open squares*), and seven quasars presented in this paper (*filled squares*). Dotted line shows the best fit to ZL sample only; dashed line shows the best-fit to ZL and our samples combined. Solid line is the fit to our 17 quasars only. See text for definitions of parameters and discussion.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



did not see any obvious reason why they were more absorbed than other objects. We note that some of the quasars in the ZL sample deviate from the overall trend by a comparable amount, though in the opposite direction. The result is that if we calculate the line number density evolution for our sample only, then we obtain a much flatter evolution index, $\gamma = 0.80 \pm 0.67$. This agrees well with the determination of $\gamma$ from the line list for our sample by the usual means (e.g., Murdoch et al. 1986; B94), $\gamma = 0.89 \pm 0.61$ for lines with $W > 0.36$ Å.

## 5. SIMULATION METHOD

We estimated the parameters of the distributions of the Lyα clouds using simulated absorption spectra. We constructed a grid of models varying the cloud distribution parameters, $\mathscr{A}_0$, $\gamma$, $\beta$, $\langle b \rangle$, and $\sigma_b$. Instead of $\mathscr{A}_0$, we use an equivalent parameter, $\mathscr{N}_3$, defined as the number of Lyα absorption lines stronger than $W_{\rm rest} = 0.36$ Å, which a hypothetical quasar with emission redshift equal to 3 would have between Lyβ and Lyα emission lines. One can easily show from equation (1) that

$$\mathscr{N}_3 = \mathscr{A}_0(1+\gamma)^{-1}\left[4^{1+\gamma} - \left(\frac{27}{8}\right)^{1+\gamma}\right]. \quad (6)$$

The virtue of this parameter is that it is "$\gamma$ independent"; it remains the same when running simulations with different $\gamma$, while using $\mathscr{A}_0$ would require a nonintuitive adjustment for each $\gamma$.

For each set of parameters and for each quasar in the observed sample, its simulated duplicate was generated. Thus, each set of simulations consisted of 17 "quasar" spectra, having identical emission redshifts, binning, continuum level, instrumental resolution, and signal-to-noise ratios as the actual observed spectra. First we constructed an absorption line list, drawing column density and Doppler parameter according to equations (3) and (4) and redshift using equation (1). All line parameter values were drawn from their respective distributions independently from one another; in particular, there was no correlation between line column density and $b$ value. Then an "observed" spectrum was generated in the following way: we took the pixel wavelengths and continuum level from the observed spectrum and calculated the deficit in all pixels from all absorption lines in the list, using calculated Voigt profiles. As a result, we had an "ideal" spectrum, with infinite signal-to-noise ratio and perfect resolution, sampled like the real spectrum. This spectrum was then convolved with a Gaussian function with FWHM corresponding to the instrumental resolution. After that noise was imposed on each pixel. We estimated the noise by interpolating between the signal-to-noise ratios in the spectrum and in the bottom lines in the corresponding observed spectrum, and then generating its value from a Gaussian distribution with this dispersion. We used the observed signal-to-noise ratios as a function of wavelength, which had been calculated and preserved during data reduction.

The ranges of the input parameters for the simulations were as follows.

*Number density redshift evolution.*—We ran the simulations for seven input values of $\gamma$, ranging from 0.5 to 3.5 with intervals of 0.5. We wanted to bracket the acceptable range of $\gamma$. The lowest value, 0.5, corresponds to nonevolving number density of clouds in a flat universe. The highest is safely higher than the largest value of $\gamma$ reported in the literature, which is near 3.

*Number density normalization.*—We used six values of $\mathscr{N}_3$, from 20 to 80 in intervals of 10. This range surrounds the values of $\mathscr{N}_3$ found in the literature, which give $\mathscr{N}_3$ near 35–40 (see § 7).

*Column density.*—We used 15 input values of $\beta$, from 1.10 to 1.80 with intervals of 0.05, which should span the actual value.

