THE STATISTICS OF DENSITY PEAKS AND THE COLUMN DENSITY DISTRIBUTION
OF THE Lyα FOREST

LAM HUI, 1, 2 NICKOLAY Y. GNEDIN, 1, 3 AND YU ZHANG 4

Received 1996 August 28; accepted 1997 April 14

ABSTRACT

We develop a method to calculate the column density distribution of the Lyα forest for column densities in the range \(10^{12.5} - 10^{14.5} \text{ cm}^{-2}\). The Zeldovich approximation, with appropriate smoothing, is used to compute the density and peculiar velocity fields. The effect of the latter on absorption profiles is discussed, and it is shown to have little effect on the column density distribution. An approximation is introduced in which the column density distribution is related to a statistic of density peaks (involving its height and first and second derivatives along the line of sight) in real space. We show that the slope of the column density distribution is determined by the temperature-density relation, as well as the power spectrum, on scales \(2 \text{ h Mpc}^{-1} \leq k \leq 20 \text{ h Mpc}^{-1}\). An expression relating the three is given. We find very good agreement between the column density distribution obtained by applying the Voigt profile fitting technique to the output of a full hydrodynamic simulation and that obtained using our approximate method for a test model. This formalism is then applied to study a group of cold dark matter, as well as cold plus hot dark matter, models. We show that the amplitude of the column density distribution depends on the combination of parameters \((\Omega, h^2)T_{\text{0}}^{0.7} J_{\text{HI}}\), which is not well constrained by independent observations. The slope of the distribution, on the other hand, can be used to distinguish between different models: those with a smaller amplitude and a steeper slope of the power spectrum on small scales give rise to steeper distributions, for the range of column densities that we study. Comparison with high-resolution Keck data is made.

Subject headings: cosmology: theory — intergalactic medium — quasars: absorption lines

1. INTRODUCTION

There is a long history of theoretical efforts to place the study of the Lyα forest within the framework of cosmological structure formation theories (Doroshkevich & Shandarin 1977; Rees 1986; Bond, Szalay, & Silk 1988; McGill 1990; Bi, Börner, & Chu 1992). Recent work making use of numerical simulations has greatly advanced our understanding in this direction (Cen et al. 1994; Zhang, Anninos, & Norman 1995; Hernquist et al. 1995; Petitjean, Mücke, & Kates 1995; Miralda-Escudé et al. 1996; see also Bi, Ge, & Fang 1995 for a linear theory calculation). The emerging picture is that it is possible to account for all the observed properties of the Lyα forest (with column densities less than about \(10^{17} \text{ cm}^{-2}\)) by assuming that it originates from the small-scale structure, including the network of filaments, pancakes, and mild density fluctuations, that arises naturally in hierarchical clustering cosmological models (Weinberg et al. 1996).

A commonly used statistic to characterize the forest is its column density distribution, the number of absorption lines per unit neutral hydrogen column density per unit redshift as a function of neutral hydrogen column density. Other useful statistics include line-line correlations and the distributions of \(b\)-values and equivalent widths (Murdoch et al. 1986; Carswell et al. 1991; Press, Rybicki, & Schneider 1993; Cristiani et al. 1995). There have also been proposals of new statistical tools (Meiksin & Bouchet 1995; Miralda-Escudé et al. 1996; Pando & Fang 1996; see Tytler 1992 for a general overview of the statistical issues concerning quasar absorption systems). We focus our attention on the column density distribution in the present work.

One of the most striking features of the observed column density distribution of quasar absorption systems is that it can be approximated by a single power law that extends over many orders of magnitude. This was emphasized by Tytler (1987), among others, who found that in the range \(10^{13} \text{ cm}^{-2} < N_{\text{HI}} < 10^{22} \text{ cm}^{-2}\) the distribution was reasonably well represented by a power law, \(N_{\text{HI}}^{-\beta}\) with \(\beta = 1.51 \pm 0.02\). However, there exists evidence of at least one break. It has been demonstrated that there is a deficit of absorption systems somewhere in the column density range \(10^{14} - 10^{17} \text{ cm}^{-2}\) compared with a power-law extrapolation of the distribution from lower column densities (Carswell et al. 1987; Petitjean et al. 1993; Hu et al. 1995; Giallongo et al. 1996). For reasons that have to do with the nature of the approximations that we make (§ 5.2), we focus our attention on absorption systems with column densities in the range \(10^{12.5} - 10^{14.5} \text{ cm}^{-2}\). Hu et al. (1995) obtained \(\beta = 1.46\) with 95% confidence limits of \(1.09 \pm 0.05\) in the column density range \(10^{12.3} - 10^{14.5} \text{ cm}^{-2}\). Lu et al. (1996) found the same best-fit \(\beta\) for the same range of column densities.

An obvious ultimate goal of recent theoretical work on the Lyα forest is to constrain theories of structure formation. The natural question is what determines the normalization and slope of the column density distribution. What are the major determining factors, in addition to the usual parameters specified by a given cosmological model? To answer these questions, another question has to be
addressed: what are the analytical and/or computational tools necessary to make accurate predictions for the column density distribution, given all the required parameters?

Accordingly, the present work can be divided into three parts, in which the tools are developed, the factors that influence the column density distribution are analyzed, and one application to a class of cosmological models is discussed.

Numerical hydrodynamic simulations (Cen et al. 1994; Zhang et al. 1995; Hernquist et al. 1995; Miralda-Escudé et al. 1996) provide the most obvious tools to study the Lyα forest. Computational costs, however, prevent one from testing extensively several cosmological models. We show in this paper that the Zeldovich approximation (Zeldovich 1970), with appropriate smoothing, is an efficient and accurate alternative. Our basic assumption is that the part of the Lyα forest with column densities less than about \(10^{14.5} \text{ cm}^{-2}\) arises mostly from regions that are slightly overdense (overdensity \(\leq 5\)), or even underdense, and that have not undergone orbit crossing. The Zeldovich approximation can then be coupled with the equations governing the thermal and ionization states of the gas to yield accurate predictions for the density of neutral hydrogen and the peculiar velocity as a function of position. Absorption spectra are then generated and analyzed. Basic expressions for the absorption optical depth are presented in § 2, and the approximations that go into its computation are discussed in § 3.

Given an absorption spectrum, the column density distribution depends on the method of identifying lines and assigning column densities. This is discussed in § 4.1. We investigate the effects of peculiar velocities on the column density distribution, using a method described by Miralda-Escudé et al. (1996). We find that although peculiar velocities can strongly influence the shapes of absorption profiles, they play a relatively minor role in determining the column density distribution. The various interesting effects of peculiar velocities are discussed in § 4.2. Motivated by this finding, a very different way of assigning column densities is introduced in § 5, in which no absorption spectrum needs to be generated. In the absence of peculiar velocities, there is a one-to-one correspondence between density peaks in real space (if they are separated by a distance larger than a minimum corresponding to the thermal broadening width) and minima of transmission (maxima in absorption) in the observed spectrum. Under such conditions, we can simply associate each density peak in real space with an absorption line and assign a column density to each based upon the height and curvature of the peak. The column density distribution is then a statistic of density peaks in real space. We apply this procedure (we call it the density-peak Ansatz) to the density field predicted by the truncated Zeldovich approximation and test the result against that of a full hydrodynamic simulation. The column density distribution obtained in this way is compared with that obtained from the hydrodynamic simulation using the Voigt profile fitting technique, which is the line identification method most widely used. The level of agreement is found to be excellent. In § 5.2, we discuss the range of parameters in which our computed column density distribution is expected to be reliable.

Armed with the right tools, we turn our attention to the second question: what factors determine the column density distribution? They can be divided into two categories. One has to do with properties of the intergalactic medium, including its temperature, the equation of state (or temperature-density relation, which we will use interchangeably; see Hui & Gnedin 1996), the ionizing radiation intensity, and the baryon density. Uncertainties in all of them have to be taken into account before one can meaningfully confront theories with observations. We distinguish between the factors that mostly affect the normalization of the column density distribution and those that mostly affect its slope: it is found that the temperature-density relation (weakly) affects the slope while the rest of the above factors influence the normalization. It is also emphasized that the temperature and the equation of state depend on the reionization history of the universe (a fuller discussion of this point and related topics will be given in a separate paper). The second set of factors affecting the column density distribution has to do with the specific cosmological model, namely, the normalization and shape of the corresponding power spectrum. We study a few variants of the cold dark matter (CDM) model in § 7 for this purpose. It is found that the amplitude and slope of the linear power spectrum on comoving scales of around 2–20 h Mpc\(^{-1}\) are the most important factors in determining the slope of the column density distribution (the equation of state also has a weak effect on it). Decreasing the amplitude and/or steepening the slope of the power spectrum tends to steepen the distribution in the column density range \(\sim 10^{12.5} - 10^{14.5} \text{ cm}^{-2}\). We introduce an expression relating the slope of the column density distribution to the equation of state and properties of the power spectrum on small scales.

We then study a class of cold plus hot matter (CHDM) models in § 8, making use of the insights gained in §§ 6 and 7. The \(\Omega = 0.2\) (density parameter in neutrino) models have steeper column density distributions compared with those with \(\Omega = 0.1\) because they have less power on the relevant scales. In particular, the low Hubble constant \((H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1})\) \(\Omega = 0.2\) models predict slopes that are steeper than the observed one for most of the parameter space specifying the properties of the intergalactic medium. Only for equations of state that are close to isothermal can the two be made consistent with each other. We emphasize however that a more detailed comparison between the models and observations, taking fully into account instrumental noise and biases of the line identification method(s), is necessary before one can firmly reject any model. We conclude in § 9.

It is appropriate that we mention here two recent pieces of work along similar lines as described above, but using a different dynamical approximation, namely, the lognormal approximation: Bi & Davidsen (1997) and Gnedin & Hui (1996). The former, in particular, contains a very comprehensive and careful analysis of the many different observational properties of the Lyα forest. One strong point of their analysis is that they tested their method using VPFIT, a spectral analysis routine that is commonly used by observers. We will discuss the predictions for the column density distribution by the lognormal and the Zeldovich approximations in § 5.

In our notation, bold-faced letters are reserved for three-dimensional vectors. The symbols \(v_{\text{pec}}\) and \(x\) denote the three-dimensional peculiar velocity and comoving position while \(v_{\text{pec}}\) and \(x\) are their counterparts along the line of sight of interest. Standard symbols are used for cosmological parameters: \(H\) for the Hubble constant as a function of \(z\),
$H_0$, for the Hubble constant today, $h$ for $H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$, $\Omega_0$ for the density parameter today, with the subscript $b$ to denote its baryonic portion and $\nu$ its neutrino content. We use the symbol $h$ (as distinct from $h$) to denote the Planck constant in the few places where it arises. The term *multiple streaming* is reserved for the situation in which a single observed redshift corresponds to more than one position in real space. We distinguish it from the term *orbit crossing*, which is commonly used interchangeably with multiple streaming in other contexts. Orbit crossing refers to the case in which a single position has more than one velocity.

2. COSMOLOGICAL Ly$\alpha$ ABSORPTION IN A FLUCTUATING MEDIUM: BASIC RESULTS

A photon emitted with energy higher than 10.196 eV (wavelength of 1216 Å) by a distant quasar is continuously redshifted as it travels through the intergalactic medium, until it reaches the observer. At some intermediate point, the photon is redshifted to around 1216 Å in the rest frame of the intervening medium, which may contain neutral hydrogen. It can then excite the Ly$\alpha$ transition and be absorbed. Let us consider a particular line of sight from the observer to the quasar. The optical depth $\tau$ (the probability of transmission is given by $e^{-\tau}$) of a photon at a given observed frequency $\nu_o$ is given by

$$\tau(\nu_o) = \int_{x_A}^{x_B} n_{\text{HI}} \sigma_z \frac{dx}{1+z},$$

where $x$ is the comoving radial coordinate of some intermediate point along the line of sight, $z$ is the redshift, and $n_{\text{HI}}$ is the proper number density of neutral hydrogen at that point. The limits of integration, $x_A$ and $x_B$, are the comoving positions of the observer and the quasar. The Ly$\alpha$ absorption cross section is denoted by $\sigma_z$. It is a function of the frequency of the photon with respect to the rest frame of the intervening medium, $\lambda$, which may contain neutral hydrogen. It can then excite the Ly$\alpha$ transition and be absorbed. Let us call this frequency $v$. The cross section is peaked when $v$ is equal to the Ly$\alpha$ frequency, $v_{\text{Ly\alpha}}$.

The frequency $v$ is related to the observed frequency $\nu_o$ by

$$v = \nu_o (1 + z) \left(1 + \frac{v_{\text{pec}}}{c}\right),$$

where $v_{\text{pec}}$ is the peculiar velocity along the line of sight at position $x$ and $1 + z$ is the redshift factor due to the uniform Hubble expansion alone at the same position. The peculiar velocity of the observer, which merely displaces the whole spectrum by a constant amount (independent of $x$), is ignored. The quantity $v_{\text{pec}}/c$, where $c$ is the speed of light, is much smaller than 1.

