STATISTICAL PROPERTIES OF THE LYMAN α FOREST

JILL BECHTOLD
STEPHEN A. SHECTMAN
The Mount Wilson-Las Campanas Observatories
The Carnegie Institute of Washington
Pasadena, California 91101

ABSTRACT. We have observed two high-redshift quasars with the echelle spectrograph and 2D-Frutti Photon Counter at Las Campanas, in order to investigate the statistical properties of the Lyman-α forest. The two-point correlation function for the Lyman-α forest lines at z $\approx$ 3 is consistent with zero, for all velocity splittings $\Delta v > 50$ km/sec. When Lyman-α lines and other metal lines from known metal-line systems are included, the correlation function shows a weak non-zero signal at small Δv . We suggest that the weak clustering of the Lyman-α forest detected by other workers may be the result of contamination by a small number of metal-line systems and their associated Lyman-α lines.

1. Introduction

The distribution of quasar absorbing clouds in redshift provides unique information about large scale structure at early epochs. One of the goals of echelle spectroscopy of the Lyman-α forest is to measure the clustering of the clouds for small velocity splittings, for which the high resolution is necessary to deblend closely spaced velocity components. Before presenting our new measurement of the two-point correlation function for the Lyman-α forest clouds at z $\approx$ 3, we review previous observational results. The predicted clustering of clouds in different theoretical models is discussed elsewhere (e.g. Dekel 1982; Salmon and Hogan 1986; Ikeuchi and Ostriker 1986; Vishniac and Bust 1987; Bond, Szalay, and Silk 1988).

2. Previous Results

Based on moderate resolution data, Sargent et al. (1980) showed that the distribution of velocity splittings, Δv, between pairs of Lyman-α forest clouds with redshifts $1.7 < z < 3.3$, is consistent with a random, Poisson distribution, for $300 < \Delta v < 30{,}000$ km/sec. The amplitude of any clustering of the Lyman-α forest is much smaller than what is observed for present-day galaxies or clusters of galaxies.

By contrast, the distribution of velocity splittings for C IV absorbers shows a strong excess of pairs with splittings of $\Delta v \approx 500$ km/sec (Young et al. 1980; Sargent, Boksenberg and Steidel 1988).

Webb (1987) measured the correlation function for the Lyman-α forest clouds for much smaller velocity splittings, $50 < \Delta v < 300$ km/sec, based on high resolution data (resolution ~ 0.25 Å, FWHM). He finds weak clustering, $< \xi > = 0.44 \pm 0.12$, at small splittings (ξ is defined in §4), but only for low redshifts, $1.9 < z < 2.4$. For higher redshifts, $2.4 < z < 2.8$, the distribution is consistent with Poisson, $< \xi > = 0.18 \pm 0.11$.

Crotts (1988) measured the *cross*-correlation function of the Lyman-α clouds in the spectra of pairs of quasars close together in the sky. This technique complements high-resolution studies along single lines-of-sight, since in principle the comparison between the cross-correlation and auto-correlation functions can distinguish between internal velocity splittings within galaxies and clustering between galaxies. Crotts finds a tantalizing, but not statistically significant excess of splittings at $\Delta v \approx 100$ km/sec for separations ~ 0.7 Mpc, with $< \xi > = 1.9 \pm 0.8$. Clearly, this is a worthwhile approach which should be pursued with other pairs of quasars.

3. New Observations and Analysis

In September 1987, we obtained spectra of two high-redshift quasars, PKS 2126-158 (z = 3.26) and Q 0130-403 (z = 3.02), with the echelle spectrograph and 2D-Frutti photon counter at the 100-inch Dupont telescope at Las Campanas (Figure 1). The spectral resolution, through a 2-arcsec slit, is 19 km/sec FWHM. Single exposure wavelength coverage is continuous from 3600Å to 6500Å and the spectra extend, with gaps, to 7500Å. For these quasars, this corresponds to rest-frame wavelengths from just redward of the C IV emission line to blueward of the Lyman limit. Details of the reduction proceedures are presented elsewhere.

