SIZES, SHAPES, AND CORRELATIONS OF LYMAN ALPHA CLOUDS AND THEIR EVOLUTION IN THE $\Lambda$CDM UNIVERSE

RENYUE CEN$^{1,2}$ AND ROBERT A. SIMcoe$^1$

Received 1996 June 10; accepted 1997 January 27

ABSTRACT

This study analyzes the sizes, shapes, and correlations of Ly$\alpha$ clouds produced by a hydrodynamic simulation of a spatially flat CDM universe with a nonzero cosmological constant ($\Omega_0 = 0.4$, $\Lambda_0 = 0.6$, $\sigma_8 = 0.79$) over the redshift range $2 \leq z \leq 4$. The Ly$\alpha$ clouds range in size from several kiloparsecs to about a hundred kiloparsecs in proper units, and they range in shape from roundish, high column density regions with $N_{H_1} \geq 10^{13}$ cm$^{-2}$ to low column density sheet-like structures with $N_{H_1} \leq 10^{13}$ cm$^{-2}$ at $z = 3$. The most common shape found in the simulation resembles that of a flattened cigar. The physical size of a typical cloud grows with time roughly as $(1 + z)^{-3/2}$, while its shape hardly evolves (except for the most dense regions with $\rho_{cut} > 30$). Collectively, the clouds form large networks of filaments and sheets. Our result indicates that any simple model with a population of spheres (or other shapes) of a uniform size is oversimplified; if such a model agrees with observational evidence, it is probably only by coincidence. We also illustrate why the use of pairs of quasar sight lines to set lower limits on cloud sizes is useful only when the perpendicular sight line separation is small ($\Delta r \leq 50 \, h^{-1} \, $kpc). Finally, we conjecture that high column density Ly$\alpha$ clouds ($N_{H_1} \geq 10^{13}$ cm$^{-2}$) may be the progenitors of the lower redshift faint blue galaxies, based on consideration of their correlation, number density, and mass.

Subject headings: cosmology: theory — hydrodynamics — intergalactic medium — large-scale structure of universe — methods: numerical — quasars: absorption lines

1. INTRODUCTION

The theories of structure formation that remain in contention are those whose parameters have been tuned to match the most up-to-date cosmological observations. At high redshift, the current leading constraint on models is the data from COBE (Smoot et al. 1992), which fixes the amplitude of the power spectrum on very large scales ($\sim 1000 \, h^{-1} \, $Mpc) to an accuracy of about 12%. At low (essentially zero) redshift, we demand that models fit current observations of our local universe, primarily those concerning the distributions of galaxies in $(x, \, t)$ space. These observations include the abundance of clusters of galaxies, which fixes the amplitude of the power spectrum on scales of $\sim 8 \, h^{-1} \, $Mpc to about 10% accuracy (Bahcall & Cen 1992, 1993; Oukbir & Blanchard 1992; White, Efstathiou, & Frenk 1993a; Viana & Liddle 1995; Bond & Myers 1996; Eke, Cole, & Frenk 1996; Pen 1997), the power spectrum of galaxies, which constrains the shape of the power spectrum on the intermediate to large scale of $\sim 10$–$100 \, h^{-1} \, $Mpc (Peacock & Dodds 1994; Feldman, Kaiser, & Peacock 1994), and the ratio of gas to total matter in galaxy clusters, which determines $\Omega_b/\Omega_m$ (White et al. 1993b). In addition, the current measurements of the Hubble constant (Fukugita, Hogan, & Peebles 1993; Freedman et al. 1994; Riess, Press, & Kirshner 1995; Hamuy et al. 1995) and the age constraint from the oldest globular clusters (Bolte & Hogan 1995) limit the range for $H_0$ and the combination of $\Omega_m$, $\Lambda_0$, and $H_0$.

This observational suite has been examined by Ostriker & Steinhardt (1995) to constrain flat cold dark matter models with a nonzero cosmological constant (Peebles 1984; Efstathiou, Bond, & White 1992; Bahcall & Cen 1992; Kohman, Gnedin, & Bahcall 1993; Cen, Gnedin, & Ostriker 1993). The exercise is repeatable for other models (not necessarily flat), including the tilted cold dark matter model (Cen et al. 1992; Liddle, Lyth, & Sutherland 1992; Lidsey & Coles 1992; Adams et al. 1993; Lucchin, Matarrese, & Mollerach 1993), the mixed dark matter model (Davis, Summers, & Schlegel 1992; Taylor & Rowan-Robinson 1992; Klypin et al. 1993; Cen & Ostriker 1994; Ma & Bertschinger 1994), the open cold dark matter model (Gott 1982; Bucher, Goldhaber, & Turok 1995), and the primeval isocurvature baryon model (Peebles 1987a, 1987b; Cen, Ostriker, & Peebles 1993). The allowed parameter space for the family of Gaussian cosmological models with the tunable parameters ($H_0$, $\Omega_m$, $\Lambda_0$, $R_{HC}$, $n$, ISO, ADIA) is thus quite limited. In this context, $H_0$, $\Omega_m$, and $\Lambda_0$ are the Hubble constant, the density parameter, and the value of cosmological constant, all in the present epoch, $R_{HC}$ is the mass ratio of hot to cold matter, $n$ is the asymptotic power spectral index on large scales, and ISO and ADIA denote isocurvature and adiabatic models. While it is possible to further tighten this parameter space simply by making more accurate observations of the aforementioned quantities, other independent tests could help distinguish between competing models, particularly at low to moderate redshift, for which more observations are available.

The Ly$\alpha$ forest lines observed in the spectra of high-redshift quasars have two unique qualities that distinguish them from other observationally accessible phenomena at redshifts between the epoch observed by COBE and our local universe. First, each line of sight indiscriminately samples the distribution of neutral hydrogen gas over a wide redshift range ($z \sim 0$–5) along random lines of sight (i.e., foreground objects are unrelated to the background quasar). Second, each individual spectrum contains a large amount of information about low-absorption regions (e.g.,
voids, fluctuating Gunn-Peterson absorption) as well as information from high-absorption “cloud” lines (number densities of clouds at different redshifts, b-parameters of the individual clouds, correlations of the clouds, relationship with other cosmic entities). With the proliferating number of observed quasars, the total amount of information available is very large, allowing for detailed statistical studies (e.g., Carswell et al. 1991; Rauch et al. 1993; Petitjean et al. 1993; Schneider et al. 1993; Cristiani et al. 1995; Hu et al. 1995; Tytler et al. 1995). These facts, taken together, suggest that the Lyα forests constitute perhaps the most rich and unbiased sample available for studying the universe at moderate redshift.