An important issue is the maximum and minimum values for the column density, $N_{\rm min}$ and $N_{\rm max}$. We used $N_{\rm max} = 10^{16}$ cm$^{-2}$ for the high cutoff value, which corresponds to the upper limit for the column density of the Lyα forest lines. This limit is of lesser importance, since the distribution is a steep function of column density and there are always very few lines near the high-$N$ end. The most important factor is the low-$N$ end.

We tested three values of $N_{\rm min}$: $10^{12.5}$, $10^{13}$, and $10^{13.5}$ cm$^{-2}$, before settling for the middle one. The high value led to obvious lack of weak lines in the spectra ($N = 10^{13.5}$ cm$^{-2}$ corresponds to the lowest rest equivalent widths for lines of the order of 0.17 Å). On the other hand, when $N_{\rm min}$ was set to $10^{12.5}$ cm$^{-2}$, there were so many weak lines that they totally dominated the spectra, creating huge blends and virtually eating out all flux in the continuum. There are some indications that the H I column density power-law distribution was a double power law (Petitjean et al. 1993, and references therein), flattening for low $N$, but this was not well established at the time we carried out the simulations. We used a single power law, which is a reasonable first approximation.

*Doppler parameter mean value.*—Six input values of $\langle b \rangle$ were used, from 20 to 45 km s$^{-1}$ with intervals of 5 km s$^{-1}$.

*Width of Doppler parameter distribution.*—We tested two values of $\sigma_b$, 10 and 20 km s$^{-1}$. Negative Doppler parameters were not allowed. It turned out that our analysis did not favor either of the two distributions, the reason probably being that the spectral resolution of the data is $\sim 50$ km s$^{-1}$.

## 6. COMPARISON OF SIMULATIONS WITH REAL SPECTRA

In order to compare the simulated spectra to the real ones, we consider lists of the "dispersion elements" with significant absorption. For each simulated and observed spectrum, we created these lists following the procedure described in Young et al. (1979). We calculated equivalent width of absorption and its uncertainties in bins of width equal to 2.5 times the FWHM instrumental resolution and flagged all bins in which absorption was significant at the 5 $\sigma$ level. Although in such a case neither the simulations nor the observations are free from blending (the resolution elements with significant absorption very often contain more than one "true" line), they are both affected in the same way. The contiguous resolution elements flagged were then combined and counted as one "absorption bin."

We then compared the samples of $W_r$, equivalent widths of absorption bins, and the samples of $x_r$, the separations between these bins, scaled to the local average line separation, in the observed and simulated data. We derive grids of test results $P_{W_r}(\gamma, \mathscr{N}_3, \beta, \langle b \rangle, \sigma_b)$ and $P_{x_r}(\gamma, \mathscr{N}_3, \beta, \langle b \rangle, \sigma_b)$, $P$ being the Kolmogorov-Smirnov (K-S) probability that the two considered samples are drawn from the same distribution. Note that such a test takes into account not only the strong lines, but the weak ones as well. In principle, we can estimate model free parameters by comparing respective distributions and finding a region in the parameter space in which these distributions agree. In reality, however, some of the parameters turned out to be poorly constrained (see below).