The continuum level was determined iteratively, by fitting the data with cubic splines, and throwing out the largest negative deviations until the root-mean-square positive deviations were equal to the negative deviations and were consistent with the deviations expected from the variance. Significant lines were found by searching for features with equivalent width larger than 3.5 times the error in the equivalent width as determined by the variance, assuming that the continuum is precisely determined. Metal systems were found by looking for pairs of lines with the spacing corresponding to various well-known transitions (C IV $\lambda\lambda 1548, 1550$Å; Mg II $\lambda\lambda 2795, 2803$Å; Lyman-$\alpha$ and Si III; etc.) and requiring that doublet ratios and other line ratios be sensible. Finally, we formed a list of Lyman-α forest lines significant above the 5σ level. Because of vignetting in the image tube, the signal-to-noise of the spectra, and hence the equivalent width limit, is a function of wavelength. Below, we consider only those lines with rest equivalent widths greater
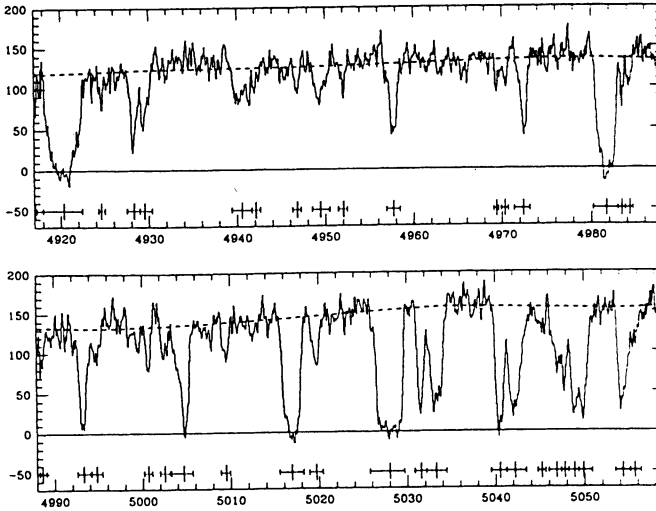
Figure 1. Portion of LCO echelle spectrum of PKS 2126-158. Continuum fit is shown by the dashed line. 3.5-σ lines are indicated below the spectrum.

than 0.15 Å, which is greater than or equal to the 5σ limit everywhere in the region of interest.

4. The Two-Point Correlation Function

With the high resolution spectra described in Section §3, we can measure the 2-point correlation function for Lyman-α forest clouds at velocity splittings as low as 50 km/sec. We define the velocity splitting, Δv, between two clouds with $z_{abs} = z_1$ and $z_{abs} = z_2$ as

$$\Delta v = c\frac{z_2 - z_1}{1+ <z>} \tag{1}$$

where

$$<z> = \frac{z_1 + z_2}{2}$$

and $z_1 < z_2$; Δv is the velocity splitting observed at the cosmic time when the photons passed through the clouds.

The correlation function, ξ, is defined as (*c.f.* Peebles, 1980)

$$\xi = \frac{N_{obs}(\Delta v)}{N_P(\Delta v)} - 1 \tag{2}$$

where $N_{obs}(\Delta v)$ is the number of pairs observed with velocity splitting Δv, and $N_P(\Delta v)$ is the number of pairs expected from a random distribution.

In Figure 2 we plot $\xi(\Delta v)$ for PKS 2126-158 and Q 0130-403 from the echelle spectra described in §3. The errorbars are plotted assuming that the error in a bin with N_{obs} pairs of lines is $(N_{obs})^{-1/2}$.

Figure 2a shows $\xi(\Delta v)$ for the Lyman-α forest lines between emission Lyman-α and Lyman-β (*i.e.* only those Lyman-α lines which have no associated metal-line absorption detected). For all $\Delta v \geq 50$ km/sec, $\xi(\Delta v)$ is consistent with zero, with $<\xi> = 0.1 \pm 0.1$ for the bin with 50 km/sec $< \Delta v <$ 100 km/sec.