Recent cosmological hydrodynamic simulations by several independent groups have consistently shown that Lyα clouds are an integral part of the cosmic structure, resulting naturally from the gravitational growth and/or collapse of density fluctuations on small to intermediate scales (~ 100 kpc to a few Mpc in comoving units, Cen et al. 1994, hereafter CMOR; Zhang, Anninos, & Norman 1995; Hernquist, Katz, & Weinberg 1996). Several simple population models have been designed to examine the (local) physical as well as (global) statistical properties of Lyα clouds (Sargent et al. 1980; Ostriker & Ikeuchi 1983; Ikeuchi & Ostriker 1986; Bahcall & Spitzer 1969; Arons 1972; Rees 1986; Ikeuchi 1986; Bond, Szalay, & Silk 1988). One of the essential simplifications in almost all these models is to assume that individual Lyα clouds are spherical. While the clouds produced in the new simulations may have some physical properties in common with these simple models, a visual inspection of the new simulation results reveals that the Lyα absorbing structures resulting from small to intermediate-scale structure formation at high redshift are far from spherical. Furthermore, they seem to have a wide range of sizes. We presented these results first in CMOR and then in much more detail in Miralda-Escudé et al. (1996, hereafter MCOR). Here, we present a quantitative study of the topological aspects of the Lyα clouds, complementing the topics covered in CMOR and MCOR.

In the redshift range $z = 2-4$, Lyα clouds exhibit a rich spectrum of structure, ranging in shape from quasispherical to filamentary and even sheet-like. The spherical structures often reside at density maxima (with $N_{HI} \geq 10^{15}$ cm$^{-2}$) located in the centers of extended structures, while the filaments and sheets tend to have a low $N_{HI} \leq 10^{15}$ cm$^{-2}$ and form a web-like, interconnecting network covering a large portion of the simulation box ($L = 10 h^{-1}$ Mpc comoving).

We also show that it is not unusual for pairs of quasar sight lines to contain absorption features at only slightly separated wavelengths. In particular, we show that “common” absorption features (coincident lines) in sight line pairs separated by 100 $h^{-1}$ proper kpc or more are probably due to absorption by different clouds, while for separations smaller than 40 $h^{-1}$ kpc, pairs of sight lines are mostly likely to actually pierce a common cloud.

Finally, we show that Lyα clouds are spatially significantly clustered, with a correlation length of roughly 1–2 $h^{-1}$ Mpc comoving (for the high column density clouds of $N_{HI} \geq 10^{15}$ cm$^{-2}$) at redshift $z \sim 0.5–1.0$. This correlation, in conjunction with their number density and mass, suggests the possibility of an intriguing connection to faint blue objects (Koo 1986; Tyson 1988; Cowie et al. 1988).

This paper is organized as follows. Some brief descriptions of the simulations (for more details see MCOR) and our cloud identification method are presented in § 2. Results and conclusions are given in § 3 and § 4.

2. SIMULATIONS AND CLOUD IDENTIFICATION

2.1. Simulations

We simulate the formation of Lyα clouds in a spatially flat cold dark matter universe with a cosmological constant ($\Lambda$CDM), using the following cosmological parameters: $H_0 = 65$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_{CDM} = 0.3645$, $\Lambda_0 = 0.6$, $\Omega_{b,0} = 0.0355$ (see Walker et al. 1991), and $\sigma_8 = 0.79$ (the simulations we use in this paper are the same as those used in CMOR and MCOR). The primary motivation for choosing this model is that it is the best fit to the available observations as summarized in the introduction. The simulation box size is $10 h^{-1}$ Mpc comoving on a side and contains $N = 288^3$ cells and $144^3$ dark matter particles. The cell size is $35 h^{-1}$ kpc comoving, corresponding to a average baryonic cell mass of $6.3 \times 10^5 M_\odot$, with the true spatial and mass resolutions being about 2 and 8 times worse than those values, respectively. At $z = 3$, the Jeans length, $\lambda_J \approx (\pi c_s^2/G \rho_m)^{1/2}$ for $c_s = v_{rms} = 10$ km s$^{-1}$, is equal to 400 $h^{-1}$ kpc in comoving units, or 11 cells. The power spectrum transfer function is computed using the method described in Cen, Gnedin, & Ostriker (1993). We use a new shock-capturing, total variation diminishing (TVD) cosmological hydrodynamic code, as described by Ryu et al. (1993).

All the atomic processes for a plasma of (H, He) of primordial composition (76%, 24%) in mass are included using the heating, cooling, and ionization terms described in Cen (1992). We calculate self-consistently the average background photoionizing radiation field as a function of frequency, assuming that the radiation field is spatially uniform (i.e., optically thin). The evolution of the radiation field is calculated given the average attenuation in the simulated box and the emission (both from the gas itself and from the assumed sources of ionizing photons). The time-dependent equations for the ionization structure of the gas are solved by iteration, using an implicit method to avoid the instabilities that arise in solving stiff equations. In general, the abundances of different species are close to ionization equilibrium between recombination and photoionization after most of the gas has been photoionized.

We model galaxy formation as in Cen & Ostriker (1992, 1993a, 1993b). The material turning into collisionless particles as “galaxies” is assumed to emit ionizing radiation, with two types of spectra: one characteristic of star forming regions and the other characteristic of quasars, with efficiencies (i.e., fraction of rest-mass energy converted into radiation) of $e_{UV,\star} = 5 \times 10^{-6}$ and $e_{UV, Q} = 6 \times 10^{-6}$, respectively. We adopt the emission spectrum of massive stars from Scalo (1986) and that of quasars from Edelson & Malkan (1986). Details of how we identify galaxy formation and follow the motions of formed galaxies have been described in Cen & Ostriker (1993a). Note that in this simulation supernova energy feedback into the intergalactic medium from aging massive stars is not included.