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



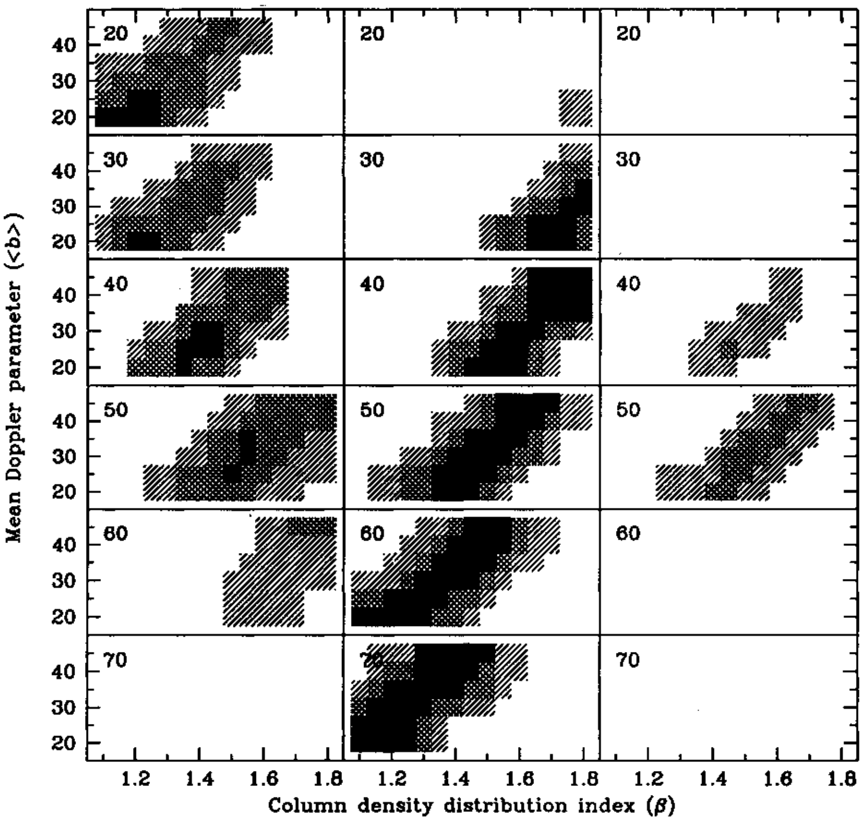



FIG. 3.—Agreement regions for the observed and simulated distributions. The points on the simulation grid are those in which the K-S probability that the two samples were drawn from the same distribution was higher than 0.0003 (*single-hatched*), 0.05 (*cross-hatched*), and 0.33 (*solid filled*). *Left*: agreement regions for equivalent widths of absorption bins; *middle*: agreement regions for scaled separations between bins; *right*: the product of the two. Numbers in upper left corners of each plot indicate the value of $\mathcal{N}_3$ used in the simulations.

Figure 3 shows a plot of $P_{W_r}$ and $P_{x_r}$ for one choice of parameters, $\sigma_b = 20$ km s$^{-1}$ and $\gamma = 1.5$. This plot presents the regions on the $\beta$-$\langle b \rangle$ plane in which the difference between the distribution from simulations and the observed distribution was less than 3, 2, and 1 $\sigma$, or, more precisely, it shows the points on the simulation grid in which the K-S probability that the two samples were drawn from the same distribution was higher than 0.0003, 0.05, and 0.33, respectively. The left panel shows the agreement regions for $W_r$, the middle panel shows the agreement regions for $x_r$, and the right panel shows the product of the two. The numbers in the upper left corners of each plot indicate the value of $\mathcal{N}_3$.

We found that we could locate the set of parameters which best describes the data fairly quickly, given the following considerations. Suppose that for a given $\mathcal{N}_3$ we have found a region on the $\beta$-$\langle b \rangle$ plane in which the observed and simulated distributions agree (or "agreement region," as we will refer to such regions in what follows). Let us now consider what happens if we increase the number of strong lines (i.e., $\mathcal{N}_3$) in our simulations.

First, consider the distribution of the equivalent widths of absorption bins. If the number of the strong lines in the simulated spectra increases, then the simulated and observed distributions can agree only if some weak lines are added as well. However, adding new lines to the spectra increases the number of occurrences of strong blends, which "swallow" some weak lines. In order to compensate for that loss, a large number of weak lines has to be added, which translates to high $\beta$. As a result, we expect that for larger $\mathcal{N}_3$ the agreement region for the $W_r$ distributions will be shifted toward larger values of the column density distribution index, $\beta$.

Second, consider the distribution of the scaled line separations, $x_r$. Adding lines to the spectra results in a tendency for large line separations to disappear. In order to compensate for that effect and preserve a good agreement between the observed and simulated distributions, one has to add as few lines as possible, which means that only a few weak lines are added. However, since the number of strong lines added is set for a given simulation, the only room for maneuvering is in limiting the number of weak lines. That means that increasing the number of strong lines in the simulated spectra will shift the agreement region toward smaller values of $\beta$.