In Figure 2b, we plot $\xi(\Delta v)$ for *all* the lines between Lyman-α and Lyman-β (Lyman-α lines plus metal-lines). A suggestive non-zero signal, significant at the 2-σ level, is evident in Figure 2b. Perhaps some of the previously reported clustering of the Lyman-α forest lines results from contamination by un-identified metal systems, which are known to cluster strongly. This result illustrates the importance of obtaining spectra which extend redward to C IV in emission, in order to identify the metal-line redshift systems. Of course, our results are in agreement with Webb's *at the same redshift* , and it could well be that the clustering of the Lyman-α forest he sees at lower redshifts is real.

5. Conclusions

It may seem surprising that the two-point correlation function for the Lyman-α forest is so nearly zero, since one's general impression, looking at spectra, is that there are clumps and voids in the distribution. This impression has lead some people to suggest that the two-point correlation function is not the appropriate tool to measure clustering, and some other statistic should be sought.

A related issue is the significance of apparent "voids" in the Lyman-α forest, which has been the subject of much spirited debate recently (Crotts 1988; Duncan, Ostriker and Bajtlik 1988; and references therein). Here, we make the following qualitative remark. Figure 3 is a plot of lines whose relative positions were determined by a random number generator. Clearly, there are seemingly large gaps and clumps in the distribution. The density of lines is far from *uniform* , even though the distribution of lines is random.

As Gould (1988) has pointed out in a different context, the average person "has no feel for the frequency and length of sequences in random data". As an example, he cites studies which have shown that no streak or slump in baseball has ever occurred "above and beyond the frequency predicted by coin-tossing models", with the important exception of Joe Dimaggio's 56 game hitting streak. Gould comments, "We see pattern, for pattern surely exits, even in a purely random world ... Our error lies not in the perception of pattern but in automatically imbuing pattern with meaning, especially with meaning that can bring us comfort, or dispel confusion."