2.2. Lyα Cloud Identification

Perhaps the most critical decision which must be made in a study such as this is how to define a Lyα “cloud.” This is not as simple as it may seem, because the smooth transition in density between the global intergalactic medium and the local structures within it can blur the distinction between actual clouds and the intercloud medium. However, in
order to examine the structure of \( Ly \) clouds in any quantitative way, it is absolutely necessary to adopt some sort of definition of the boundary of a cloud. The simplest approach that we believe to be meaningful is to identify clouds as regions with densities above a chosen threshold.

There are two major motivations for defining our clouds in this manner. First of all, this method allows us to associate the clouds found at each density cut with a particular column density, once we obtain information about the characteristic sizes of the clouds. This feature is desirable because we can only observe \( Ly \) forest clouds through absorption lines in quasar spectra. Second, density perturbations with amplitudes larger than a certain threshold become self-gravitating bound clouds. For a spherical cloud, the average overdensity at which a cloud breaks away from the general Hubble expansion is \( \sim 5.55 \). Thus, our cloud definition also has physical motivation in that structures above some high density threshold are gravitationally bound, distinct systems. Having chosen a general strategy for defining our clouds, we now describe the details of the grouping procedure.

Cells with a baryonic density below a chosen value (3, 10, and 30 will be used) in units of the global mean of the baryonic density are cut out of the original density array. Then clouds are defined by grouping the remaining cells using the DENMAX scheme (Bertschinger \& Gelb 1991) as follows.

First, two sets of baryonic densities are stored. One is the original density array from the simulation \( \rho_{\text{org}} \), 288\(^3\) elements, and the second is a smoothed version of \( \rho_{\text{org}} \) with a Gaussian smoothing window of radius 1.5 cells \( \rho_{\text{smooth}} \), 288\(^3\) elements. The smoothing is performed to eliminate any small cell-to-cell density fluctuations, which could be physical, in the form of small discontinuities or oscillating sound waves, or could only be numerical noise. Such small-scale fluctuations could result in the identification of small unwanted clouds with sizes on the order of one cell. At the same time, since the true resolution of the simulation is close to 2–3 cells, a Gaussian smoothing window of radius 1.5 cells should preserve all information regarding “real” adjacent structures. A larger smoothing window could merge real and separate entities that are actually well-resolved in the simulation. We have made rather extensive experiments on the smoothing operation and conclude that the adopted smoothing window size is appropriate, as will be shown below.

After the smoothing operation is performed, we cut out all elements of \( \rho_{\text{smooth}} \) and \( \rho_{\text{org}} \) whose \( \rho_{\text{org}} \) values are below a chosen value, \( \rho_{\text{cut}} \). We then collect the uncut cells of \( \rho_{\text{org}} \) (using the DENMAX scheme) by propagating each cell along the gradient of \( \rho_{\text{smooth}} \) until it reaches a local density maximum. All the cells collected at a particular local maximum are grouped into one “cloud.” Once the cells have been grouped, we examine the shapes and sizes of the resultant “clouds” by approximating them as ellipsoids. We create a symmetric \( 3 \times 3 \) moment tensor of the form

\[
T_{ij} = \sum_{\text{cells}} x^i y^j \rho_{\text{org}} dV
\]

and so on for each cloud. This tensor is diagonalized and the square roots of the eigenvalues yield the lengths of the semimajor and semiminor axes, \( a', b', c' \), in descending order of length. For cells where \( \Delta x, \Delta y, \text{ or } \Delta z = 0 \), we account for the finite size of a single cell by finding \( \int x^2 dy dx / dy dx = 1/12 \) for a 1 cell square and adding that value to the tensor component for a particular cloud. This method underestimates the actual axial lengths of the ellipsoid because it only computes moments. We correct this problem by scaling up the values of \( a', b', \) and \( c' \) determined above by a factor of \( (3V/4\pi a'b'c')^{1/3} \):

\[
(a, b, c) = (a', b', c')(3V/(4\pi a'b'c'))^{1/3}
\]

where \( V \) is the actual volume of the cloud computed by summing over all the member cells of the cloud.

3. Results

3.1. Visual Inspection

We first present a three-dimensional visual description of the structure of the \( Ly \) clouds produced in the simulation. Figure 1a, b, and c (Plates 1–3) displays the isodensity contour surfaces at \( z = 3 \) for \( \rho_{\text{cut}} = 3, 10, \) and 30, respectively. Here \( \rho_{\text{cut}} \) is the baryonic density in units of its global mean. The box size is \( 10 h^{-1} \) Mpc comoving. The most striking visual feature is the network of sheets and filaments spanning the box. In Figure 1a we see that most of the structures are connected over the entire box, and it is apparent that most of the covering areas (i.e., cross sections for quasar lines of sight at low column densities [\( \lesssim 10^{12} \text{ cm}^{-2} \]; see below]) are sheet-like objects. Although this result was obtained for the particular \( \Lambda \)CDM universe simulated here, we expect this qualitative picture to be fairly generic for any CDM-like model (HDN-like models should show even more prominent sheets, while models like PBI are likely to be less coherent). At a higher density of \( \rho_{\text{cut}} = 10 \) (Fig. 1b), most structures become filamentary and more isolated. When the isodensity reaches \( \rho_{\text{cut}} = 30 \) (Fig. 1c), most of the filaments are replaced by relatively round systems that are typically isolated, with separations of about 1 Mpc comoving.

The second noticeable visual feature is the existence of a large low-density region at the center of the box, occupying more than half the simulation volume. Since the initial state of the simulation is not constrained (i.e., the simulation is only a random realization of the cosmological model in a certain volume), this void indicates that the simulation box may still not be large enough to contain a “fair” volume for the structures under consideration. It may be necessary to use a larger simulation box, perhaps 20–30 \( h^{-1} \) Mpc on a side, in order to properly sample the objects in question. Different properties will be affected by the simulation volume to varied degrees. We expect that the most significant differences will be in quantities like void sizes and correlations.