It proves convenient for later discussion to expand $z$ around some mean redshift of interest $\bar{z}$, which could be the redshift of a simulation output or the average redshift of an observed spectrum with limited redshift range. Using $dx = c dt/a$, where $a$ is the Hubble scale factor and $t$ is the proper time, it can be shown that

$$v = v_o \left(1 + \bar{z}\right) \left(1 + \frac{u/c}{1 + \bar{z}}\right), \quad u \equiv \frac{H}{1 + \bar{z}} (x - \bar{x}) + v_{\text{pec}}(x),$$

where $\bar{x}$ is the position at which the redshift due to Hubble expansion coincides exactly with $\bar{z}$. The Hubble constant at the same redshift is denoted by $\bar{H}$. We assume that the range of $x$ is small enough so that $u/c \ll 1$. The convention that $a = 1$ today is adopted.

The velocity coordinate $u$ defined above contains contributions from both the Hubble expansion and the peculiar motion. Without peculiar motion, $u$ increases monotonically with $x$ and is in fact linear in $x$. Peculiar velocities destroy the linear relation and could give rise to situations in which a given $u$ corresponds to more than one position $x$. This implies that a photon of a given observed frequency $\nu_o$ can have the same rest-frame frequency $v$ at more than one place in its trajectory from the quasar to the observer. As noted above, we reserve the term *multiple streaming* to describe this situation and distinguish it from *orbit crossing*, in which a given $x$ carries more than one $v_{\text{pec}}$ or $u$. We will return to the subject of multiple streaming in § 4.2 and that of orbit crossing in § 3.1.

We define one more velocity coordinate, $u$, which is related to the observed frequency $\nu_o$ by

$$v_o = \frac{v_o}{1 + \bar{z}} \left(1 - \frac{u}{c}\right),$$

where $v_o$ is the Ly$\alpha$ frequency. The velocity coordinate $u$ is simply equal to $u$ when $v$ coincides exactly with $v_o$ (this can be seen by comparing eqs. [3] and [4], bearing in mind that $u/c$ and $u/c$ are both assumed to be much less than 1).

With the definitions in place, we change the variable from $x$ to $u$ in equation (1), which results in the following expression for $\tau$, now a function of $u$:

$$\tau(u) = \int_{u_A}^{u_B} n_{\text{HI}} \frac{du}{1 + z} \left(1 - \frac{u}{c}\right) dx \left(1 + \frac{v_{\text{pec}}}{c}\right),$$

where $\sigma_z$ is reserved for the situation in which $\nu$ coincides exactly with $v_o$ (as distinct from $\sigma_z$). We define one more velocity coordinate, $u$, which is related to the observed frequency $\nu_o$ by

$$v_o = \frac{v_o}{1 + \bar{z}} \left(1 - \frac{u}{c}\right),$$

where $v_o$ is the Ly$\alpha$ frequency. The velocity coordinate $u$ is simply equal to $u$ when $v$ coincides exactly with $v_o$ (this can be seen by comparing eqs. [3] and [4], bearing in mind that $u/c$ and $u/c$ are both assumed to be much less than 1).

With the definitions in place, we change the variable from $x$ to $u$ in equation (1), which results in the following expression for $\tau$, now a function of $u$:

$$\tau(u) = \int_{u_A}^{u_B} n_{\text{HI}} \frac{du}{1 + z} \left(1 - \frac{u}{c}\right) dx \left(1 + \frac{v_{\text{pec}}}{c}\right),$$

The summation refers to a sum over multiple streams (all the $x$'s within the range $x_A - x_B$ that correspond to a given $u$), and $n_{\text{HI}}$, $z$, and $|du/dx|$ are now functions of $u$. The limits of integration $u_A$ and $u_B$ are the velocity coordinates corresponding to the positions $x_A$ and $x_B$ (assuming no orbit crossing, so that each $x$ carries one $u$). Note that, in practice, only a limited range of $u$ contributes to $\tau$ for a limited range of $u$, so that one can replace the redshift $z$ with $\bar{z}$. The same is also true for equation (1).

The Ly$\alpha$ cross section is expressed as a function of $u - u_o$. The constant $\sigma_{\text{HI}}$ is equal to the combination of fundamental physical constants $0.416\pi \sigma^2/(m_e c v), \pi$, where $q$ is the charge of an electron and $m_e$ is its mass. It is about $4.5 \times 10^{-18}$ cm$^2$. The parameter $b$ is equal to $(2k_B T/m_e)^{1/2}$, where $k_B$ is the Boltzmann constant, $m_e$ is the mass of a proton, and $T$ is the temperature of the gas at the velocity coordinate $u$.

The form of the line profile function above takes into account thermal broadening but ignores the natural line width. A more general profile function involves a convolution of the two, resulting in the Voigt profile (Rybicki & Lightman 1979). However, Voigt profiles are accurately thermal profiles for column densities less than about $10^{17}$ cm$^{-2}$. The reader is referred to Spitzer (1978) and Press & Rybicki (1993) for discussions of curve-of-growth analysis.

Note also that it is sometimes assumed that $b$ contains a component due to turbulent motion. We do not include it explicitly in our formalism. Bulk motion, on the other hand, is accounted for by $v_{\text{pec}}$ or $u$. No. 2, 1997 COLUMN DENSITY DISTRIBUTION OF Ly$\alpha$ FOREST
Let us consider two different limits of equation (5).

Suppose there is a high local maximum in \( n_{HI} |du/dx|^{-1} \) at some \( u = u_{\text{max}} \) with width in velocity space much smaller than the thermal width \( b \). Then one can take the line profile function associated with \( \sigma_x \) out of the integral in equation (5) because \( n_{HI} |du/dx|^{-1} \) varies much more rapidly than the thermal profile:

\[
\tau(u_o) = \left[ \int_{u_{\text{max}}}^{u_o} \frac{dx}{1 + z} \right] \sigma_{x,0} \frac{c}{b\sqrt{\pi}} e^{-\left((u-u_{\text{max}})^2/b^2\right)},
\]

where the variable of integration has been changed back from \( u \) to \( x \). This equation holds when \( u_o \) is close enough to \( u_{\text{max}} \). The integral is over the local maximum, assuming that the amount of neutral hydrogen away from the maximum does not cause significant absorption (until another maximum is encountered). One then sees an absorption line with a Gaussian profile in optical depth. While the width of the line tells us about \( b \), which is proportional to the square root of the temperature, the depth of the line provides information about both \( b \) and the column density, which is the integral inside the brackets on the right-hand side. Let us call this the narrow-maximum limit.

Consider another limit of the integral (eq. [5]), in which \( n_{HI} |du/dx|^{-1} \) varies slowly with \( u \). Suppose the scale of variation is much larger than the thermal width. In this case, one can leave the line profile function inside the integral but take the rest outside:

\[
\tau(u_o) = \sum \frac{n_{HI}}{1 + z} \left( \frac{du}{dx} \right)^{-1} c\sigma_{x,0}.
\]

The velocity-dependent terms on the right-hand side are evaluated at \( u_o \). The profile function has been integrated out.

In the above limit, \( \tau \) does not necessarily have the thermal profile around its maxima. We will call this the broad-maximum limit. An extreme example is that of a homogeneous medium, which gives rise to featureless and uniform absorption (Gunn & Peterson 1965).

Conventional analysis of quasar spectra involves identifying those parts of the spectra that are due to the \( \text{Ly} \alpha \) absorption and fitting them with superpositions of Voigt profiles (of which the thermal profiles are a subset) until the residual signal is consistent with noise. This technique was motivated by the picture of the intergalactic medium as consisting of a smooth component that causes relatively little absorption and a set of clouds that satisfy the narrow-maximum limit. For each cloud, the best-fit Voigt profile then yields its temperature and column density according to equation (6).

However, it is clear that for a general, fluctuating medium not all maxima in \( \tau \) necessarily satisfy the conditions leading to equation (6). In fact, according to most structure formation theories, there invariably exist fluctuations in the intergalactic medium on scales larger than the thermal width. In the broad-maximum limit, the shape of a local maximum in optical depth is determined by the distributions of \( n_{HI} \) and \( |du/dx| \) around it. Each maximum in \( \tau \) can still be identified as an absorption line, and one can even apply standard techniques and try to fit its shape with superpositions of Voigt profiles. Given the best-fit Voigt profiles, one can assign a \( b \)-value (width of the profile) and a column density to each profile, but it is no longer true, for instance, that the \( b \)-value thus obtained is equal to \((2k_B T/m_p)^{1/2} \) (eq. [5]). A reasonable question to ask is whether there are other practical methods of identifying absorption lines and assigning column densities without assuming that every absorption line consists of a superposition of Voigt profiles. This will be discussed in § 4.1. It should be borne in mind, however, that all existing observational data on the column density distribution are obtained using the Voigt profile fitting technique, so for the purpose of comparing theory with these observations, it is important that the line identification algorithm one uses yield results consistent with the profile-fitting technique.

In general, there are regions of high density and limited extent, galaxies, for instance, that give rise to absorption profiles well approximated by the narrow-maximum limit, but there are also regions in the intergalactic medium with gentle fluctuations where the broad-maximum limit holds. Then there are those places where neither limit applies, in which case a full integration of equation (5) has to be carried out to compute the optical depth. To do so, one needs to know how the neutral hydrogen density, peculiar velocity, and temperature vary with spatial position. This is the subject of the next section. In any case, the above discussion should make it clear that the quasar absorption spectrum contains a wealth of information about the intergalactic medium.

3. INGREDIENTS FOR GENERATING QUASAR ABSORPTION SPECTRA

There are four quantities that go into the computation of the optical depth: temperature, peculiar velocity, overdensity, and neutral fraction. That the temperature and peculiar velocity are important should be obvious from the expression for the absorption cross section in equation (5). The temperature determines the extent of thermal broadening (\( b \)), and the peculiar velocity changes the frequency of the photon in the fluid rest frame (eq. [3]). Let us define carefully what we mean by the other two quantities, the overdensity and the neutral fraction.

Suppose \( n_\text{HI}(x) \) is the total proper number density of all hydrogen species at position \( x \) and \( \tilde{n}_\text{HI} \) is its spatial average. The overdensity \( \delta_b \), which describes the variation in space of \( n_\text{HI}(x) \), satisfies

\[
\tilde{n}_\text{HI}(x) = \bar{n}_\text{HI}[1 + \delta_b(x)], \quad \rho_b(x) = \bar{\rho}_b[1 + \delta_b(x)].
\]

In the first expression, \( \delta_b \) is defined as the number overdensity of hydrogen. In the second expression, we equate \( \delta_b \) with the mass overdensity of baryons (\( \rho_b \) is the proper mass density of baryons and \( \bar{\rho}_b \) is its mean), which is an excellent approximation for our application because there is no significant conversion of hydrogen into other elements, nor is there any interaction that could cause the spatial distribution of hydrogen to deviate significantly from that of other types of baryons.

What the \( \text{Ly} \alpha \) absorption directly probes is not the total hydrogen density but its neutral component. The neutral fraction \( X_{HI} \) is defined by the following relation:

\[
n_{\text{HI}}(x) = n_\text{HI}(x) X_{HI}(x),
\]

where \( n_{HI} \) is the proper number density of neutral hydrogen as a function of space. The neutral fraction is determined by
the balance between recombination and ionization, the rates of which are dictated by the temperature and radiation intensity, respectively.

In general, all four quantities, overdensity $\delta_b$, peculiar velocity $v_{pec}$, temperature $T$, and neutral fraction $X_{HI}$, are functions of position. In the next two subsections, we discuss first how to determine the spatial distributions of $\delta_b$ and $v_{pec}$ and, second, how $T$ and $X_{HI}$ vary with position through their dependence on $\delta_b$. All quantities are evaluated at $z = 3$. Although most of the material in this section is standard textbook fare, it consists of a somewhat unusual combination of methods, so it is worth going through the basic equations and stating our assumptions carefully.

3.1. The Zeldovich Approximation

In cosmological models in which dark matter (a term that we use interchangeably with noninteracting matter) dominates the mass density of the universe, $\delta_b$ as defined in equation (8) coincides with the dark matter overdensity $\delta_{DM}$ on large scales. We define $\delta_{DM}$ in an analogous manner as before (eq. [8]):

$$\rho_{DM}(x) = \rho_{DM}[1 + \delta_{DM}(x)] ,$$

where $\rho_{DM}$ is the mass density of dark matter at position $x$ and $\rho_{DM}$ is its mean. The equality $\delta_b = \delta_{DM}$ is equivalent to the statement that the hydrogen density (which we assume is simply proportional to the baryon density) varies with position in the same manner as the dark matter density does. This is true on large scales where gas pressure is insignificant compared with the gravitational attraction of the dark matter, provided the baryons and dark matter start out having the same spatial distribution, which is approximately true for most popular cosmological models. Moreover, without significant interaction that distinguishes between the two on large scales, the baryons and dark matter share the same peculiar velocity field. On small scales, however, gas pressure can cause the spatial distributions of baryons and dark matter and their velocities to differ. We will return to this point below.

Hence, on sufficiently large scales (how large is “large,” an obviously important question, will be addressed later), it is adequate to know the overdensity and peculiar velocity of the dark matter. The Zeldovich approximation (Zeldovich 1970) is a well-tested approximation to compute the density and velocity distributions of dark matter in the mildly non-linear regime ($\delta_{DM} \lesssim 5$) before orbit crossing. The reader is referred to the article by Shandarin & Zeldovich (1989) for a comprehensive review (see also Hui & Bertschinger 1996 for an alternative formulation of the approximation).