Note that the shapes of lines have virtually no influence on this distribution, since it does not matter whether the line which limits a given separation is broad or narrow. Therefore, we do not expect the distribution of $x_r$ to be a strong function of $\langle b \rangle$ or $\sigma_b$.

The two trends described above go in opposite directions (see Fig. 3), which means that finding an area on the $\beta$-$\langle b \rangle$ plane in which two agreement regions (for $x_r$ and $W_r$) coincide will allow us to estimate $\mathcal{N}_3$, while the shape of the region should give us an idea of how well $\beta$ and $\langle b \rangle$ are constrained.

## 7. RESULTS

First of all, our results are insensitive to the value of $\gamma$ used. We recall that we ran the simulations for seven values of $\gamma$, and

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



only the extreme cases, 0.5 and 3.5, gave slightly lower probabilities (but the difference was statistically insignificant) than the other five cases. Although poor limits on $\gamma$ are not surprising in view of the fact that our sample has a limited redshift range, we found it encouraging that this test, completely independent of the tests described earlier, produced a sensible outcome. Therefore, we adopted $\gamma = 1.5$. The results for $\sigma_b = 10$ and 20 km s$^{-1}$ turned out to be basically indistinguishable. This is also not surprising, since we do not expect the data used here to strongly constrain the Doppler parameter distribution.

In Figure 3, the agreement regions for $x_r$ and $W_r$ coincide for $\mathcal{N}_3 = 45 \pm 10$. This agrees with what we expect from the distributions obtained from line lists. If we translate the best-fit $\mathcal{A}_0$ for $W_{\rm thr} = 0.36$ Å from the sample in Table 3 into $\mathcal{N}_3$ (eq. [6]), we obtain $\mathcal{N}_3 = 38 \pm 30$. These two values are consistent but within large uncertainties. For the sample of Bajtlik et al. (1988) (all parameter values quoted below are as recalculated by Frye et al. 1993), $\gamma$ was equal to $2.28 \pm 0.42$ with $\mathcal{A}_0 = 3.4$, which translates to $\mathcal{N}_3 = 42 \pm 23$. The sample of Lu et al. (1991) had $\gamma = 2.61 \pm 0.29$ with $\mathcal{A}_0 = 2.1$, which is equivalent to $\mathcal{N}_3 = 40 \pm 15$. The echelle sample of Rauch et al. (1992) had $\gamma = 1.50 \pm 0.71$ and $\mathcal{A}_0 = 8.5$, resulting in $\mathcal{N}_3 = 38 \pm 35$. Finally, B94 had $\gamma = 1.85 \pm 0.27$ with $\mathcal{A}_0 = 5.04$, which translates to $\mathcal{N}_3 = 35 \pm 12$.

Clearly, in this respect, the number of strong lines per unit redshift, our results are consistent with the determinations obtained with the use of the line lists. Although the latter tend to give smaller values for $\mathcal{N}_3$, the differences are not significant.

## 8. DISCUSSION

Our best estimate for the power-law index of the cloud column distribution is $\beta = 1.4 \pm 0.1$. This result is quite interesting. Basically all values of $\beta$, based on profile fitting analysis of the Ly$\alpha$ forest lines, are higher, $\beta = 1.7$–$1.9$ (e.g., Petitjean et al. 1993, and references therein). However, indirect studies (Webb et al. 1992; Chernomordik & Ozernoy 1993; Press & Rybicki 1993) indicated that $\beta$ may in fact be lower, $\beta = 1.3$–$1.4$. In addition, the photoionization model for the proximity effect (Bajtlik et al. 1988) predicted that the density of observed Ly$\alpha$ lines in "coevolving redshift," defined as $X_\gamma \equiv \int (1 + z)^\gamma dz$, should be $d\mathcal{N}/dX_\gamma = \mathcal{A}_0(1 + \omega)^{1-\beta}$ (their eq. [25]), where $\omega$ is the Lyman-limit flux density ratio (quasar flux/background). As noted by B94, the influence of $\beta$ on this result is quite strong, and, in fact, a low $\beta$ leads to a better fit to the proximity effect data (see Fig. 33 in B94).