Let us now examine the individual clouds to assess the general accuracy of our grouping scheme. Figure 2a shows clouds at \( \rho_{\text{cut}} = 3 \) for two randomly selected slices of size \( 5 \times 5 h^{-2} \) Mpc\(^2\) comoving with thickness of 175 \( h^{-1} \) kpc comoving at \( z = 3 \). The two left panels show the density contours at levels \( 10^{10^{10^{10}} \rho_{\text{cut}}^i} = 1, 2, 3, \ldots \), and the right panels show the \( Ly \) clouds identified in the same slices. Two symbols are used to show two types of clouds: the filled dots represent “large” clouds, with the dot size roughly proportional to the actual size of the cloud and the open circles represent the “small” clouds. A cloud is called “small” if its smallest axis does not exceed 3 cells in length, and “large” otherwise. Figure 2b is similar to Figure 2a, but with \( \rho_{\text{cut}} = 10 \) and using two other random slices. We note that “large” clouds are typically embedded in larger
extended density structures while “small” ones are more isolated. A close one-by-one examination of the density maxima in the density contour plots indicates that each identified cloud corresponds to a well-defined density maximum. (Some clouds, especially “small” ones, have no corresponding density maxima simply because their densities fall below the contour levels when averaged over the slice.) It appears that our smoothing operation and cloud identification scheme indeed yield well-defined, distinct clouds with one particularly desirable feature: they are neither over-merged in the sense that a single cloud contains multiple structures of comparable size (oversmoothed), nor over-separated in the sense that artificial clouds are created by small-scale amplitude density fluctuations (undersmoothed).

When the clouds are classified into “large” and “small” groups, we find that most of the clouds are “small.” The fractions of “small” clouds at $\rho_{\text{cut}} = 3, 10,$ and 30 at $z = 3$ are 62%, 92%, and 98%, respectively, containing 14%, 49%, and 63% of the total cloud mass. The fact that a large fraction of the baryonic mass is in “small” clouds indicates that a higher resolution simulation is perhaps needed before we can be absolutely sure that these clouds are resolved properly. Inclusion of more initial small-scale power, which is limited by the Nyquist frequency of a simulation, might further increase the fraction of “small” clouds.

We note in passing that we also experimented with the conventional friends-of-friends grouping algorithm, with a linking length of one cell. The resulting clouds were obviously not useful for our analyses. For example, at $\rho_{\text{cut}} = 3$, most of the volume, as well as the mass, wound up in a single supercloud because the distinct clouds were linked together by touching boundaries, as shown in Figure 2. In other words, the friends-of-friends grouping scheme picks out large, interconnected networks (see but not the Fig. 1a) but not the individual structures embedded within them.

3.2. Quantitative Measures

We have chosen to focus on three quantitative aspects of the identified Ly$\alpha$ clouds: sizes, shapes, and correlations. In addition, an analysis of pairs of quasar sight lines is performed.
To facilitate quantitative discussions, we relate the mass density of a cloud to its column density. Assuming that the clouds are in photoionization equilibrium, the column density of a cloud can be related to the parallel size and density of the cloud (assuming that the density is uniform across the cloud along the line of sight) as

$$N_{\text{HI}} = 8.18 \times 10^{11} \left( \frac{\Omega_b}{0.0125 \, h^{-2}} \right)^2 \left( \frac{j_{\text{HI}}}{10^{-12}} \right)^{-1} \times \left( \frac{T}{10^2} \right)^{-0.7} \frac{L}{100} \left( \frac{\rho_{\text{cut}}}{\langle \rho_b \rangle} \right)^2 \left( \frac{1 + z}{4} \right)^6 \text{cm}^{-2},$$

where $\Omega_b$ is the baryonic density parameter, $T$ the temperature in Kelvin, $j_{\text{HI}}$ the hydrogen photoionization rate, $\rho_b$ and $\langle \rho_b \rangle$ the baryonic density of the cloud and the mean baryonic density at the redshift in question, $L$ the proper size of the cloud along the line of sight in kpc, and $z$ the redshift. For a photoionization radiation field with power-law form $\nu^{-1}, j_{\text{HI}} = 4.34 \times 10^{-12} J_{\text{LL}}$ s$^{-1}$, where $J_{\text{LL}}$ is the intensity of the photoionization field at the Lyman limit in the usual units ($10^{-21}$ ergs cm$^{-2}$ Hz$^{-1}$ s$^{-1}$ sr$^{-1}$). Taking $\Omega_b = 0.0125 \, h^{-2}, j_{\text{HI}} = 7.0 \times 10^{-13}$ s$^{-1}, T = 2 \times 10^4$ K, and $L = 40$ kpc as typical values at $z = 3$, one finds that $N_{\text{HI}} = 2.5 \times 10^{12}, 2.7 \times 10^{13},$ and $2.5 \times 10^{14}$ cm$^{-2}$ for $\rho_{\text{cut}} = 3, 10,$ and $30$, respectively. Exact column densities would depend on exact values of the assumed quantities as well as the actual density distribution. We expect that the actual density distribution, which is nonuniform, will increase the estimated column densities. For example, redistribution of the same mass within $L$ into a one-dimensional single power-law distribution would result in a column density $(1 + z)^2/(1 + 2z)$ times that of the uniform density distribution, where $z$ is the slope of the density profile. We simply increase the above estimates by a factor of 4 to account for the gradient in the density distribution of a cloud (note that this factor of 4 is fairly plausible for a reasonable $z \sim 0.5$ in the one-dimensional singular case or a lower $z$-value for a nonsingular profile. We also note that this factor of 4 can also be achieved if $\Omega_b$ is twice the value adopted here, in light of recent new measurements on the deuterium abundance (Tytler, Fan, & Burles). Thus we obtain

$$N_{\text{HI}} = (1.0 \times 10^{13}, 1.1 \times 10^{14}, 1.0 \times 10^{15}) \text{cm}^{-2}$$

for $\rho_{\text{cut}} = (3, 10, 30)$ at $z = 3$. (3)

We will use the above relation for quantitative discussions.

3.2.1. A Few Simple Global Quantities

Before engaging in a more detailed quantitative discussion of cloud properties, we present a few simple global
quantities. Listed in Table 1 are the numbers of clouds at each redshift and density threshold, the baryonic mass fraction of such clouds, the mean intercloud separation in proper units, and the mean cloud mass. Four points are interesting. First, while the average cloud increases in mass by a factor of about 3–4 from $z = 4$ to $z = 2$, the number of clouds decreases by a factor of 2. This result indicates that the merging of old clouds and the creation of new clouds appear to shrink due to gravitational collapse.

To summarize, we find that Ly$\alpha$ clouds exhibit a wide range of sizes, from a few proper kiloparsecs to about a hundred proper kiloparsecs at $z = 3$. The comoving size of these clouds tends to increase slowly with time, implying that the proper physical size increases with time more rapidly than $(1 + z)^{-3}$.