The starting point of the Zeldovich approximation is the following equation for the displacement of a given mass element or particle:

$$x(q, t) = q + D_+(t)\nabla_x \psi(q) ,$$

The coordinate $q$ is the initial position of the mass element and $x$ is its comoving position as a function of time. The displacement is then $D_+(t)\nabla_x \psi(q)$. Its time-dependent part $D_+(t)$ is the linear growth factor (Peebles 1980), which, for a universe with critical matter density, can be equated with $a$, the Hubble scale factor. The time-independent function $\nabla_x \psi(q)$ is determined by initial conditions. Growing-mode initial conditions dictate that it is curl-free, hence its form as the gradient of the potential $\psi$ ($V_q$ is the spatial gradient in $q$-space).

Expressions for the peculiar velocity and overdensity follow immediately from equation (11):

$$v_{pec} = aD_+[\nabla_x \psi] , \quad 1 + \delta_{DM} = \det^{-1} \left[ \delta_{ij} + D_+(t) \frac{\partial^2 \psi}{\partial q_i \partial q_j} \right].$$

(12)

The overdot in the first expression denotes differentiation with respect to proper time $t$. The peculiar velocity is defined by $v_{pec} = a \dot{\mathbf{r}} / dt$. The second expression follows from mass conservation, i.e., $(1 + \delta_{DM})d^3x = d^3q$. The right-hand side of the second expression is simply the Jacobian of the $q$-$x$ transformation matrix.

The function $\psi(q)$ contains all the information about the specific cosmological model one chooses to study. For the cosmological models that we study in this paper, it is a Gaussian random field in $q$-space. Suppose $\psi(k)$ is its Fourier counterpart, defined by $\psi = \int d^3k \psi(k) e^{i \mathbf{k} \cdot \mathbf{r}}$. The two-point correlation of $\psi$ is related to the commonly used power spectrum $P$ by

$$\langle \psi(k) \psi^*(k') \rangle = k^{-4} P(k) \delta^3(k - k') ,$$

(13)

where $P$ is related to the root mean square (rms) linear overdensity fluctuation by

$$\langle \delta^2 \rangle = D_+^2(t) \int_0^\infty 4\pi P(k)k^2 dk .$$

(14)

Note that $D_+^2 = (1 + z)^{-2}$ for a universe with critical matter density, choosing $D_+ = 1$ today.

To produce a realization of the density and velocity fields for a given cosmological model, we employ the following procedure: first, we use the corresponding power spectrum to generate the Gaussian random field $\psi(q)$ on a grid; second, we displace particles from their initial grid positions $(q)$ according to equation (11) for $D_+(t)$ corresponding to $z = 3$; third, a peculiar velocity is assigned to each particle according to the first expression in equation (12); finally, we use the TSC (triangular-shaped density cloud) scheme (Hockney & Eastwood 1988) to interpolate the particle positions and velocities to become momentum and mass densities on the grid and divide one by the other to obtain the velocity itself. The interpolation to obtain mass density is our way of enforcing mass conservation, as is expressed in the second formula of equation (11). In the last procedure, we smooth the momentum and mass density fields over a small number of grid cells (in fact, we use one and have checked that the precise number is not important as long as it is small) before performing the division to obtain the velocity field, so that we have well-defined velocities even in places with zero density after the TSC interpolation (Kofman et al. 1994). Any line of sight can then be chosen through the simulation box, and the above set of steps yield the overdensity and peculiar velocity (in fact, only the component parallel to the line of sight is needed) at each grid point on it.

The procedure just outlined is very efficient because there is no need to integrate any equation of motion. One simply multiplies the displacement of each particle by an appropriate factor of $D_+(t)$. However, the first step of the procedure has to be slightly modified to address two problems.

The first problem is orbit crossing. The Zeldovich approximation is known to predict too-rapid growth of the
thickness of the post-collapse pancake (Shandarin & Zeldovich 1989). A number of cures have been proposed (Kofman, Pogosyan, & Shandarin 1990; Matarrese et al. 1992; Brainerd, Scherrer, & Villumsen 1993; Bagla & Padmanabhan 1994), but the one that consistently gives good agreement with N-body simulations is the truncated Zeldovich approximation (Kofman 1991; Coles, Melott, & Shandarin 1993; Melott, Buchert, & Weiss 1995). The basic idea is to smooth the initial power spectrum on small scales so that the amount of orbit crossing that might have occurred by the time of interest is not significant enough to destroy the accuracy of the Zeldovich approximation on large scales, where the fluctuations are still mildly nonlinear. The initial power spectrum \( P(k) \) is multiplied by a Gaussian smoothing kernel of the form \( \exp \left( -k^2/k_0^2 \right) \) before it is used to generate the Zeldovich displacement field (eq. [13]). This is equivalent to smoothing the initial density field \( \delta(x) \) through the convolution \( \left( 2\pi k_0^2 \right)^{-0.5} \int \delta(x') \exp \left( -k_0^2 |x - x'|^2 / 2d^3x' \right) \). The smoothing wavenumber \( k_s \) is chosen according to the following prescription:

\[
k_s = 1.5k_{NL} ,
\]

where

\[
1 = D_\Lambda^2(t) \int_0^{k_{NL}} 4\pi P(k)k^2 dk .
\]

Note that \( P(k) \) above is the initial power spectrum before any smoothing. The proportionality constant between \( k_s \) and \( k_{NL} \) actually depends somewhat on the power spectrum, with more smoothing (smaller \( k_s \)) required for models that have relatively more power on small scales (Melott 1994). The choice above has been shown to yield good agreement with N-body simulations for CDM models (Melott et al. 1995). We will see that for those CHDM models with relatively little power on small scales, the precise value of \( k_s \) is not important. The procedure described above is commonly called the truncated Zeldovich approximation.

The second problem is one we have pointed out before, namely, that \( \delta_{DM} \) is not necessarily equal to \( \delta_b \) (which is what we are interested in ultimately) on small scales. In linear theory, it is possible to show that for a dark matter-dominated universe, the Fourier components of the two quantities obey \( \delta_b(k) = \delta_{DM}(k) \) when \( k \ll k_J \) and \( \delta_b(k) = k_J^2 \delta_{DM}(k)/k^2 \) when \( k \gg k_J \). Under some restrictive assumptions (see Appendix A), it can be shown that \( \delta_b(k) = \delta_{DM}(k)(1 + k^2/k_J^2)^{-1} \). The quantity \( k_J^{-1} \) is known as the Jeans scale and is defined by

\[
k_J^{-1} = \sqrt{\frac{\gamma k_B T}{4\pi a^3 G\mu \rho_b / \rho_{DM}}} ,
\]

where \( k_B \) is the Boltzmann constant, \( T \) is the spatially averaged temperature of the gas, \( \mu \) is the mean mass per particle (for fully ionized gas with primordial abundances it is about 0.6\( m_p \), where \( m_p \) is the mass of the proton), and \( \gamma \) describes the relation between the temperature \( T \) (the actual, not average, value) and \( 1 + \delta_b ; T \propto (1 + \delta_b)^{\gamma-1} \). Note that \( \gamma \) does not necessarily equal 5/3 unless the gas behaves adiabatically. The proofs of the above assertions can be found in Appendix A (see also Bi et al. 1992; Peebles 1993). It is sufficient to note here that in the linear regime, the baryon density field is smoother than that of the dark matter on small scales because of the effect of gas pressure.

Now, the above relations between \( \delta_b \) and \( \delta_{DM} \) hold only in the linear regime when both quantities are small. To take into account the effect of gas pressure in the mildly nonlinear regime, one possibility is to smooth the initial power spectrum by a factor of \( (1 + k^2/k_J^2)^{-\frac{1}{2}} \) before generating the displacement field, similar to what is done in the case of the truncated Zeldovich approximation. This method was used by Reisenegger & Miralda-Escudé (1994) to study the fluctuating Gunn-Peterson effect. In practice, we smooth the initial power spectrum by a Gaussian kernel \( \exp \left( -k^2/k_J^2 \right) \) and find that the two ways of smoothing yield very similar column density distributions.

To give an idea of scale, for \( \gamma = 1.5, T = 10^4 K \), and a universe at critical density, \( k_J \) is equal to 16.8 h Mpc\(^{-1} \). It turns out that for all models considered in this paper except the \( \Omega_c = 0.2 \) CHDM models, the truncation scales \( k_J^{-1} \) according to equation (15) are larger than \( k_J^{-1} \) (eq. [16]), for reasonable ranges of temperature and \( \gamma \). For these models it is unnecessary to smooth the initial power spectrum over the Jeans scale, because the truncated Zeldovich approximation already prescribes more smoothing. The opposite is true for the \( \Omega_c = 0.2 \) CHDM models. In fact, the amount of small-scale power is so insignificant for these models that the precise scale of smoothing does not affect the column density distribution for column densities of interest (§ 8). Orbit crossing is probably not very severe for this class of models. Hence, uncertainty in the Jeans smoothing scale due to uncertainties in the temperature and equation of state of the intergalactic medium is not a concern.

To sum up, we smooth the initial power spectrum on the scale of \( k_J^{-1} \) (eq. [15]) or \( k_J^{-1} \) (eq. [16]), depending upon which is larger, before it is used to generate the displacement field (eq. [13]) (except for a few test cases discussed in § 7). The rest of the procedure to obtain the overdensity \( \delta_b \) and peculiar velocity \( \nu_{pec} \) on a grid follows as before. The implicit assumption underlying the whole procedure is that fluctuations on scales smaller than the smoothing scale do not contribute significantly to the number of absorption lines at our column densities of interest, about \( 10^{12.5} \)–\( 10^{14.5} \) cm\(^{-2} \). The upper limit is related to the maximum overdensity (\( \delta_b \approx 5 \)) beyond which the Zeldovich approximation is not expected to be reliable, and the lower limit is set by our resolution (see § 5 for more details). Note that while the Jeans-scale smoothing is meant to capture the actual smoothing of the density field by gas pressure, the nonlinear-scale smoothing is an approximation technique to avoid the problem of orbit crossing. As such, the validity of the latter in the present application has to be checked.

We show in § 5 a comparison between the column density distribution computed using the approximate method described here and that using a full hydrodynamic simulation. The level of agreement lends support to our assumption. Another consistency check is to see whether shock heating is important for regions with overdensities (or underdensities) associated with the above range of column densities. A plot of density versus temperature like Figure 2 in Weinberg et al. (1996) shows that shock heating, and by extension orbit crossing, is not important for regions of underdensity or low overdensity.

### 3.2. The Thermal and Ionization State

Given the evolution of \( \delta_b \) predicted by the Zeldovich approximation, it is possible to integrate evolution equations for \( T \) and for \( X_{HI} \), as well as the abundance of other
species, to obtain their relations with \( \delta_b \). Details of the computation will be given in a separate paper (Hui & Gnedin 1996). A brief discussion can be found in Appendix B of this paper. We summarize the main relevant conclusions here.

First, ionization equilibrium is maintained at high accuracy except during the period of initial reionization. Ionization equilibrium implies that the neutral hydrogen fraction (eq. [9]) satisfies

\[
X_{\text{H}} \sim 1.6 \times 10^{-6} \left( \frac{T}{10^4 \text{ K}} \right)^{-0.7} \left( \frac{\Omega_b \Omega^2}{0.0125} \right) \times \left( \frac{J_{\text{H}}}{0.5} \right)^{-1} \left( 1 + \delta_b \right)^{1 + \frac{z^2}{4}} \left( 1 + \frac{x_{\text{H}}}{2} \right)^3, \tag{17}
\]

where we have adopted the approximate form for the recombination coefficient of hydrogen \( 4.29 \times 10^{-13} [T / (10^4 \text{ K})]^{-0.7} \text{ cm}^3 \text{ s}^{-1} \), which is sufficient for our purpose (see Hui & Gnedin 1996 for a more accurate analytical fit). The quantity \( J_{\text{H}} \) is a measure of the radiation intensity defined as follows (analogous to the definition in Miralda-Escudé et al. 1996 but differing by a factor of \( 10^{-21} \) ergs Hz\(^{-1}\) s\(^{-1}\) cm\(^{-2}\) sr\(^{-1}\)):

\[
J_{\text{H}} = \frac{\int_0^\infty 4\pi J_{\text{H}} \, dv/hv}{\int_0^\infty 4\pi \sigma_{\text{H}} \, dv/hv} \times (10^{-21} \text{ ergs Hz}^{-1} \text{ s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1})^{-1}, \tag{18}
\]

where \( J_{\text{H}} \) is the specific intensity as a function of frequency \( v \) in the units given above, \( h \) is the Planck constant, \( 2\pi \hbar v_{\text{H}} \) is \( 13.6 \) eV, and \( \sigma_{\text{H}} \) is the ionization cross section. The photo-ionization rate is simply equal to \( 4 \times 10^{-12} J_{\text{H}} \) s\(^{-1}\).

Observations indicate that \( J_{\text{H}} \) is between about 0.1 and 2.0 for \( z = 2-4 \) (Bakalov, Dickey, & Ostriker 1988; Lu, Wolfe, & Turnshek 1991; Bouchard 1994; Giallongo et al. 1996; Cooke, Espy, & Carswell 1997). A perhaps more common way of characterizing the radiation intensity is to quote its value, often referred to as \( J_{\text{H}1} \) at \( v = v_{\text{H}1} \), or at wavelength 912 Å, in units of ergs Hz\(^{-1}\) s\(^{-1}\) cm\(^{-2}\) sr\(^{-1}\). The relation between \( J_{\text{H}1} \) and \( J_{\text{H}} \) depends on the spectrum. A good approximation for reasonable slopes of the spectrum right above \( v_{\text{H}1} \) (\( J \propto v^{-m} \) for \( m \) between 1 and 1.5) is \( J_{\text{H}} = 0.75 J_{\text{H}1}/10^{-21} \).