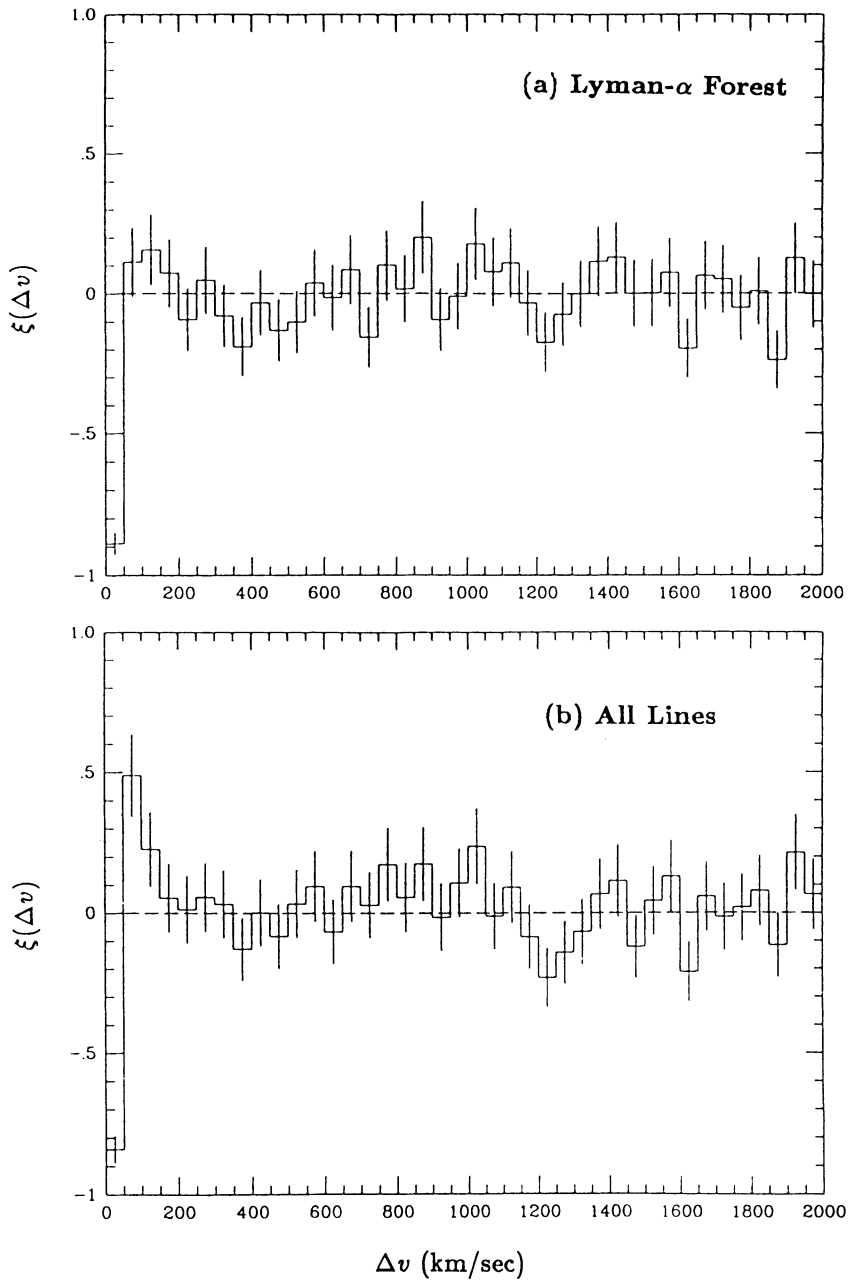## References

Crotts, A.P.S. 1988, preprint.

Figure 2. Two-point correlation function, $\xi(\Delta v)$. (a) Lyman-α forest lines only. (b) All lines between emission Lyman-α and Lyman-β.

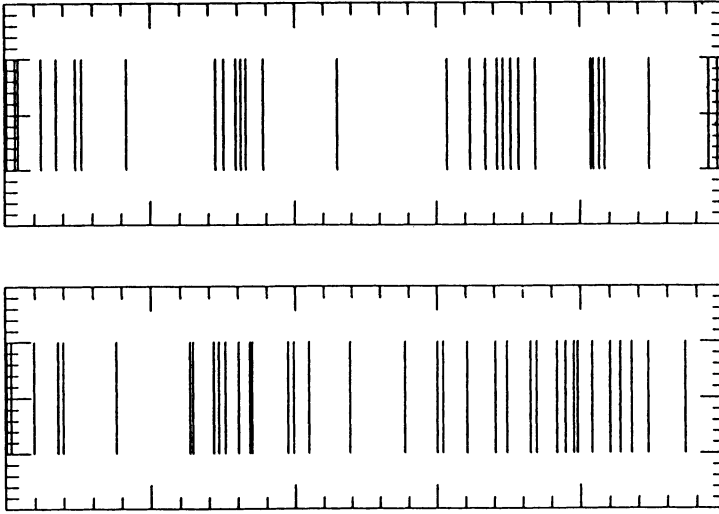**Figure 3.** Randomly placed lines.

Dekel, A. 1982, *Ap.J.(Letters)*, **261**, L13.

Duncan, R.C., Ostriker, J.P., and Bajtlik, S. 1988, preprint.

Ikeuchi, S. and Ostriker, J.P. 1986, *Ap.J.*, **301**, 522.

Gould, S.J. 1988, "The Streak of Streaks", *The New York Review of Books* , **XXXV** , No. 13, p.8.

Peebles, P.J.E. 1980. *The Large Scale Structure of the Universe* , Princeton University Press, Princeton..

Salmon, J. and Hogan, C. 1986, *M.N.R.A.S.*, **221**, 93.

Sargent, W.L.W., Young, P.J., Boksenberg, A., and Tytler, D. 1980, *Ap.J.Suppl.*, **42**, 41.

Sargent, W.L.W., Boksenberg, A., and Steidel, C. 1988, preprint.

Vishniac,E.T. and Bust,G.S. 1987, *Ap.J.*, **319**, 14.

Webb, J.K. 1987, in *Proc. IAU Symp. 124, "Observational Cosmology"* eds. A. Hewitt, G. Burbidge, and L. Fang (Dordrecht:Reidel), p.803.

Young, P.J., Sargent, W.L.W., and Boksenberg, A. 1982, *Ap.J.Suppl.*, **48**, 455.