3.2.3. Shapes of Ly$\alpha$ Clouds

Figure 5a, b, and c shows the distributions of cloud shapes in the $c/b$-$b/a$ plane for clouds with $\rho_{\text{cut}} = 3, 10$, and 30 at $z = 3$. Each dot represents one cloud and the contours indicate the number density of clouds (weighted by $A$) in the plane. The contour levels are incremented up linearly from outside to inside. We see that the most common cloud at $\rho_{\text{cut}} = 3$ and 10 is an ellipsoid with axial ratios of $\sim 1:2:4$. However, at $\rho = 30$ the situation is interestingly different.

![Figure 3](image-url)

**Fig. 3.**—Cumulative distributions of the cloud sizes at $z = 3$ for $\rho_{\text{cut}} = 3, 10,$ and 30, weighted by $A$ (thin curves) and by mass (thick curves).
Two major concentrations of clouds are seen at \((b/a, c/b) = (0.35, 0.85)\) and \((0.75, 0.85)\), indicating the existence of two distinct populations: filaments and near-spherical ellipsoids.

In order to conveniently show the evolution of shapes of the clouds and to relate shapes to other quantities, we define a simple "shape" parameter of a cloud as
\[
\eta \equiv \frac{b^2 + c^2}{a^2}.
\]

For an ideal spherical, filamentary, or disk cloud, \(\eta = 2, 0,\) or \(1\), respectively. Some degeneracy exists for any value of \(\eta\) with such a simple definition. For example, \((a, b, c) = (1, 1, 0)\) and \((a, b, c) = [1, 1/(2)^{1/2}, 1/(2)^{1/2}]\) both give \(\eta = 1\).

Figure 6 shows \(\eta\) as a function of size \(S\) at \(z = 3\) for \(\rho_{\text{cut}} = 3, 10,\) and \(30\). There is a trend, albeit with a large scatter, that larger clouds tend to be less spherical.

Figure 7 represents the \(A\)-weighted cumulative distribution of \(\eta\) for three cases at \(z = 3\). We note two points. First, we see that there is negligibly small fraction of clouds with the shape parameter \(\eta\) greater than \(1.5\), indicating that spherical clouds \((\eta = 2)\) contribute only a tiny percentage to the total covering area of the clouds at all three density cuts. Second, we find a larger fraction of round clouds \((\eta \approx 0)\) and filamentary clouds \((\eta \approx \text{close to zero})\) at high densities than at lower densities. This result comes about because the large sheets and filaments present at lower densities are broken up into filaments and turn into roundish regions at higher densities, respectively. This quantitative result is consistent with the visual impression that structures go from sheets to filaments and spheres as the density goes from low to high (Fig. 1). The fact that there are more filaments than spheres at all densities is probably a consequence of nonlinear evolution coupled with hydrodynamic (pressure) effects. The situation for dark halos might be different, since dark matter distributions tend to be more unstable in lower dimensional structures.

Finally, we show in Figure 8 the redshift evolution of the median shape parameter \(\eta_{\text{med}}\). It indicates that at \(\rho_{\text{cut}} = 3\) the median shape of the clouds do not change with time, while at \(\rho_{\text{cut}} = 30\) the clouds progress toward the shape of spheres (larger \(\eta\)) in time, with a dramatic upturn from \(z = 3\) to \(z = 2\), possibly related to the collapse and/or merging of gas along the longest axis. The situation at the intermediate \(\rho_{\text{cut}} = 10\) is between the above two cases.

In summary, we have shown that Ly\(\alpha\) clouds display a variety of shapes ranging from quasispherical clouds to filaments and sheets, with the most common shape being a "flattened cigar" with an axial ratio of \(\sim 1:2:4\) at a column density of \(10^{13}-10^{14}\) cm\(^{-2}\) and the most common shapes being a thin cigar with an axial ratio of \(1:3:4\) and a near sphere with an axial ratio of \(1:1.3:1.5\) at a column density of more than \(10^{15}\) cm\(^{-2}\). Larger clouds tend to be less spherical. The shapes of clouds with densities a few times the mean density of the universe evolve very weakly with time, whereas the clouds at higher densities \((\rho_{\text{cut}} \geq 30)\) grow more spherical with time.

3.2.4. Analysis of Pairs of Quasar Lines of Sight

CMOR and MCOR examined many properties of Ly\(\alpha\) clouds in detail and found that the simulation results match observations very well. Here, we focus on the observations of pairs of quasar sight lines (Smette et al. 1992, 1995; Dinshaw et al. 1994, 1995; Fang et al. 1996; Bechtold et al. 1994; Bechtold & Yee 1995). Let us first describe our procedure for selecting pairs of sight lines.

1. Once three-dimensional clouds in the simulation box are identified (see §2.2), each cell is labeled with an integer \(n\), meaning that it belongs to cloud \(n\) (if \(n = 0\), it means that the cell does not belong to any cloud, i.e., its density is below \(\rho_{\text{cut}}\)).

2. A random pair of sight lines separated by \(\Delta r\) is selected. Since the simulation box has no preferred orientation, we simply choose a direction perpendicular to one of three faces of the simulation box as the direction for the double sight lines.

3. Along each of the two sight lines we identify cells whose labels \(n\) (see item 1) are nonzero, and then separate regions of different \(n\) into separate clouds, which we call "clouds along the line of sight" (CALOSs). Note that each CALOS inherits its three-dimensional parent cloud label \(n\).

4. We define the center of each CALOS in \(z\)-space by \(\rho_{\text{cut}}^2\)-weighted averaging over the positions of all the cells in that CALOS that lie along the sight line.