Second, we find that

\[
T = T_0 (1 + \delta_b)^{-1}, \tag{19}
\]

where \( T_0 \) is not dependent on position, is a good approximation for overdensities of interest, \( \delta_b \ll 5 \), with a little flattening at the low end (\( \delta_b \) close to 0.1) for some reionization scenarios. We will call this our equation of state. Note that this implies that the spatial dependence of \( T \) (and by extension \( X_{\text{H}} \)) is completely determined by that of \( \delta_b \), which is true for unshocked gas. Similar power-law relations between the overdensity and the temperature can be seen in Figure 2 of Weinberg et al. (1996) for low overdensity.

Since these equations are local, in the sense that each mass element evolves independently of the others, there is actually no need to generate a full three-dimensional realization for the purpose of studying the thermal and ionization evolution. A simpler approach is to generate a set of eigenvalues of the deformation matrix \( \partial^2 \psi / \partial q \partial q \) according to the prescription of Doroshkevich (1970) and determine the density evolution through the second part of eq. (12).

Third, both \( T_0 \) and \( \delta_b \) depend on the reionization history, the reasonable ranges being \( 1.2 < \gamma < 1.7 \) and \( 3000 \text{ K} < T_0 < 30,000 \text{ K} \) at \( z = 3 \). It is shown in Hui & Gnedin (1996) that \( 1.3 < \gamma < 1.62 \) at \( z = 3 \) if the universe reionizes before \( z = 5 \), assuming a uniform radiation field. We allow for a larger range here. Combining equations (19) and (17), it can be seen that the neutral hydrogen fraction is proportional to \( (1 + \delta_b)^{1 - 0.7 (T_0 - 1)} \).

In conclusion to § 3, we have outlined a procedure to use the Zeldovich approximation, with appropriate initial smoothing, to produce a realization of \( \delta_b \) and \( \nu_{\text{rec}} \) as a function of position, and we have also shown how the relations between \( T, X_{\text{H}} \), and \( \delta_b \) can be obtained (eqs. [19] and [17]). All of them enter into the calculation of the optical depth \( \tau \) (eq. [1] or [5]). We can compute \( e^{-\tau} \), called the transmission, which is the ratio of the observed to the emitted intensities. Observationally, its measurement requires knowledge of the quasar emission spectrum. Moreover, one must carefully choose the range of frequencies to consider if one is to limit the source of absorption to that due to the Lyz transition. For a discussion of these issues, the reader is referred to Press et al. (1993). To produce a realistic spectrum, one should also add noise and convolve the transmission with a window function to mimic instrumental resolution. This is important for a detailed comparison between theories and observations, which we defer to later work. Our x-space grid cells, depending upon the particular simulation, have sizes ranging from 0.028 to 0.075 comoving Mpc. Note that the true resolution in velocity space is not uniform, because peculiar velocity varies from one place to another. Without peculiar velocity, the above grid cell sizes correspond to velocity cells of 2.8–7.5 km s\(^{-1}\), for \( h = 0.5 \) at \( z = 3 \) (eq. [3]). The true velocity resolution is probably a little worse than that. As a comparison, high-quality Keck Telescope data have a full width at half-maximum of about 7 km s\(^{-1}\) and signal-to-noise ratio per pixel of the order of 30 or higher (Hu et al. 1995; Lu et al. 1996).

4. The Peculiar Velocity: Its Effects on Line Shapes and the Column Density Distribution

We show in this section that while the peculiar velocity plays an important role in determining the absorption profiles, its effect on the column density distribution is minor. The procedures to obtain the column density distribution are discussed first.

4.1. Line Identification and the Column Density Distribution

Figures 1 and 2 show the velocity, density, and transmission (\( e^{-\tau} \)) along two lines of sight for a \( q = 0.7 \) CDM simulation, with \( h = 0.5 \) (see Table 1). The significance of the dashed transmission profile will be explained in § 4.2. The thermal and ionization parameters are described in the legend of Figure 1. The truncation scale \( k_{\text{c}} \) (eq. [15]) is 2.3 Mpc\(^{-1}\). The transfer function is taken from Ma (1996). We find that using instead the transfer function of Bardeen et al. (1986) makes almost no difference in the resulting column density distribution, for the range of column densities considered.

The first thing to note is that, for the given parameters,

\[
b = 13 (1 + \delta_b)^{1/4} \text{ km s}^{-1}. \tag{20}
\]
might originate from regions at higher $\delta_b$. It is true, though, that the $1/2$ power of $1 + \delta_b$ does not help very much. Second, a distinction should be made between the observed $b$-value and the $b$ defined above. The observed $b$-value is obtained by fitting the quasar spectrum with superpositions of Voigt profiles. Each Voigt profile yields a column density and a $b$-value. All the density peaks that give rise to absorption troughs in Figures 1 and 2 have velocity widths larger than or comparable to the small thermal width defined in equation (20). Therefore the narrow-maximum limit (eq. [6]) does not apply, and the absorption troughs do not exactly have Voigt profile shapes. The $b$-value obtained from the best-fit Voigt profile of a given absorption trough does not necessarily correspond to the thermal width in equation (20). It should also be emphasized that the recent hydrodynamic simulations of the Ly$\alpha$ forest, which have been so successful in accounting for many of its observed properties, have similarly low temperatures (see, e.g., Weinberg et al. 1996$^7$).

One might wonder if there exists an alternative spectral reduction method in which the Voigt profile is not assumed to be the fundamental shape of absorption troughs and, for such a method, how the column density is assigned to each trough. The Voigt profile fitting technique is nonetheless very important because it is how all existing observational data on the column density distribution are obtained.

An alternative line identification algorithm was proposed by Miralda-Escudé et al. (1996) and was also used by Hernquist et al. (1995). First, a transmission ($e^{-\tau}$) threshold is chosen. Any part of the spectrum that is continuously below the threshold is identified as an absorption line. The column density associated with it is defined by

$$ N_{HI} = \frac{1}{\sigma_{x,0}} \int_{line} \tau(u) \frac{du}{c}, $$

where $\sigma_{x,0}$ is defined below equation (5). The limits of integration are taken to be over the absorption line, i.e., where the transmission is continuously below the threshold. Note
that if the narrow-maximum limit or the thin-cloud assumption were to hold, equation (6) could be substituted into equation (21) to show that $N_{HI}$ does correspond to $\int n_{HI}dx/(1+z)$, assuming that the threshold is high enough so that most of the Voigt profile is included in the definition of the absorption line.

Let us call the above procedure the threshold algorithm. We show in Figure 3 the column density distribution computed according to the algorithm (crosses). The quantity $d^2 N_{LyA}/dN_{HI}dz$ denotes the number of absorption lines per unit column density per unit redshift. The reason for the chosen range of column densities will be given in § 5. The transmission is chosen to be at the mean value of 0.89.

The threshold algorithm has the tendency to underestimate the number of absorption lines compared with the Voigt profile fitting technique. One reason is that it does not deblend. In other words, a given absorption line according to the threshold algorithm may contain more than one minimum in transmission. Such an absorption line would be broken up into a few lines if the Voigt profile fitting technique was employed. To demonstrate this effect, we modify the threshold algorithm: for each (parent) absorption line identified, we break it up into individual components (children), where each component is bordered by local maxima in the transmission within the confines of the parent. The column density for each child component is defined similarly as in equation (21), and the limits of integration are taken to be the boundaries of each component. We will call this the threshold-deblending algorithm.

The resulting column density distribution is denoted by squares in Figure 3 for the transmission threshold of 0.89. One can see that indeed the number of lines of low column densities goes up. We should emphasize, however, that the threshold-deblending algorithm cannot be used to analyze observational data without modifications because, in real life, noise creates local transmission maxima within any parent absorption line. For now, the threshold algorithm is adopted as a simple way to identify lines and assign column densities, which we will use to study the effects of the peculiar velocity on the column density distribution.

4.2. The Role of Peculiar Velocities

The following experiment is performed to investigate the importance of peculiar velocities. We generate absorption spectra and compute the column density distribution using the same density field as that used to produce the solid curves in Figures 1 and 2 and the crosses in Figure 3, but we set all peculiar velocities to zero.

Let us first examine some examples of the absorption spectra. The dashed curves in Figures 1 and 2 are the resulting spectra after setting all peculiar velocities to zero.

A comparison of the dashed absorption spectrum with its solid counterpart in each figure shows that the peculiar velocities play an important role in determining the shapes of absorption lines. Without peculiar velocities, the shapes of absorption troughs mirror closely (perhaps a little thermally broadened compared with) those of the density peaks while, with nonzero peculiar velocities, the absorption troughs can have quite different shapes from the underlying density field. Peculiar velocities can add or erase structures. An example of the former can be found in the pair of density peaks around $x = 9$ Mpc and their corresponding absorption profiles in Figure 1. An example of the latter can be found in the density peak(s) around $x = 7$ Mpc and the corresponding absorption trough(s) in Figure 2.

Broadly speaking, the effects of peculiar velocities on absorption spectra fall into three categories. They are distinguished by the value of $du/du$ (is defined in eq. [3]). First, there are regions in space where the peculiar velocity gradient is small, so that $du/du$ is almost equal to its Hubble value $H/(1+z)$ (eq. [3]). An example is the density peak around $x = 2.2$ Mpc in Figure 2. The peculiar velocity shifts the position of the associated absorption trough but does not affect its shape.

Second, there are places where the peculiar velocity gradient is opposite in sign and comparable in magnitude to $H/(1+z)$, in which case $|du/du|$ becomes very small. Suppose also that $|d^2u/du^2|$ is small. The implication is that a small range in $u$ corresponds to a relatively large range in $x$. See, for instance, the density peak(s) around $x = 7$ Mpc in Figure 2, which is a very broad structure in $x$-space but relatively narrow in $x$-space if peculiar velocities are not set to zero. The small $|du/du|$ or the converging peculiar velocity flow around it helps to produce a narrow absorption trough (Fig. 2, second panel from top). Contrast it with the corresponding absorption feature in the top panel of the same figure, where peculiar velocities are set to zero. The limiting case in which $|du/du|$ exactly vanishes is called a velocity caustic (McGill 1990).

Third, there are regions where the peculiar velocity gradient dominates in such a way that $du/du$ is negative and $|du/du|$ is not small. An example can be found around the

---

8 In fact, numerical noise can also have the same effect. We checked for this by deblending local maxima in two ways: local maxima over three cells and local maxima over five cells, with the slope on either side of the maxima not changing signs. It turns out the resulting column density distributions are almost the same.
pair of density peaks at \( x = 9 \) Mpc in Figure 1. This is where multiple streaming occurs. A given range in \( u \) corresponds to disjoint pieces in \( x \)-space. As a result, the shapes of the associated absorption troughs are significantly different from those of the underlying density peaks.

The three categories can be shown to correspond to the three evolutionary stages of a pancake collapsing along the line of sight (McGill 1990). Restricting equations (11) and (12) to one dimension and substituting into equation (3), one obtains

\[
\frac{du}{dx} = a \left( 1 + 2a \frac{d^2 \psi}{dq^2} \right)^{-1} \left( 1 + a \frac{d^2 \psi}{dq^2} \right)^{-1},
\]

where we have assumed a universe of critical matter density so that \( D_0 = a \). Restricting equation (12) to one dimension, it can also be shown that

\[
1 + \delta_b = \left( 1 + a \frac{d^2 \psi}{dq^2} \right)^{-1},
\]

where we have replaced \( \delta_{DM} \) by \( \delta_b \), assuming the appropriate initial smoothing has been carried out as indicated in § 3.1. As \( a \) grows, it can be seen that \( \frac{du}{dx} \) goes through the three different regimes outlined above for negative \( d^2 \psi / dq^2 \).

At the velocity caustic, where \( du/dx = 0 \), it can be shown that \( \delta_b = 1 \) (McGill 1990). This conclusion does not hold in general, of course, because pancakes can collapse in directions different from the line of sight. But it is true that velocity caustics are often found in regions of slight overdensities.

In principle, at a velocity caustic an absorption line can arise even without any variation in the density field at all (eq. [5]). In practice, one expects that converging peculiar velocity flows are accompanied by density peaks. This is consistent with the few examples we have seen.

Next we consider how the column density distribution changes when the same density field is used but all the peculiar velocities are set to zero. This is shown in Figure 3. The mean transmission of the spectra computed with zero peculiar velocities differs from the mean transmission of the full spectra by less than 1%. It is used as the transmission threshold in the line identification procedure for both analyses. The resulting column density distributions are very similar.

Hence the peculiar velocity plays a relatively minor role in determining the column density distribution. It changes the shapes of absorption troughs without altering the overall number of lines and their column densities. This serves to motivate an approximation that we will introduce in the next section.