There are two explanations for this discrepancy which come to mind: (1) line blending and (2) the behavior of the distribution at the low end of the cloud column density range. Line blending would explain the discrepancy if for some reason profile fitting–based analyses had a tendency to "overdeblend" strong lines, or if indirect analyses tended to neglect the influence of weak lines.

On the other hand, the discrepancy can be well explained by flattening of the H I column density distribution at low $N$. All studies which yielded high values of $\beta$ were done on samples with lower limits on log $N$ near 13.7 cm$^{-2}$. It was already suggested by Carswell et al. (1987) and Petitjean et al. (1993) that below that value there may be some flattening of the H I column density distribution, which was recently confirmed by the observations with HIRES and the Keck telescope (Tytler 1995). Tytler (1995) reports that the distribution is indeed a double power law, with a break point at $N_{\rm break} \approx 10^{13.8}$ cm$^{-2}$, with $\beta \approx 1.5$ for $N \gtrsim N_{\rm break}$ and $\beta \approx 1.2$ for $N \lesssim N_{\rm break}$. Our result is in excellent agreement with these findings. If one approximates Tytler's result with a single power law between $10^{13}$ and $10^{16}$ cm$^{-2}$, the resulting power-law index is indeed close to 1.4. This indicates that our method is capable of discovering effects hidden below the resolution of the sample.

As expected, our analysis does not yield strong constraints on the distribution of the Doppler parameter $b$. The distribution of separations between absorption bins turned out to be basically insensitive to $b$, and the only information about its distribution is buried in the shape of the agreement region for the equivalent widths of the absorption bins. This result is consistent with previous determinations of typical Doppler parameter for Ly$\alpha$ lines (e.g., Oke & Korycansky 1982; Carswell et al. 1991; Press & Rybicki 1993).

The agreement regions for both equivalent width and separation distributions appear skewed with respect to the plot axes (Fig. 3). This is not an indication of a correlation between column density and Doppler parameter, since we did *not* introduce such a correlation in the simulations (which would be the only way we could make such claims). This effect can be easily explained: suppose that for given and fixed $\mathcal{N}_3$, we find a good agreement between the observed and simulated distributions for a certain pair of $\beta$ and $\langle b \rangle$. If one increases $\beta$ in the simulations (more weak lines), one will obtain a better agreement for the distribution of $W_r$ if $\langle b \rangle$ is increased as well (lines are thicker). We should observe the same effect for the distribution of separations: an increased number of weak lines is going to create excess of small separations, which has to be countered by these lines being broader, which results in a higher occurrence of blends.

## 9. LINE BLENDING

It is interesting to look at a simulated spectrum, and we show an example in Figure 4. It is a very typical simulation, and it was *not* chosen as having any special features. This particular simulation was aimed at reproducing the spectrum of Q1607+183 with the best-fit parameters described above (see Fig. 1*d*). This spectrum clearly shows the blending problems that can be encountered in the analysis of the line list.

We compared the traditional line fitting analysis for this spectrum and several others. To generate an "observed" line list, we fitted Gaussian absorption lines with the code FINDSL, kindly provided to us by T. Aldcroft. On Figure 4 there are two set of tick marks below the spectrum. The lower ticks indicate the positions of all input absorption lines in the simulation, before blending and noise have been imposed. The upper ticks mark the lines detected with FINDSL for the 3 $\sigma$ significance limit. One can clearly see that there are numerous unrecognizable blends. In this particular case, the final number of detected lines is equal to only 44% of the input lines (172 out of 391; note that few input lines are too weak to be detected in this spectrum). We found that in our simulated spectra this ratio was usually somewhere between 40% and 60%.