5. For each CALOS along the first sight line, if a CALOS that lies nearer to the CALOS on the first sight line than a velocity space separation \(D\) is found along the second sight line, an integer counter \(N_{\text{common}}\) is incremented if the two CALOSs share the same \(n\), i.e., if they belong to the same three-dimensional cloud. Otherwise a "clustering" integer counter \(N_{\text{cluster}}\) is incremented by 1. If both a common and a clustering line are found, we consider the pairing common. The above exercise is then repeated for each CALOS along the second sight line. We note that the bulk peculiar velocity of each CALOS is ignored in this calculation, i.e., the line-of-sight velocity difference between a pair of coincident lines simply reflects the line-of-sight real-space distance difference. This approximation should hold because the peculiar velocity gradient is typically smaller than the Hubble constant on the scale \(D\) (see below) in which we are interested.
We perform this exercise on 10,000 randomly selected pairs of sight lines for \( \rho_{\text{cut}} = 3, 10, \) and 30 at six different perpendicular separations, \( \Delta r = 1, 3, 6, 12, 24, \) and 48 cells. Figure 9 shows the fraction of lines that are coincident at \( z = 3, {\cal N}_{\text{co,tot}}/{\cal N}_{\text{tot}}, \) as a function of \( \Delta r, \) where \( {\cal N}_{\text{co,tot}} = {\cal N}_{\text{common}} + {\cal N}_{\text{cluster}} \) and \( {\cal N}_{\text{tot}} \) is the total number of lines (clouds). Six curves are shown, corresponding to combinations of \( \rho_{\text{cut}} = 3, 10, \) and 30 with \( D = 50 \) and 150 km s\(^{-1}\). We see that the line coincidence rate decreases with increasing column density. At a perpendicular proper separation of 100 h\(^{-1}\) kpc and \( \rho_{\text{cut}} = 3 \) (corresponding to clouds with \( N_{\text{HI}} = 10^{13} \) cm\(^{-2}\)), about \( \frac{1}{2} \) to \( \frac{3}{4} \) of the sight lines should have coincident absorption features for \( D = 50 \)–150 km s\(^{-1}\). The line coincidence rates drop to 27% and 38%, respectively, for \( \rho_{\text{cut}} = 10 \) (\( N_{\text{HI}} = 10^{14} \) cm\(^{-2}\)) and to 11% and 17% for \( \rho_{\text{cut}} = 30 \) (\( N_{\text{HI}} = 10^{15} \) cm\(^{-2}\)) at \( D = 50 \) and 150 km s\(^{-1}\).

Figure 10 shows the redshift evolution of \( {\cal N}_{\text{co,tot}}/{\cal N}_{\text{tot}} \) at \( \Delta r = 100 h^{-1} \) kpc in proper units for four cases with \((N_{\text{HI}}, D) = (3 \times 10^{13}, 150), (3 \times 10^{13}, 50), (3 \times 10^{14}, 150), \) and \((3 \times 10^{14}, 50).\) The indicated neutral hydrogen column densities in Figure 9 are approximately computed as follows. Combining equation (3) for \( N_{\text{HI}} \) at \( z = 3 \) with the result that the physical sizes of clouds go roughly as \((1 + z)^{-3/2} \) (Fig. 4) gives \( N_{\text{HI}}(\rho_{\text{cut}}, z) = 1.0 \times 10^{13}(\rho_{\text{cut}}/3.0)^2(1 + z)^{9/2} \) cm\(^{-2}\). We see that, at a fixed column density, the rate of coincident lines
Shape parameter $\eta$ (see eq. [6]) as a function of size $S$ at $z = 3$ for $\rho_{\text{cut}} = 3$ (top), 10 (middle), and 30 (bottom). There is a trend, with large scatters, for larger clouds to be less spherical and more filamentary.

(at $\Delta r = 100$ km s$^{-1}$) evolves with redshift very weakly, assuming that $j_{\text{HI}}$ is constant over the redshift range in question. The assumption of constant $j_{\text{HI}}$ is merely for the convenience of illustration. In fact, the self-consistent photoionization field produced during the simulation is constant from $z = 4$ to $z = 2$ to within a factor of 2. In other words, the results would not have significantly differed if the actual photoionization field had been used.

Given the three-dimensional nature of the simulations, we can distinguish between two kinds of coincident lines, one of which ($N_{\text{common}}$) occurs when a pair shares the same three-dimensional parent cloud, the other of which ($N_{\text{cluster}}$) occurs when the lines intersect two separate three-dimensional clouds. The former kind of coincident lines arises from the extended size of a single cloud, while the latter kind of coincident lines arises from the clustering of distinct clouds. Figure 11 shows $N_{\text{common}}/N_{\text{tot}}$ at $z = 3$. Six curves are shown for $\rho_{\text{cut}} = 3$, 10, and 30 with $D = 50$ and

$$\frac{N_{\text{common}}}{N_{\text{tot}}}$$ as a function of $\Delta r$. Six curves are shown for combinations of $\rho = 3, 10, 30$ with $D = 50$ and $150$ km s$^{-1}$.

Redshift evolution of the median shape parameter $\eta_{\text{med}}$ for $\rho_{\text{cut}} = 3, 10$, and 30.
150 km s$^{-1}$. We see that at $\Delta r \approx 30-60$ h$^{-1}$ kpc about half of the coincident line pairs share the same clouds and the other half pierce separate clouds. At $\Delta r = 100$ h$^{-1}$ kpc only 10%-20% of the total coincident line pairs share the same clouds, i.e., clustering of separate clouds dominates the coincident events at this and larger separations.

The observational signature of clustering-dominated coincident lines is that the difference in line of sight velocity between the two absorption lines should be a weak function of the separation perpendicular to the line of sight. Figure 12 shows rms velocity difference of coincident line pairs along the line of sight, $R_{||,\text{rms}}$, as a function of the perpendicular separation $\Delta r$ for $D = 50$ and 150 km s$^{-1}$ at $\rho_{\text{cut}} = 3, 10, 30$. The two long-dashed horizontal lines show the rms difference for the two $D$ values if clouds are small (compared to perpendicular separation) and randomly distributed. The combined results of our analysis of pairs of quasar sight lines suggest a few basic trends which should be kept in mind when analyzing QSO spectra. The most important of these is that line-coincidence events can be caused by two different phenomena. At small perpendicular separations ($\Delta r < 40$ h$^{-1}$ kpc proper) a pair of coincident lines is most likely to pierce a common cloud, as is usually assumed when analyzing such observations to infer the actual cloud number density (crowding of clouds) at high redshift controls the type of line-coincidence events. This fact is consistent with results shown in Figure 11.
size. However, at larger $\Delta r$, a pair of coincident lines is more likely to penetrate two different clouds that are spatially clustered. At $\Delta r = 100$ h$^{-1}$ kpc, only 10%–20% of the total coincident line pairs share the same clouds. At very large separations of $\Delta r > 500$ h$^{-1}$ kpc, all the coincident line events can be explained as the random intersection of two unrelated clouds whose rms velocity difference $R_{\text{rms}}$ should be related to the velocity interval $D$ as in equation (7).