A final note on velocity caustics: The reader might worry that at a velocity caustic the optical depth may diverge, while it is clear from equation (1) that, for a finite number density of neutral hydrogen, the optical depth should always be a finite quantity. The resolution is that close to a velocity caustic \( u = u_c \), \( du/dx \) varies as \( |u - u_c|^{1/2} \) [provided the second derivative of \( u \) with respect to \( x \) is nonzero—otherwise it will be \((u - u_c)^{3/2}\) if the third derivative does not vanish, and so on, by simple Taylor series expansion; see Shandarin & Zeldovich (1989) for a similar analysis applied to real caustics, as opposed to velocity caustics; orbit crossing occurs in the former but not in the latter]. So, under the integration in equation (5), the optical depth remains finite. We note also that because of the singular nature of \((du/dx)^{-1}\) around \( u = u_c \), the derivation leading to equation (7) breaks down at a velocity caustic.

5. THE STATISTICS OF DENSITY PEAKS

In this section, we explore a simple approximation in which each density peak in \( x \)-space is identified with an absorption line. This is motivated by the facts that peculiar velocities do not play a major role in determining the column density distribution and that each maximum in density corresponds to a minimum in transmission, or maximum in absorption, if the peculiar velocities are set to zero and if the maximum in density is separated from other maxima by a distance larger than that given by the thermal broadening width.

To calculate \( d^2 N_{Ly} / dN_{HI} dz \), we relate \( dz \) and \( dx \) by ignoring peculiar velocities: \( dz = c^{-1}H dx \). Hence

\[
\frac{d^2 N_{Ly}}{dN_{HI} dx} = \frac{d^2 N_{pk} c}{dN_{HI} dx H},
\]

where \( dN_{pk} / dN_{HI} dx \) is the average comoving number density of density peaks along a random line of sight per unit column density and \( H \) is the Hubble constant at the redshift of interest.

For each density peak, we need a simple prescription for assigning a column density. To that end, we perform the following expansion around each density maximum:

\[
\ln n_{HI}(x) = \ln n_{HI}(x_{pk}) + \frac{1}{2} \frac{d^2 \ln n_{HI}}{dx^2} (x - x_{pk})^2.
\]

It is a straightforward Taylor expansion around the position of the peak, \( x_{pk} \). The second derivative in the last term is negative. The rationale behind expanding \( \ln n_{HI} \) rather than \( n_{HI} \) itself is that \( n_{HI} \) is supposed to fall off quickly far away from the peak (until, of course, another peak is encountered). In other words, the above expansion implies that \( n_{HI} \) has a Gaussian falloff (instead of a power-law one if \( n_{HI} \) itself were Taylor expanded). In a sense, this is close in spirit to the Voigt profile fitting technique. Suppose the broad-maximum limit (eq. [7]) holds, so that the local optical depth is simply proportional to the number density of neutral hydrogen if one ignores peculiar velocities. Then, fitting a minimum in optical depth with a Voigt or thermal profile (eq. [5]) is equivalent to fitting the corresponding neutral hydrogen density peak with a Gaussian.

We then assign the following column density to the density peak:

\[
N_{HI1} = \int_{x_{pk}} dx \frac{n_{HI}(x)}{1 + x} = \frac{n_{HI}(x_{pk})}{1 + x_{pk}} \sqrt{2\pi \left( -\frac{d^2 \ln n_{HI}}{dx^2} \right)^{-1}} |_{x = x_{pk}},
\]

where equation (25) has been used and where \( \int_{x_{pk}} \) denotes integration around the peak until it decays sufficiently. All \( x \)-dependent terms on the right-hand side are evaluated at \( x = x_{pk} \). The above equation is also derived by the authors in a separate paper, using the stationary-phase method (Gnedin & Hui 1996).

Using equations (17) and (19), the above can be rewritten
\[ N_{H_1} = 1.63 \times 10^{13} \left( \frac{T_0}{10^4 \text{ K}} \right)^{-0.7} \left( \frac{\Omega_k h^2}{0.0125} \right)^2 \left( \frac{J_{H_1}}{0.5} \right)^{-1} \times \left( \frac{1 + \frac{z}{4}}{4} \right)^{2 \left( \frac{2 - 0.7(y - 1)}{1.65} \right) - 0.5} A \text{ cm}^{-2}, \]  

(27)

where \( A \) is defined by

\[ A \equiv (1 + \delta_0)^2 - 0.7(y - 1) \left( \frac{d^2 \ln (1 + \delta_0)}{dx^2} \right)^{-1/2} \mid_{x = x_{pk}}, \]  

(28)

with \( x \) being measured in Mpc.

We will refer to our method as the density-peak Ansatz. It consists of two parts: (1) associate each density peak in \( x \)-space with an absorption line; (2) assign a column density to each density peak according to equation (27).\(^9\)

Making use of equations (24) and (27), the column density distribution can be written as

\[ \frac{d^2 N_{Ly} \gamma}{d N_{Ly} dx} = 6.25 \times 10^{-14} \left( \frac{T_0}{10^4 \text{ K}} \right)^{0.7} \left( \frac{\Omega_k h^2}{0.0125} \right)^2 \left( \frac{J_{H_1}}{0.5} \right)^{-1} \times \left( \frac{1 + \frac{z}{4}}{4} \right)^{2 \left( \frac{2 - 0.7(y - 1)}{1.65} \right) - 0.5} \frac{c^2 d^2 N_{pk}^2}{H \ d A \ dx} \text{ cm}^{-2}, \]  

(29)

where \( x \) and \( cH^{-1} \) are in the same unit of Mpc. Most of the factors above come from the scaling between \( A \) and \( N_{H_1} \), and \( cH \) provides the conversion from comoving coordinate \( x \) to redshift \( z \). The last factor, \( d^2 N_{pk}/dA \ dx \), is the number density of peaks in \( x \)-space having the quantity \( A \) within the range \( dA \).

Let us define \( \xi \equiv \ln (1 + \delta_0) \). Suppose one is given \( P(\xi, \xi', \xi') d\xi d\xi' d\xi'' \), which is the probability that \( \xi \) and its first and second derivatives with respect to \( x \) fall in the specified ranges at a point. Then

\[ \frac{dN_{pk}}{dx} = \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{0} d\xi'' \mid \xi'' \mid P(\xi, \xi' = 0, \xi''), \]  

(30)

where \( dN_{pk}/dx \) is the integral of \( d^2 N_{pk}/dA \ dx \) over all \( A \) (Bardeen et al. 1986).

By a change of variable and a differentiation and making use of equation (28), one can obtain

\[ \frac{d^2 N_{pk}}{dA \ dx} = \frac{1}{[2 - 0.7(y - 1)]A} \int_{-\infty}^{0} d\xi'' \mid \xi'' \mid P(\xi, \xi' = 0, \xi''), \]  

(31)

where \( \xi \) should be expressed in terms of \( \xi'' \) and \( A \) using equation (28).

Note that the above two equations are completely general and that no assumption about the Gaussianity of the underlying fields has been made. The hard part is of course to come up with the probability function \( P \). The one-point probability distribution of just \( \xi \) or density has been calculated for the Zeldovich approximation (Kofman et al. 1994). We find the one-point joint probability distribution of density and its first and second derivatives along a line of sight difficult to calculate analytically for the Zeldovich approximation. A numerical approach is adopted in this paper, and the number of peaks is counted along random lines of sight in actual three-dimensional realizations. In a separate paper, we discuss an analytical calculation based upon not the Zeldovich approximation but the lognormal approximation, where \( \xi \) is assumed to be a Gaussian random field (Gnedin & Hui 1996). A comparison between the two will be made here.

5.1. Testing the Density-Peak Ansatz

We test the density-peak Ansatz in two different ways. First, we make a scatter plot of the column density obtained using the threshold-deblending algorithm versus the column density obtained by searching for the maximum density peak that contributes to each absorption line identified by using the threshold method and then applying equation (27). The result is shown in Figure 4. It shows that while the agreement is far from perfect, the column densities assigned using the density-peak Ansatz and using the threshold algorithm are broadly consistent.

The important question, however, is whether the density-peak Ansatz, coupled with the Zeldovich approximation with appropriate initial smoothing, yields the correct number of absorption lines as a function of column density. We compare the column density distribution obtained using our approximate methods with that obtained by applying the Voigt profile fitting technique to synthetic spectra from a full hydrodynamic simulation (see discussion below and Zhang et al. 1997 for details). This is shown in Figure 5. Note that the values of \( \gamma \) and \( T_0 \) given in the legend of Figure 5 were obtained directly from the hydrodynamic simulation. The temperature-density relation is not an exact power law but is well approximated by one for the relevant range of densities. We also show in the same figure the column density distribution obtained using the threshold algorithm, coupled with the Zeldovich approximation.
The predictions of the lognormal approximation are shown as well for comparison.

The level of agreement between the exact hydrodynamic computation and our calculation based upon the density-peak Ansatz coupled with the Zeldovich approximation is encouraging. Two sets of points are shown for our approximate calculation using the density-peak Ansatz, one (open triangles) with exactly the same box size and grid spacing as the hydrodynamic simulation and the other (open squares) with a larger box size and smaller grid spacing. They both agree very well with the exact computation. We will explore the effects of changing the resolution in the next subsection. A third set of points (crosses) shows that the threshold algorithm described in §4.2 underestimates the number of lines at low column densities.

The agreement between the results of the hydrodynamic simulation and our approximation method is perhaps telling us something interesting about the low column density systems. Note that for the Zeldovich approximation computations presented in Figure 5, an initial smoothing scale ($k_s = 2.3\ Mpc^{-1}$) larger than the Jeans length ($k_J = 8.4\ Mpc^{-1}$) was chosen to deal with the problem of orbit crossing. The agreement can then be understood as a result of the Jeans length scales being smaller than the Jeans length. The low column density systems mostly consist of relatively broad density peaks.

Some explanation is also in order regarding the spectral analysis method used by Zhang et al. (1997) in obtaining the distribution shown in Figure 5. It consists of identifying absorption features above a specified opacity and deblending them into individual lines centered at local maxima in optical depth and fitting each with a suitable Voigt profile. This procedure is designed to be similar to most observers’ analysis techniques but is not exactly the same. In particular, most observers’ method consists of seeking a “minimal” set of Voigt profiles, the superposition of which reproduces the input absorption spectrum to within some specified error consistent with noise. It should be noted, however, that the definition of minimal tends to vary from one observer to another. The superposition of Voigt profiles identified using the method of Zhang et al. seems to reproduce the input spectrum quite well (see Fig. 2 of Zhang et al. 1997), although the agreement has not been carefully quantified. One should keep in mind possible differences in the column density distributions obtained using different methods (see Davé et al. 1997 for related discussion).

We also show in Figure 5 two sets of curves based upon the lognormal approximation but using the same density-peak Ansatz (see Gnedin & Hui 1996). One of them has the same amount of initial smoothing as that of the truncated Zeldovich approximation, and the other has less smoothing so as to match the final (not linear) rms density fluctuation of the Zeldovich computation. In both cases, the lognormal approximation tends to predict too much flattening of the column density distribution at low column densities. In general, the lognormal approximation tends to yield column density distributions that deviate quite significantly from power-law unless a very small smoothing length $k_s^{-1}$ is chosen. (From the figure, it might appear that the lognormal approximation yields more lines than the Zeldovich approximation at the very low column densities, but it is really a resolution effect; see §5.2.)

The reader may have noted that we included in Figure 5 a wider range of column densities than is warranted by the nature of our approximations. For instance, objects with column densities higher than $10^{16}\ cm^{-2}$ are almost certainly highly nonlinear, and we do not expect the truncated Zeldovich approximation to work well in this regime. For the low column densities, the finite resolution should cause us to underestimate the number of absorption lines. In the next subsection, we give quantitative estimates of the range of column densities within which the density-peak Ansatz, used in conjunction with the truncated Zeldovich approximation, can be counted upon to yield reliable column density distributions.

We will also discuss two different ways of defining a density peak in the next subsection: local maxima over three cells or local maxima over five cells with the slope on either side of the maxima not changing sign. The three-cell criterion is used in Figure 5. One expects, however, that some of the three-cell peaks are not real but merely artifacts of numerical noise, especially those with low column densities. The five-cell criterion, on the other hand, probably fails to include some narrow peaks that are real. We will see that the two different criteria produce almost identical results above a certain column density.

One aspect of the density-peak Ansatz that we have glossed over is that two density peaks separated by a distance in velocity space much less than the thermal width should be counted as contributing to not two but one
absorption line. A more sophisticated approach would be to group together such density peaks and use the sum of their column densities as the column density of a single absorption line. We find that for the range of validity discussed in the following subsection, it makes little difference. It is conceivable, however, that this effect cannot be ignored for simulations with higher resolution than we have, or at higher redshifts, where line blending is more important.

5.2. The Range of Validity

For the computation presented in Figure 5, the column density (given by the density-peak Ansatz) above which the mean \( \delta_h \) exceeds 5 is about \( 10^{14.1} \) cm\(^{-2}\). For the parameters listed in the legend of Figure 5, \( N_{HI} = 3.6 \times 10^{13} \) cm\(^{-2}\) (eq. [27]). We therefore take \( A = 3.5 \) as an upper limit beyond which we cannot expect our approximations to be reliable. Shock heating should be relatively unimportant for \( \delta_h \) less than about 5 (see Hui & Gnedin 1996).

Note that, according to Figure 5, comparing with the hydrodynamic simulation data, the density-peak Ansatz coupled with the truncated Zeldovich approximation seems to yield a reliable number density of absorption lines even for column densities above \( 10^{14.1} \) cm\(^{-2}\). The level of agreement at such high column densities (and, by extension, such high \( \delta_h \)) is surprising. We will adopt the conservative upper limit of \( A = 3.5 \).