## 10. SUMMARY

The spectra of seven quasars presented here form the completion of the large sample of Ly$\alpha$ forest spectra of quasars, which was analyzed in depth in B94. In the present paper, we combined these data with 10 other high S/N spectra from the B94 sample and presented a nontraditional method of the analysis of the data.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJ...457..102D



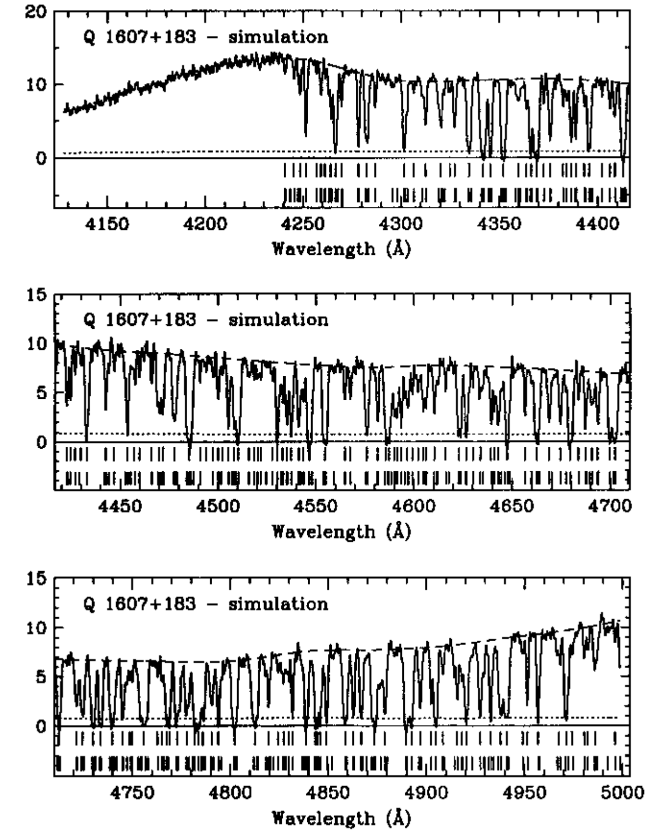



FIG. 4.—Simulated spectrum of Q1607+183. The code simulates lines between Ly$\beta$ and Ly$\alpha$ emission lines only, hence the lack of lines below 4240 Å. Dashed line indicates the continuum; dotted line indicates 1 $\sigma$ uncertainties. The lower ticks indicate the positions of *input* absorption lines. The upper ticks mark the lines detected with absorption line analysis code. See text for discussion.

We have shown that one can obtain reliable results for the cloud physical parameters using a moderate resolution sample of spectra, but with a high signal-to-noise ratio, virtually eliminating the subjective part of the analysis. Our simulation-based method allowed us to derive distributions of various cloud parameters. We estimated the value of the H I column density distribution (eq. [3]) index, $\beta$, with good accuracy, reaching down to column densities of $10^{13}$ cm$^2$. The limited redshift range of the sample prevented us from estimating a reliable line number density evolution parameter $\gamma$. Also, the instrumental resolution of the sample is such that we can only claim that our results for Doppler parameter distribution are consistent with previous determinations.

We found $\beta = 1.4 \pm 0.1$, which is somewhat lower than the value obtained with the use of line lists. Interestingly, this is in excellent agreement with the result of the similar analysis by Press & Rybicki (1993) of a sample of low-resolution spectra at higher redshift. We think, however, that our sample is better suited for that type of study, since it allows for examination of the spectra and exclusion of regions containing metal and damped Ly$\alpha$ systems from the subsequent analysis, which is practically impossible with low-resolution spectra. Moreover, our result, obtained by nonstandard analysis of intermediate-resolution data, is in excellent agreement with recent results obtained from very high-resolution data.

We have presented a new method of analysis of the Ly$\alpha$ forest lines, which enables one to determine physical properties of the population of Ly$\alpha$ clouds from moderate-resolution but high signal-to-noise spectra. This method can be applied, with practically no adjustments, to much larger sets of data. Extending the redshift range of the sample will enable us to determine the line number density evolution independent of blending. The method is particularly well suited for the analysis of the spectra of high-redshift quasars, for which there are many lines (and, inevitably, blends; see Storrie-Lombardi et al. 1995). Also, one can introduce into the simulations some nonrandom effects such as voids in the redshift distribution of lines. The analysis of such effects will be presented in a future paper in this series.