3.2.5. Correlations of Ly$\alpha$ Clouds

It is tempting to make a connection between Ly$\alpha$ clouds and dwarf galaxies and/or moderate-redshift faint blue galaxies. We approach this problem by examining clustering properties of the Ly$\alpha$ clouds.

A few cautionary words about the limitations of the simulation are in order here. The simulation box size ($L = 10$ h$^{-1}$ Mpc comoving) places severe limits on our ability to study the clustering properties of clouds on large scales with a high degree of accuracy because it sets an upper limit on the scale of the input power spectrum. Waves longer than the simulation box size would have made a considerable contribution to clustering on larger scales, since the density fluctuations on scales comparable to the simulation box length started to approach nonlinearity even at redshift $z = 2$ ($\sigma \sim 0.4$ at $z = 2$ on the box scale). The situation became even more serious at lower redshift.

Nevertheless, we may gain some insight into the distribution of mass on smaller scales and say something useful about the more local clustering properties of clouds. The three-dimensional correlation of the Ly$\alpha$ clouds at $z = 3$ is shown in Figure 14 for $\rho_{\text{cut}} = 10$ and 30. Higher column density clouds seem to be more strongly clustered than lower column density ones. Since most of the faint blue galaxies are thought to be in the redshift range $z \sim 0.5$–1.0, we show in Figure 15 the evolution of correlation length $r_o$ (solid curve), defined as the length for which the correlation is unity, plotted for the redshift range $z = 2$–4, for which we have the most confidence in the accuracy of the simulation. All lengths are shown in comoving units.

In order to estimate the effect of missing power at box-sized scales, we computed the correlation function of the clouds in eight subboxes. These subboxes were created by dividing the original simulation box into eight equal cubes and pretending that each subbox also had periodic boundary conditions. The dotted line in Figure 15 shows the median value of the correlation length for the subboxes and the vertical bars indicate the full width of the distribution at each redshift (i.e., the highest and lowest values of $r_o$). We see that the median correlation length drops about 30%–40% for a box of size $L = 5$ h$^{-1}$ Mpc. Our best estimate is that the true correlation length at each redshift
would be larger than values indicated by the solid curves by perhaps 50% if the simulation box were sufficiently large.

Simple extrapolation of the solid curve in Figure 15 to lower redshifts is unlikely to give us accurate values of \( r_p \) at \( z \sim 0.5-1.0 \). However, it is reasonable to expect the value of \( r_p \) at these redshifts to be significantly larger than 1.5 \( h^{-1} \) Mpc comoving, when one takes into account the effect of the missing longer waves. This correlation length may be close to the observed correlation length of the faint blue galaxies (Efstathiou et al. 1991; Nechaev, Windhorst, & Dressler 1991; Couch, Jurcevic, & Boyle 1993; Roche et al. 1993; Infante & Pritchet 1995; Brainerd, Smail, & Mould 1995), although an accurate conversion from the observed angular correlation function \( \omega(\theta) \) to the three-dimensional correlation function computed here would require a detailed knowledge of the redshift distribution of the observed faint blue objects that is currently unavailable.

Furthermore, the average mass of the clouds is close to \( 10^8 - 10^{10} M_\odot \) (see Table 1), in accord with the mass that the faint blue galaxies are thought to possess or are predicted to have (e.g., the starburst dwarf galaxy model of Babul & Rees 1992). Finally, the number density of the clouds (with a mean separation of \( \sim 2 h^{-1} \) Mpc at \( z = 2 \) for \( \rho_{cut} = 30 \)) and the tendency of the number density to increase with redshift indicate that the number density is probably quite close to the observed value for faint blue objects. These three considerations lend tentative support to the conjecture that high-redshift Ly\( \alpha \) clouds (\( N_{HI} \geq 10^{15} \) cm\(^{-2} \)) are the progenitors of faint blue galaxies. Fang et al. (1996) made a similar suggestion based on their Ly\( \alpha \) cloud observations.

4. CONCLUSIONS

We have presented a quantitative study of the sizes, shapes, and correlations of Ly\( \alpha \) clouds in a spatially flat cold dark matter universe with a cosmological constant, utilizing state of the art cosmological hydrodynamic simulations, including detailed atomic physics, for a plasma of primordial composition. Keeping in mind the unavoidable limitations of such numerical experiments, e.g., limited box size and limited resolution, as well as some approximate treatments of physical processes, a few findings are probably fairly reliable.

Structures formed at high redshift \( z \sim 2-4 \) by gravitational growth and collapse of cosmic density perturbations at small to intermediate scales (100 kpc to a few Mpc comoving) are responsible for the observed Ly\( \alpha \) forest. The sizes and shapes of these Ly\( \alpha \) clouds cannot be characterized by a simple model. Their sizes vary in a wide range, from a few kiloparsecs to about one hundred kiloparsecs in proper units, with the median size being \( \sim 15-35 h^{-1} \) kpc at \( z = 3 \), and their shapes vary from nearly spherical ellipsoids to filaments and pancakes. A typical Ly\( \alpha \) cloud resembles a flattened cigar, with an axis ratio of \( \sim 1:2:4 \) at column densities of \( 10^{15}-10^{16} \) cm\(^{-2} \), and is either a thin cigar with an axis ratio of \( 1:3:3.4 \) or a near sphere with an axis ratio of \( 1:1:3:15 \) at column densities greater than \( 10^{15} \) cm\(^{-2} \).

Larger clouds are, on average, less spherical than smaller ones. The physical size of a typical cloud grows with time roughly as \((1+z)^{-3/2}\), while its shape hardly evolves (except for the most dense regions \( \rho_{cut} > 30 \), which tend to become more spherical with time).

An analysis of simulated quasar double sight lines indicates that coincident absorption features observed in the two spectra can have two different causes. At small perpendicular sight line separations (\( \Delta r < 40 h^{-1} \) kpc proper), a pair of coincident lines most likely represents absorption by a common cloud. This case is usually assumed when analyzing such observations to infer the actual cloud size. However, at larger \( \Delta r \), a pair of coincident lines most likely samples two separate clouds which belong to the same cloud cluster. In other words, clustering of separate clouds dominates the coincident events at larger sight line separations. At \( \Delta r = 100 h^{-1} \) kpc, 80%-90% of coincident line pairs can be explained in this fashion. At very large separations, \( \Delta r > 500 h^{-1} \) kpc, the coincident line events are entirely due to random intersections of two unrelated, uncorrelated clouds, whose rms velocity difference \( R_{rms}^* \) should relate to the velocity interval being examined, \( D \), as \( R_{rms}^* \sim 0.58 D \).