To determine the column density below which finite resolution results in an underestimate of the number of absorption lines, we performed a simulation using the truncated Zeldovich approximation with the same parameters as for the open squares in Figure 5 but with higher resolution: the same box size of 12.8 Mpc but a smaller grid spacing of 0.0284 Mpc. A comparison of the resulting column density distributions is shown in Figure 6.

Note that we have included two definitions of density peaks (three-cell and five-cell). For each simulation, the true column density distribution is probably somewhere between the two in the places where they differ.

We take the low column density cutoff to be \( 10^{12.8} \) cm\(^{-2}\) for the lower resolution simulation (box size of 12.8 Mpc, grid spacing of 0.05 Mpc) using the three-cell definition of peaks. It can be seen that the higher resolution simulation differs from the lower one only at column densities less than roughly this cutoff value. Moreover, above this column density the three-cell and five-cell criteria yield almost identical results.

The parameters in the simulations in Figure 6 are such that \( N_{HI} = 3.6 \times 10^{13} \) cm\(^{-2}\) (eq. [27]). Hence the above column density cutoff implies a lower limit of 0.18 for \( A \). From now on, we will use the three-cell definition of density peaks.

For readers interested in applying our formalism at different redshifts, we recommend choosing the appropriate range of \( A \) by using the same methods as above: the upper limit set by nonlinearity and the lower limit set by running simulations of varying resolution. The range is not expected to change significantly with redshift.

In the following section, we systematically investigate how the column density distribution depends on the cosmological parameters and properties of the intergalactic medium. All the simulations presented in the next two sections have the same resolution and box size, 256\(^3\) grid points with a grid spacing of 0.05 Mpc. For each of them, we will only plot the part of the column density distribution that falls within the limits 0.18 < \( A < 3.5 \). The column densities that these limits correspond to depend on the properties of the intergalactic medium and the redshift (eq. [27]). Note that our conservative limits for \( A \) greatly reduce the range of column densities that we can examine, but within these limits we can be reasonably confident that the truncated Zeldovich approximation together with the density-peak Ansatz should yield accurate predictions for the column density distribution.

6. THE COLUMN DENSITY DISTRIBUTION: DEPENDENCE ON IONIZATION FLUX, TEMPERATURE, EQUATION OF STATE, AND MEAN BARYON DENSITY

We use the CDM model presented in Figure 5 to study systematically how the column density distribution depends on properties of the intergalactic medium: the level of radiation background, the mean baryon density, and its thermal properties. The tools that we use to calculate the column density distributions are the truncated Zeldovich approximation and the density-peak Ansatz.

The transfer functions for this model and all the other models considered in this paper are taken from Ma (1996). A summary of parameters of all models can be found in Tables 1 and 2.

6.1. Dependence on Overall Temperature, Ionization Flux, and Baryon Density

Let us first consider the case in which the equation of state is fixed at \( T \propto (1 + \delta_h)^{0.5} \). As is shown in equation
The column density of a density peak with a given \( \delta_b \) (overdensity) is proportional to the following combination of parameters:

\[
F = \left( \frac{T_0}{10^4 \text{ K}} \right)^{-0.7} \left( \frac{\Omega_b h^2}{0.0125} \right) \left( \frac{J_{HI}}{0.5} \right)^{-1}.
\]

Hence, by equation (30), if \( F \) is rescaled by a certain factor (by changing \( T_0 \), \( \Omega_b \), or \( J_{HI} \) or their combinations), the number of absorption lines is also changed by the same factor at an appropriately rescaled column density.

Suppose \( F \) is rescaled to \( F' \) such that \( F' = rF \); then

\[
\frac{d^2N_{Ly\alpha}}{dN_{HI}dz} \bigg{|}_{N_{HI} = rN_{HI}} = \frac{1}{r} \frac{d^2N_{Ly\alpha}}{dN_{HI}dz} \bigg{|}_{N_{HI}}.
\]

This implies that the column density distribution is a pure power law, then in a log-log plot of the number of absorption lines per unit column density per unit redshift versus column density, the straight line would simply be shifted to the right or left (or up/down) by rescaling. In reality, the column density distribution only approximately obeys a power law, and so there should be a slight change of slope at any given column density as a result of rescaling.

The effects of rescaling can be seen clearly in Figure 7, where \( F \) is allowed to take the values 0.25, 1, and 5. Keeping \( \Omega_b h^2 = 0.0125 \) and \( J_{HI} = 0.5 \), this corresponds to changing \( T_0 \) from about 72,000 K to 1000 K. Alternatively, keeping \( T_0 \) and \( \Omega_b h^2 \) fixed at their canonical values (as shown in eq. [32]), it corresponds to allowing \( J_{HI} \) to vary between 2 and 0.1. (See Hui & Gnedin 1996 for a discussion of the dependence of \( T_0 \) on reionization history. \( T_0 \) is expected to fall within the range quoted above.)

The conventional value of \( \Omega_b h^2 = 0.0125 \) has been challenged by recent measurements of light-element abundances in high-redshift absorption systems. Tytler & Burles (1996) obtained a value of 0.024, which for \( T_0 = 10^4 \text{ K} \) and \( J_{HI} = 0.5 \) implies \( F = 3.686 \), well within the range of \( F \) plotted in Figure 7. The analysis of Rugers & Hogan (1996), on the other hand, favors the value 0.006, which means \( F = 0.23 \) for the same values of \( T \) and \( J_{HI} \). The lowest set of points in Figure 7 has to be lowered further to accommodate this value of the baryon density.

The observational data are taken from Hu et al. (1995), measured at a redshift of about 3 and corrected for incompleteness. The incompleteness correction was obtained by applying the same spectral analysis technique for the observed data to simulations of randomly distributed Voigt profiles with a known column density distribution and measuring how much the analysis method underestimates the number of low column density lines. The amount ranges from about no correction necessary at \( N_{HI} > 2 \times 10^{13} \text{ cm}^{-2} \) to a factor of 5 increase in the number of lines at \( N_{HI} < 3 \times 10^{13} \text{ cm}^{-2} \).

We note in passing that, strictly speaking, altering \( \Omega_b \), in addition to rescaling the number of absorption lines as discussed above, also changes the transfer function in a non-trivial way. But the effect is very small for models in which the dark matter (nonbaryonic) dominates. In fact, using the BBKS (Bardeen et al. 1986) transfer function, which does

![Figure 7](image)
not take into account the effect of baryons at all, instead of the Ma (1996) transfer function, which does take it into account, for $\Omega_b h^2 = 0.0125$ with $h = 0.5$, yields essentially the same column density distribution for the range of column densities discussed here. For models in which the baryon content is proportionally higher, low-density CDM models, for instance, changing $\Omega_b h^2$ has a more pronounced effect on the transfer function.

6.2. Dependence on the Equation of State or the Temperature-Density Relation

Let us hold fixed $T_0$, $\Omega_b h^2$, and $J_{HI}$, at their canonical values as shown in equation (32) but allow the equation of state to change, for the same CDM model as above. As pointed out in § 3.2, the temperature-density relation for low enough overdensity is well approximated by a power law in which the power index is around 0.5, but it can change slightly depending upon the reionization history. We plot in Figure 8 the column density distributions for $\gamma = 1.2$, 1.5, and 1.7, where $\gamma$ is defined by $T \propto (1 + \delta)^{\gamma-1}$. This should adequately cover the possible range of $\gamma$ (Hui & Gnedin 1996).

The first thing to note is that the column density distribution remains almost the same for the three different values of $\gamma$. This is because $\gamma$ affects column density through the power index of $1 + \delta$, which is $2 - 0.7(\gamma - 1)$ (eq. [28]). The index does not change significantly for the range of $\gamma$ considered. A larger index (smaller $\gamma$) means, for a density peak with a given $1 + \delta$, (and its second derivative), that the column density is larger or smaller depending upon whether $1 + \delta$ is larger or smaller than 1. The net effect is to decrease the slope of the column density distribution. The effect, though very small for the values of $\gamma$ plotted, can still be seen in Figure 8. We also show the approximate slopes given by an analytical formula (eq. [41]), which will be discussed later. Note how the column density distribution does not exactly follow a power law but can be approximated by one.

Hence, as a crude approximation, we conclude that the mean temperature, radiation intensity, and baryon density mainly determine the overall normalization of the column density distribution. The equation of state, on the other hand, mostly affects the slope of the column density distribution, but its effect is small for a reasonable range of $\gamma$.

7. The Slope of the Column Density Distribution

It has been shown that the normalization of the column density distribution is influenced by the thermal and ionization states of the intergalactic medium, which are not well constrained observationally. The slope of the distribution, on the other hand, is only weakly affected by the equation of state or the temperature-density relation.

The slope of the column density distribution is therefore relatively free of uncertainties due to our ignorance of the thermal and ionization properties of the intergalactic medium. We now turn our attention to the effect of the power spectrum on the slope of the distribution.

From equations (27) and (28), it can be seen that the column density $N_{HI}$ is proportional to $(1 + \delta)^{2 - 0.7(\gamma - 1)}$, times $1/(\xi^\gamma)^{1/2}$, which basically defines a length scale. Taking into account the correlation between this length scale and the overdensity, we find from our simulations (which use the Zeldovich approximation) a useful approximate relation for column densities between $10^{12.5}$ and $10^{14.5}$ cm$^{-2}$ (or, more accurately, for the range of validity discussed in § 5.2):

$$N_{HI} \propto (1 + \delta)^{1.68 - 0.7(\gamma - 1)}.$$ (34)

which roughly means that the length scale $1/(\xi^\gamma)^{1/2}$ is approximately proportional $^{10}$ to $(1 + \delta)^{-0.32}$.

Since we are interested in the slope of the column density distribution, the relevant quantity to consider is

$$-\beta \equiv \frac{d \ln (d^2 N_{Ly}\) / d$N_{HI}\) dz}}{d \ln N_{HI}} = -1 + \frac{m}{1.68 - 0.7(\gamma - 1)},$$

$$m \equiv \frac{d \ln \int_0^{\infty} d\xi^\gamma P(\xi, \xi^\gamma = 0, \xi^\gamma = 0)}{d\xi^\gamma}. \quad (35)$$

The equality follows from equations (34), (27), (30), and (31) and noting that $\xi \equiv \ln(1 + \delta)$. The column density distribution can be approximated by the simple power law $N_{HI}^{-\beta}$ if $\beta$ defined above is only weakly dependent on $\xi$ or $N_{HI}$.

Lacking an analytical expression for $P$ under the Zeldovich approximation, we cannot determine whether the general properties of the quantity $m$ are. First of all, $m$ depends on $\xi$ in general because the integral $\int_0^{\infty} d\xi^\gamma P(\xi, \xi^\gamma = 0, \xi^\gamma = 0)$ cannot be a simple power law in $\xi$. This is because we expect the integral to vanish for very large and very small $\xi$ and to peak at some intermediate $\xi$. This implies that one should not expect an exact power law for the column density distribution, although pieces of it might be approximated by a power law. Suppose $\xi_{pk}$ is the value of $\xi$ at which the integral $\int_0^{\infty} d\xi^\gamma |P(\xi, \xi^\gamma = 0, \xi^\gamma = 0)|$ reaches its maximum value; then the column densities we are interested in must correspond to $\xi > \xi_{pk}$. where $d \ln \int_0^{\infty} d\xi^\gamma |P(\xi, \xi^\gamma = 0, \xi^\gamma = 0)|/d\xi^\gamma$ is negative. This is

10 A log-log plot of $1/(\xi^\gamma)^{1/2}$ vs. $1 + \delta$ actually shows much scatter, but eq. (34) appears to capture the overall dependence of $N_{HI}$ on $1 + \delta$. 

---

Fig. 8.—Column density distributions of the CDM1a model (Table 1) for three different equations of state (eq. [19]). Open squares, $\gamma = 1.2$; crosses (same as crosses in Fig. 7), $\gamma = 1.5$; open triangles $\gamma = 1.7$. F = 1 (defined in eq. [32]) for all three. Points with error bars are the same observational data as in Fig. 7. Long-dashed and short-dashed lines have the approximate slopes ($\beta = 1.62$ and $\beta = 1.48$) (normalization is chosen by hand) as given in eq. (41) for the open triangles and open squares, respectively.
based upon the knowledge that the computed (as well as observed) slope in equation (35) is less than $-1$ [the factor $1.68 - 0.7(\gamma - 1)$ is always positive for reasonable values of $\gamma$].