We would like to thank the MMT staff for their assistance with observations and T. Aldcroft for providing us with his absorption line code FINDSL. A. D. wishes to thank the director and staff of the Steward Observatory for their hospitality during his stay there, when this project was initiated, S. Bajtlik, B. Czerny, D. Eisenstein, O. Kuhn, A. Siemiginowska, and A. Soltan for helpful discussions, and D. Latham for support and allowing the use of his computer at the early stage of this project. This work is supported in part by a Chretien International Research Grant. A. D. acknowledges support from NASA Contract no. NAS 8-39073 (AXAF Science Center), and J. B. was supported by AST 90-58510 from the National Science Foundation, and a gift from Sun Microsystems.

REFERENCES

Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, ApJ, 327, 570
Barcons, X., & Webb, J. K. 1991, MNRAS, 253, 207
Bechtold, J. 1994, ApJS, 91, 1 (B94)
Carswell, R. F., Lanzetta, K. M., Parnell, H. C., & Webb, J. K. 1991, ApJ, 371, 36
Carswell, R. F., Webb, J. K., Baldwin, J. A., & Atwood, B. 1987, ApJ, 319, 709
Chernomordik, V. V., & Ozernoy, L. M. 1993, ApJ, 404, L5
Frye, B. L., Bechtold, J., Moustakas, L. A., & Dobrzycki, A. 1993, MNRAS, 263, 575
Jenkins, E. B., & Ostriker, J. P. 1991, ApJ, 376, 33
Kuhn, O., Bechtold, J., Cutri, R., Elvis, M., & Rieke, M. 1995, ApJ, 438, 643
Latham, D. W. 1982, in IAU Coll. 67, Instrumentation for Astronomy with Large Optical Telescopes, ed. C. M. Humphries (Dordrecht: Reidel), 245
Liu, X. D., & Jones, B. J. T. 1988, MNRAS, 230, 481
Lu, L., Wolfe, A. M., & Turnshek, D. A. 1991, ApJ, 367, 19
Murdoch, H. S., Hunstead, R. W., Pettini, M., & Blades, J. C. 1986, ApJ, 309, 19
Oke, J. B., & Korycansky, D. G. 1982, ApJ, 255, 11
Parnell, H. C., & Carswell, R. F. 1988, MNRAS, 230, 491
Peterson, B. A. 1978, in IAU Symp. 79, The Large Scale Structure of the Universe, ed. M. S. Longair & J. Einasto (Dordrecht: Reidel), 389
Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F., & Lanzetta, K. 1993, MNRAS, 262, 499
Press, W. H., & Rybicki, G. B. 1993, ApJ, 418, 585
Press, W. H., Rybicki, G. B., & Schneider, D. P. 1993, ApJ, 414, 64
Rauch, M., Carswell, R. F., Chaffee, F. H., Foltz, C. B., Webb, J. K., Weymann, R. J., Bechtold, J., & Green, R. F. 1992, ApJ, 390, 387
Schneider, D. P., Schmidt, M., & Gunn, J. E. 1991, AJ, 101, 2004
Siemiginowska, A., Kuhn, O., Elvis, M., Fiore, F., McDowell, J., & Wilkes, B. J. 1995, ApJ, 454, 77
Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J., & Hazard, C. 1995, in Proc. ESO Workshop on QSO Absorption Lines, ed. J. Bergeron, G. Meylan, & J. Wampler (Heidelberg: Springer), in press
Trevese, D., Giallongo, E., & Camurani, L. 1992, ApJ, 398, 491
Tytler, D. 1995, in Proc. ESO Workshop on QSO Absorption Lines, ed. J. Bergeron, G. Meylan, & J. Wamper (Heidelberg: Springer), in press
Webb, J. K., Barcons, X., Carswell, R. F., & Parnell, H. C. 1992, MNRAS, 255, 319
Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F., & Whelan, J. A. J. 1979, ApJ, 229, 891
Zuo, L., & Lu, L. 1993, ApJ, 418, 601 (ZL)

© American Astronomical Society • Provided by the NASA Astrophysics Data System