Analyzing observed coincident absorption lines to infer the actual sizes of Ly\( \alpha \) clouds by assuming a population of clouds with simplified geometry (e.g., spheres, disks) is reasonably accurate only when the sight line separations are small (\( \Delta r \leq 50 h^{-1} \) kpc). For larger \( \Delta r \), this exercise is not very meaningful. The inferred sizes are not related to the true sizes of the clouds and are grossly inflated.

We attempt to make a connection between high column density Ly\( \alpha \) clouds (\( N_{HI} \geq 10^{15} \)) and faint blue galaxies. The evidence supporting such a conjecture is threefold. First, the correlation of the high column density Ly\( \alpha \) clouds (\( \geq 1.5 h^{-1} \) Mpc comoving), when extrapolated to \( z \sim 0.5-1.0 \), seems close to that of the observed faint blue galaxies. Second, the typical mass of these objects (\( 10^8 - 10^{10} M_\odot \)) is close to that of a dwarf galaxy (see, e.g., the scenario of Babul & Rees 1992). Finally, the number density of the clouds is reasonably close to what is observed for faint blue objects.

Finally, we caution that it may be necessary to extend the simulation’s dynamic range to 1000–3000 cells in each dimension, in order to have both a larger box (\( L \sim 20–30 \) \( h^{-1} \) Mpc).
$h^{-1}\text{Mpc comoving}$ and finer resolution ($\Delta l \sim 10-20\ h^{-1}\text{kpc comoving}$) to ensure both a "fair" sample and full resolution of the structures of the objects in question.

This work was supported in part by NASA grant NAG 5-2759 and NSF grants AST 91-08103 and ASC 93-18185. We thank the referee, George Efstathiou, for a constructive and pertinent report which improved the quality of the work. Discussions with J.R. Gott, N. Katz, G. Lake, J. Miralda-Escudé and J.P. Ostriker are gratefully acknowledged. R.C. would like to thank G. Lake and the Department of Astronomy at the University of Washington for their warm hospitality and the NASA High-Performance Computing and Communications/Earth and Space Sciences Program for financial support during a visit when much of this work was completed.

REFERENCES

Adams, F., Bond, J. R., Freese, K., Frieman, J., & Olinto, A. 1993, preprint

Arons, J. 1972, ApJ, 172, 553

Babul, A., & Rees, M. 1992, MNRAS, 255, 346

Bahcall, N., & Cen, R. 1992, ApJ, 398, L81

1993, ApJ, 407, L49

Bahcall, J. N., & Spitzer, L. 1969, ApJ, 156, L63

Bechtold, J., Crotts, A. P. S., Duncan, R. C., & Fang, Y. 1994, ApJ, 437, L83

Bechtold, J., & Yee, H. K. C. 1995, AJ, 110, 1984

Berlind, E., & Gelb, J. M. 1991, Comput. in Phys., 5, 164

Bolte, M., & Hogan, C. J. 1995, Nature, 376, 399

Bond, J. R., & Myers, S. T. 1996, ApJ, 457

Bond, J. R., Szalay, A. S., & Silk, J. 1988, ApJ, 324, 627

Brainerd, T. G., Smail, I., & Mould, J. 1995, 7,81B

Bucher, A., Goldhaber, G., & Turok, N. 1995, preprint

Carswell, R. F., Lanzetta, K. M., Parnell, H. C., & Webb, J. K. 1991, ApJ, 371, 36

Cen, R. 1992, ApJS, 78, 341

Cen, R., Gnedin, N. Y., Kohan, L. A., & Ostriker, J. P. 1992, ApJ, 399, L11

Cen, R., Gnedin, N. Y., & Ostriker, J. P. 1993, ApJ, 417, 387

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9 (CMOR)

Cen, R., & Ostriker, J. P. 1992, ApJ, 399, L113

1993a, ApJ, 417, 404

1993b, ApJ, 417, 415

Cen, R., Ostriker, J. P., & Peebles, P. E. J. 1993, ApJ, 415, 423

Couch, W. J., Jurcevic, J. S., & Boyle, B. J. 1993, MNRAS, 260, 241

Cole, S., Frenk, C. S., & White, S. D. M. 1994, ApJ, 423, 129

Cristiani, S., D'Odorico, S., Fontanot, A., Giallongo, E., & Savaglio, S. 1995, MNRAS, 273, 1016

Davis, M., Summers, F. J., & Schlegel, D. 1992, Nature, 359, 393

Dinshaw, N.,oltz, C. D., Impye, C. D., Weymann, R. J., & Morris, S. L. 1995, Nature, 373, 223

Dinshaw, N., Impye, C. D., Foltz, C. B., Weymann, R. J., & Chaffee, F. H. 1994, ApJ, 437, L37

Edelson, R. A., & Malkan, M. A. 1986, ApJ, 308, 59

Efstathiou, G., Bernstein, G., Tyson, J. A., Katz, N., & Guhathakurta, P. 1991, ApJ, 380, L47

Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1P

Eke, V. R., Cole, S., & Frenk, C. S. 1996, preprint

Fang, Y., Duncan, R. C., Crotts, A. P. S., & Bechtold, J. 1996, ApJ, 462, 77

Fieldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 437, 56

Freedman, W. L., et al. 1994, Nature, 371, 757

Fukugita, M., Hogan, C. J., & Peebles, P. J. E. 1993, Nature, 366, 304

Gott, J. R., III, 1982, Nature, 295, 304

Hamuy, M., et al. 1995, AJ, 104, 1

Hernquist, L., Katz, N., & Weinberg, D. H. 1996, ApJ, 457, L51

Ha, E. M., Kim, T.-S., Cowie, L. L., Songaila, A., & Rauch, M. 1995, AJ, 110, 1526

Ikeuchi, S. 1986, Ap&SS, 118, 509
Fig. 1a

Fig. 1.—Three-dimensional isodensity surfaces for $\rho_{\text{cut}} = (a) 3, (b) 10, \text{and} (c) 30 \text{ at } z = 3$

CEN & SIMCOE (see 483, 10)
Fig. 1c

Cen & Simcor (see 483, 10)