Under the Zeldovich approximation, the quantity $\xi$ and its derivatives can be related to the displacement potential $\psi(q)$ by using equations (11) and (12), and so one can express $\mathcal{P}(\xi, \xi', \xi'')$ in terms of the probability density for derivatives of $\psi$: 

$$
\int_{-\infty}^{\infty} \frac{d\xi''}{\xi'' | P(\xi, \xi = 0, \xi'') = \int_{-\infty}^{\infty} \frac{d\xi'}{\xi'} d\psi_{\psi, \psi} (d\psi_{\psi, \psi, \psi, \psi}) | e^{-\xi} P(\psi_{\psi, \psi, \psi}) \times \delta_0(\xi' - \xi(\psi_{\psi, \psi, \psi, \psi})) \times \delta_0(\xi'' - \xi''(\psi_{\psi, \psi, \psi, \psi})) ,
$$

(36)

where $\delta_0$ denotes the Dirac delta function, $\psi_{\psi}$ is the derivative of the displacement potential $\psi$ with respect to $q_i$ and $q_j$ (eq. [11]), and similarly for $\psi_{\psi, \psi, \psi, \psi}$ and $\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi}$, and $\delta''$ is the unit vector pointing along the line of sight. The quantity $\mathcal{P}(\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi})$ is the probability density of $\psi_{\psi, \psi, \psi, \psi}$ in Lagrangian space, and the factor $e^{-\xi}$ converts it into its counterpart in Eulerian space (see Kofman et al. 1994). The quantity $\xi'$ and its derivatives along the line of sight can be expressed as functions of $\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi}$ and $\delta''$. Isotropy of the universe implies that one can average the above expression over all possible orientations of $\delta''$.

For the cosmological models that we consider in this paper, the probability density $\mathcal{P}(\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi})$ is a multivariate Gaussian function that depends on three parameters, $\sigma_0, \sigma_1, \sigma_2$, defined as follows:

$$
\sigma_j = \frac{D_+(t)}{\sqrt{4\pi k^2}} \int_0^\infty 4\pi k^2 e^{-\xi(k^2)^2} dk, \quad j = 0, 1, 2 \quad (37)
$$

(see Bardeen et al. 1986), where $D_+$ is the linear growth factor, which is equal to $(1 + z)^{-1}$ for a universe at critical matter density and $k$ is an appropriate smoothing scale. This follows from the structure of the various expectation values: $\langle \psi_{\psi, \psi, \psi, \psi} \rangle \propto \sigma_0^2, \langle \psi_{\psi, \psi, \psi, \psi, \psi, \psi} \rangle \propto \sigma_1^2, \langle \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} \rangle \propto \sigma_2^2$ and $\langle \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} \rangle \propto \sigma_1$. The fact that $\langle \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} \rangle$ and $\langle \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} \rangle$ vanish by isotropy means that the probability density factors into two separate multivariate Gaussian functions: $\mathcal{P}(\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi}) = \mathcal{P}(\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi})$, with the first factor depending upon $\sigma_0, \sigma_1$, and $\sigma_2$ and the second factor depending upon $\sigma_1$.

One can replace $\sigma_1$ and $\sigma_2$ by combinations of $\sigma_0$ and the following two new parameters, which Bardeen et al. (1986) defined:

$$
R_* = \sqrt{3} \frac{\sigma_1}{\sigma_2}, \quad \gamma_B = \frac{\sigma_1^2}{\sigma_2 \sigma_0}, \quad (38)
$$

where we have renamed $\gamma_B$ to distinguish it from $\gamma$ that we use in this paper.

While $R_*$ defines a length scale, the quantity $\gamma_B$ is a measure of the slope of the power spectrum. We find it convenient to use the following quantity, $n_{\text{eff}}$, in place of $\gamma_B$:

$$
n_{\text{eff}} = 5\gamma_B - 3 \quad (39)
$$

(Bardeen et al. 1986). It is easy to show that the above quantity coincides exactly with the slope of the power spectrum if it obeys a pure power law. For $\gamma_B = 0.5$, we have $n_{\text{eff}} = -2.33$. The $n_{\text{eff}}$ defined here should be distinguished from $n$ in Tables 1 and 2: $n$ is the slope of the power spectrum at large scales whereas $n_{\text{eff}}$ is the slope at the (small) smoothing scale.

The advantage of the new notation is that it is possible to show, by changing variables $\psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} = \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} R_*, \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} = \psi_{\psi, \psi, \psi, \psi, \psi, \psi, \psi, \psi} R_*$, and $\xi' = \xi' R_*$, that the integrals in equation (36) are independent of $R_*$ except for a normalization factor. This implies, by equations (35) and (37), that the slope $\beta$ depends on the equation of state through $\gamma$ and on the power spectrum only through $\sigma_0$ and $n_{\text{eff}}$.

It is hard to make analytic progress from this point on, because $\xi$ and its derivatives are complicated functions of the derivatives of $\psi$. We resort to our numerical simulations to extract the parameter dependence of $\beta$. Using the above arguments and observing that, for all models considered in this paper, $\sigma_0 \sim 1$ and $n_{\text{eff}} \sim -2.33 (\gamma_B \sim 0.5)$, we assume the following form for $\beta$:

$$
\beta = 1 + \frac{B_1 + B_2 (\gamma_0 - 1) + B_3 (n_{\text{eff}} + 2.33)}{1.68 - 0.7(\gamma - 1)}, \quad (40)
$$

where $B_1, B_2, B_3$ are constant coefficients of a Taylor series expansion of $m$ (eq. [35]).

We determine these constants by computing $\beta$ for a series of CDM models of varying $\sigma_0$ and $\gamma_B$. The column density distributions are obtained by using the Zeldovich approximation together with the density-peak Ansatz, as described in previous sections. We pick out in particular models with similar $n_{\text{eff}}$’s but very different $\sigma_0$’s, and vice versa. It is found that the following expression for $\beta$ fits reasonably well the slope of the column density distributions for the CDM models, as well as the CHDM models that we will also study later:

$$
\beta = 1 + \frac{0.96 - 0.8(\gamma_0 - 1) - 0.4(n_{\text{eff}} + 2.33)}{1.68 - 0.7(\gamma - 1)}, \quad (41)
$$

where $\beta, \gamma, \sigma_0, \text{ and } n_{\text{eff}}$ are defined in equations (35), (19), (37), and (39), respectively.

We show two sets of examples in Figures 9 and 10. The former has two CDM models with the same $n_{\text{eff}}$ but different $\sigma_0$’s, while the opposite is true for the latter. To find models with such properties, we find it necessary sometimes to smooth on scales larger than the orbit-crossing scale $k_*^{-1}$, defined by equation (15) (examples are models CDM3 and CDM4 shown in the figures; see also Table 1). For this reason, the column density distributions computed in this section should be viewed as predictions of the corresponding CDM models at the specified smoothing scales only: they do not necessarily coincide with the predictions of these models if more appropriate smoothing scales are chosen. Nor should the CDM models shown in the two figures be considered realistic models of the universe.

Examples of how well equation (41) describes the variation with the equation of state for a CDM model can be found in Figure 8. More examples for CHDM models are also shown in Figures 15, 16, 17, and 18 below.

Before we go on to discuss physical interpretations of the above expression for $\beta$, let us make some general remarks and state a few caveats. First, as we discussed before, $\beta$ is expected to vary with column density but can be approximated as a constant over a restricted range. That a power-law column density distribution is only an approximation is
clear even in some of the figures mentioned above. We find it appropriate to allow for a maximum error bar of $\pm 0.1$ for $\beta$ in equation (41), given $\gamma$, $\sigma_0$, and $\gamma_B$. The coefficients $B_1$, $B_2$, and $B_3$ (eq. [40]) should be viewed as correspondingly uncertain. Furthermore, it should be kept in mind that equation (41) holds only for $\sigma_0 \sim 1$ and $n_{\text{eff}} \sim 2.33$, as the constants $B_1$, $B_2$, and $B_3$ are only meant to be coefficients of a Taylor series expansion.

Finally, one obvious problem of using equation (41) to make predictions for a given cosmological model is that both $\sigma_0$ and $n_{\text{eff}}$ vary with the smoothing scale. An example of how the choice of smoothing scale can affect the column density distribution is shown in Figure 11. The comparison between results of a hydrodynamic simulation and our Zeldovich computation in Figure 5 seems to support the choice of smoothing in equation (15). However, one could reasonably argue that the Jeans scale (eq. [16]) is more physically motivated smoothing scale, while the choice in equation (15) was made only as a device to counter the effect of orbit crossing and hence one cannot be sure that one is not thereby erasing real structures on scales smaller than the orbit-crossing scale but larger than the Jeans length, which can contribute significantly to the low column density Ly$\alpha$ forest. One way to settle this question is to make more detailed comparisons with more hydrodynamic simulations, which is outside the scope of the present work and left for a future paper. It suffices to say that, for models with sufficiently small-scale power such as the CHDM models considered in § 8, which have orbit-crossing scales close to or smaller than the Jeans length, it is probably safe to smooth according to the prescription laid down in § 3.1 and that the precise smoothing scale might not matter very much (see Fig. 13 below). For more nonlinear models, we can only offer the comparison in Figure 5 as one piece of evidence that smoothing on the orbit-crossing scale works reasonably well. However, for whatever smoothing scale one decides upon, we expect equation (41) to hold approximately, provided that $\sigma_0 \sim 1$ and $n_{\text{eff}} \sim 2.33$.

With these being said, let us try to understand qualitatively the parameter dependence of $\beta$ as expressed in equation (41). First, the dependence on $\gamma$: Recall that $N_{\text{HI}} \propto (1 + \delta_s)^{1.68 - 0.7(\gamma - 1)}$ (taking into account the correlation between $1 + \delta_s$ and its second derivative). This means that higher density peaks translate into higher $N_{\text{HI}}$ and vice versa. Lowering $\gamma$ increases the effect of this translation, i.e., a given density peak with $1 + \delta_s > 1$ is associated with a higher column density if $\gamma$ is reduced (the opposite is true for...
1 + δs < 1). The net effect is to stretch a power-law-like column density distribution and make it flatter. The effect is small for a reasonable range of γ. It should be between 1.3 and 1.62 at \(\bar{z} = 3\) if the universe reionizes before \(\bar{z} = 5\) (see Hui & Gnedin 1996). A somewhat larger range is shown in Figure 8.

Second, the dependence on \(\sigma_0\): A cosmological model with lower \(\sigma_0\) is in a more “linear” state of evolution compared with higher \(\sigma_0\) models. In other words, a lower \(\sigma_0\) model has proportionally more intermediate-density peaks compared with high-density ones, hence the steeper column density distribution [associating once again high-density peaks with high column densities using the \(N_{HI} \propto (1 + \delta_b)^{1.68 - 0.7(\gamma - 1)}\) scaling]. For sufficiently low column densities, however, the absorption lines arise from very underdense regions, which should be more common in higher \(\sigma_0\) models. Hence, at very low column densities the high-\(\sigma_0\) model should win: it has more very low density peaks. Where this might occur we cannot tell from our simulations because of the limited resolution. For the range of column densities that we can measure reliably, the slope of the column density distribution simply steepens as \(\sigma_0\) is lowered.

The dependence of \(\beta\) on \(n_{eff}\) is more subtle. To understand it, we resort to a Press-Schechter-type argument (Press & Schechter 1974). According to the Press-Schechter theory, the low-mass slope of \(n_\gamma(M)\) (the number density of clumps of mass \(M\) per unit range \(dM\)) is \(-1.5 + n_{eff}/6\), which means that smaller \(n_{eff}\) (\(n_{eff} < -1\)) implies a steeper slope of the clump number density distribution, exhibiting the same trend as in the case of equation (41). The \(n_{eff}\) dependence arises from the fact that \(n_\gamma(k) \propto k^{(\omega_{lin} + 3)/2}\) \(\sigma_0\) with smoothing scale \(\delta\) and assuming \(M \propto k^{-3}\), together with \(M\gamma(M) \propto k^3/\sigma_0(k)\) (for high \(k\)). The factor of \(k^3\) appears because \(n_\gamma(M)\) is a number density. The dependence of \(n_\gamma(M)\) on \(\sigma_0(k)\) follows from a simple conjecture: the probability that a given point belongs to a clump of mass \(M\) or higher is equal to \(2 \int_{\delta_t}^{\infty} P(\delta) d\delta\), where \(\delta_t\) is a fixed threshold and \(P(\delta)\) is the Gaussian probability distribution for overdensity \(\delta\) with dispersion \(\sigma_0(k)\).

One can carry the above reasoning over to absorption lines, keeping in mind something like the threshold algorithm for identifying lines as the analog of the threshold criterion of the Press-Schechter theory for a collapsed clump. Assume that \(N_{HI} \propto k^{-1}\) (because in the linear regime it is the peak width, not the peak height, that provides the lowest order contribution to column density); it “follows” that \(N_{HI} (d^2 N_{Ly} / dN_{HI} d\bar{z}) \propto k/\sigma_0(k)\), again in the high-\(k\) limit. The factor of \(k\) is to account for the quantity of interest being a one-dimensional number density. It is easy to see that this implies \(\beta = (1 - n_{eff})/2\), i.e., smaller \(n_{eff}\) (larger \(n_{eff}\)) implies a steeper column density distribution. This admittedly crude argument actually yields a value for \(\beta\) that agrees surprisingly well with what we have found using the Zeldovich approximation and density-peak Ansatz. For example, for the CDM model in Figure 5, \(n_{eff} = -2\) (see Table 1) and \(\beta = 1.5\) according to the adapted Press-Schechter argument. This of course brings up the question why there does not seem to be any \(\sigma_0\)-dependence of \(\beta\) in the above argument. It is in fact possible to develop this line of reasoning further to include the \(\sigma_0\)-dependence, but we will leave this to a later paper.

Finally, it is interesting to point out that because most realistic cosmological models have similar power on large scales (COBE or cluster scale), small \(\sigma_0\) on small scales almost necessarily means a steeper falloff of the power spectrum, namely, smaller \(n_{eff}\). So, in practice, both the amplitude \(\sigma_0\) and the slope \(n_{eff}\) of the power spectrum work in the same direction in decreasing/increasing the slope of the column density distribution. Hence, as a rule of thumb, less small-scale power implies a steeper column density distribution, for the column densities of interest in this paper.

8. THE COLUMN DENSITY DISTRIBUTION FOR CHDM MODELS

In this section, we turn our attention to CHDM models, which have relatively little small-scale power and for which the method of truncated Zeldovich approximation is ideally suited. The orbit-crossing scales for them are either smaller than or only slightly larger than the Jeans length, and so one has reasons to believe that any structure erased by our smoothing procedure does not contribute significantly to the number of absorption lines in the range of column densities we are interested in. Predictions for the CHDM models are also of current interest because no hydrodynamic simulations of the Lyα forest have been performed for these models.

We list in Table 2 the CHDM models considered in this section. They are all \(\Omega_0 = 1\) models with \(\Omega_b = 0.05\). Both the \(\Omega_0 = 0.2\) and \(\Omega_0 = 0.1\) versions are considered. They have been shown to agree well with the observational data on large scales (\(k \approx 0.02 - 0.4\) Mpc\(^{-1}\); see Figs. 6 and 7 of Ma 1996). The \(\Omega_0 = 0.3\) models seem to conflict with the observed abundance of damped Lyα systems, which correspond roughly to \(k \approx 0.1 - 1.0\) Mpc\(^{-1}\) comoving in the linear power spectrum (Mo & Miralda-Escudé 1994; Kauffmann & Charlot 1994; Ma & Bertschinger 1994). We include one \(\Omega_0 = 0.3\) CHDM model in Table 2 for the sake of comparison. As shown convincingly by Ma (1996), all models need some amount of tilt to match observations.

We compute as before the column density distribution for each model using the density-peak Ansatz and the Zeldovich approximation with appropriate smoothing. The (density weighted) power spectrum for each CHDM model is taken from Ma (1996).

As discussed in § 7, a plot of rms smoothed linear density fluctuation \(\sigma_0\) versus smoothing scale \(k_s\) (eq. [37]) is a very good indicator of what column density distribution to expect. This is done in Figure 12 for the CHDM models tabulated in Table 2. The no-tilt \(\sigma_0 = 0.7\) CDM model is also plotted for comparison.

Because of neutrino free-streaming, all CHDM models have less power (smaller \(\sigma_0\)) and steeper spectral slope (smaller \(n_{eff}\)) than the CDM model on small scales. Those with more neutrino content (\(\Omega_b = 0.2\)) have even less power than the others. In fact, the \(\Omega_0 = 0.2\) models have \(\sigma_0 < 1\) on all scales larger than the Jeans scale (\(k_s < k_j \approx 10\) Mpc\(^{-1}\)). One expects the Zeldovich approximation to work particularly well for these models because the amount of orbit crossing will not be significant, even without initial truncation.

This is borne out by the next test: we compute the column density distribution for one CHDM model (B2) and examine the effect of choosing different smoothing scales. The result is plotted in Figure 13. The column density distribution in the range plotted does not change much at all for the three different smoothing scales plotted. Contrast this with the case of \(\sigma_0 = 0.7\) CDM (Fig. 11), where the Column Density Distribution
column density distribution is more sensitive to changes in the smoothing scale. This is why the truncation scale has to be chosen with some care for the CDM model: not too small, so that too much small-scale structure is erased. The standard prescription (§ 3.1) seems to work well, according to Figure 5.

For the CHDM model considered (in fact, it holds true for all other \( \Omega_s = 0.2 \) models here), the amount of small-scale power is so insignificant that excluding them by smoothing does not affect the overall column density distribution at all (except possibly that one loses the small-scale fluctuations that can give rise to very low column density absorption, i.e., lower than our resolution limit). We have also performed similar tests for the \( \Omega_s = 0.1 \) models; their response to changes in the truncation scale is somewhere between the \( \sigma_s = 0.7 \) CDM model and the \( \Omega_s = 0.2 \) CHDM models, as can be expected based upon their difference in Figure 12.

We adopt the following truncation scales for the CHDM models: For the \( \Omega_s = 0.1 \) models, the standard prescription described in § 3.1 is used (i.e., \( k_s = 1.5k_{JN} \)). The \( \Omega_s = 0.2 \) models, according to the above prescription, would have truncation scales less than the Jeans length \( (k_s > k_J) \), and so by the arguments presented in § 3.1, \( k_s = k_J \) is adopted. Again, we emphasize that for this class of models, which have relatively little power on small scales, the precise truncation scale is not important. A summary of the truncation scales for all models can be found in Table 2.

The CHDM models with \( \Omega_s = 0.1 \) are plotted in Figure 14. Values of \( F \) that yield a reasonable match to the observational data are chosen for each model. Note how the low Hubble constant models (\( h = 0.5 \)) require a slightly lower \( F \) (eq. [32]) than the higher Hubble constant models. The equation of state is chosen to be the same for all models (\( \gamma = 1.5 \); see eq. [19]). The level of agreement with the observational data, for the given choices of parameters, is satisfactory. Note how the low Hubble constant (\( h = 0.5 \)) models tend to have steeper column density distributions, because they have less power on the relevant scales (see Fig. 12). Their slopes can be brought into better agreement with that of the observational data if a smaller \( \gamma \) is used.

For the \( \Omega_s = 0.2 \) models, we cannot find values of \( F \) that yield the same level of agreement with observations for \( \gamma = 1.5 \). Two examples are shown in Figures 15 and 16. Both have \( h = 0.5 \) and small amounts of tilt. For each, three sets of theoretical predictions are plotted, one for each value of \( F: 1, 2.5, \) or \( 5 \). For \( \Omega_s h^2 = 0.0125 \), the conventional big bang nucleosynthesis value, and \( T_0 = 10^4 \text{ K} \), they corre-
F. 15.—Column density distribution for the A2 CHDM model (Table 3). Three values of $F$ (eq. [32]) are shown: $F = 1$ (open squares), $F = 2.5$ (crosses), and $F = 5$ (open triangles). We choose $\gamma = 1.5$ for all three (eq. [19]). Points with error bars are the observational data as in Fig. 7. The dashed line has the slope of $\beta = 1.86$, as given in eq. (41). The normalization of the line is chosen by hand.

F. 16.—Column density distribution for the B2 CHDM model (Table 2). Three values of $F$ are shown (eq. [32]): $F = 1$ (open squares), $F = 2.5$ (crosses), and $F = 5$ (open triangles), $\gamma = 1.5$ for all three (eq. [19]). Points with error bars are the observational data as in Fig. 7. The dashed line has a slope of $\beta = 1.86$, as given in eq. (41).

F. 17.—Column density distribution for the D2 (Table 2) CHDM model. Three values of $F$ are shown (eq. [32]): $F = 7.14$ (open triangles), $F = 3.57$ (crosses), and $F = 1.79$ (open squares). $\gamma = 1.5$ for all three (eq. [19]). Points with error bars are the same observational data as in Fig. 7. The dashed line has a slope of $\beta = 1.78$, as given in eq. (41).

The response to radiation intensity $J_{H_1}$ of 0.5, 0.2, and 0.1 (eq. [32]). As we have shown before, changing $F$ mainly shifts the sets of points without altering the slope significantly. For column densities between about $10^{12.5}$ and $10^{14.5}$ cm$^{-2}$, the slope of the predicted distribution seems to be too steep compared with the observational data; $\beta$ is about 1.86, with some flattening at the lower column densities, compared with the observed value of about 1.5.

Another $\Omega_r = 0.2$ CHDM model (D2 in Table 2), which has a higher Hubble constant ($h = 0.65$), is shown in Figure 17. (The C2 CHDM model, which also has $h = 0.65$, yields very similar column density distribution.) The slope of its column density distribution is not as steep as the previous ones. This is expected because the higher Hubble constant models have slightly more power on relevant scales, as is evident in Figure 12. In fact, one might argue that the middle set of points in Figure 17, the one having $F = 3.57$, matches the observational data reasonably well if both observational and theoretical errors are carefully taken into account. However, it is still true that these two models predict a steeper column density distribution for $N_{\text{HI}}$ between about $10^{12.5}$ and $10^{14.5}$ cm$^{-2}$, compared with the $\Omega_r = 0.1$ CHDM models (Fig. 14).

It is not hard to understand the column density distributions of the CHDM models presented if one refers back to Figure 12 or Table 2. The $\Omega_r = 0.2$ models have less power (smaller $\sigma_0$) and steeper spectral slope (larger $|n_{\text{eff}}|$) than those with $\Omega_r = 0.1$ on scales $1 \text{ Mpc}^{-1} < k \lesssim 10 \text{ Mpc}^{-1}$, which are relevant for the range of column densities we are interested in. As explained above, the column density distributions are therefore steeper for the $\Omega_r = 0.2$ models at this range of column densities. Among the $\Omega_r = 0.2$ models, those with a lower Hubble constant produce comparatively steeper column density distributions because they have even less small-scale power and steeper spectral slope than those with a higher Hubble constant.

In discussing the predictions of the slope of the column density distribution for these models, one should bear in mind uncertainties due to the choice of smoothing scale, as we have presented in the previous section. However, the fact that these CHDM models have relatively little power on small scales works in our favor: the exact choice of the smoothing scale does not affect the slope of the distribution very much (see Fig. 13). This is especially true for the $\Omega_r = 0.2$ models that have orbit-crossing scales (eq. [15]) smaller than the Jeans length. So, the general conclusion of steeper column density distribution for the tilted CHDM models considered here compared with, say, the CDM1a model (see Table 1 and Fig. 5, the results of which compare favorably with a hydrodynamic simulation) should be robust.
and observations can be made by choosing $D_2$, on the other hand, have intrinsically flatter techniques, is necessary before any model can be considered account noise and biases of the line identification tech-
predictions of the models and observations, taking into

Three values of $\gamma$ are shown: $\gamma = 1.2$ (open squares), $\gamma = 1.5$ (crosses), and $\gamma = 1.7$ (open triangles). Points with error bars are the observational data as in Fig. 7. The dashed line has a slope of $\beta = 1.75$, which is the value given in eq. (41) for $\gamma = 1.2$ and $\sigma_0$ as given in Table 2.

In § 6, we discussed how the equation of state or temperature-density relation can also change the slope of the column density distribution, although the effect is small for a reasonable range of $\gamma$. We show in Figure 18 the effects of altering the equation of state on the column density distribution for one particular CHDM model (A2). $F$ is fixed at 2.5, the value that seems to yield a column density distribution closest to the observational data. Smaller $\gamma$, as we have noted before, helps flatten the column density distribution, but the flattening seems to be not quite enough even for $\gamma = 1.2$. We show in the same figure a dashed line with a slope of $-1.75$ (which follows from eq. [41] by setting $\gamma = 1.2$ and substituting the appropriate values of $\sigma_0$ and $n_{\text{eff}}$ as given in Table 2). It should be kept in mind that we can always shift the column density distribution up and down by rescaling $F$ (§ 6), so the normalization is not important. It seems that $\gamma < 1.2$ is needed for this model to yield the right slope of the distribution, at least the right slope to within the 95% confidence limits of the observed $\beta (1.46^{+0.05}_{-0.03})$. The same conclusion holds for the other low Hubble constant $\Omega_0 = 0.2$ model (B2). We should emphasize, however, that a more detailed comparison between the predictions of the models and observations, taking into account noise and biases of the line identification tech-
iques, is necessary before any model can be considered ruled out.

The high Hubble constant ($h = 0.65$) $\Omega_0 = 0.2$ models C2 and D2, on the other hand, have intrinsically flatter distributions, and a reasonable match between theory and observations can be made by choosing $\gamma$ in the range 1.2–1.7.

9. CONCLUSION

We have systematically developed a set of tools to compute in an efficient manner the column density distribution, given a cosmological model. One fundamental assumption of the approximations involved is that most of the Ly$\alpha$ forest with column densities in the range $10^{12.5}$–$10^{14.5}$ cm$^{-2}$ originates from regions of low overdensity, or even underdensity, that have not undergone orbit crossing. The results of a comparison with a hydrodynamic simulation lend support to our method.

One major conclusion we reach, in the process of developing the tools, is that the peculiar velocities play a minor role in determining the column density distribution at our column densities of interest, even though they are very important in determining the shapes of individual absorption line profiles. We take advantage of this fact and develop a method that we call the density-peak Ansatz, in which each density peak is identified as an absorption line and assigned a column density based upon its local properties. The column density distribution then becomes a statistic of density peaks.

In §§ 6 and 7, we investigate the factors controlling the column density distribution, which can be divided into two categories. One mostly affects the normalization, while the other mostly influences the slope. Those that fall into the former category include the ionizing radiation intensity, the mean temperature of the intergalactic medium, and the mean baryon density. Uncertainties in their values are such that almost any viable cosmological model that has the correct slope of the column density distribution can be made to match observations by a judicious choice of parameters.

The factors that mostly affect the slope of the distribution include the equation of state and, more strongly so, the amplitude and slope of the (linear) power spectrum on scales $1$ Mpc$^{-1} \lesssim k \lesssim 10$ Mpc$^{-1}$. Models that have less power on these scales tend to have comparatively more intermediate-density peaks than high-density ones and hence have relatively steeper column density distributions. Models with a steeper power spectrum on these scales have $k^2/\sigma_0(k)$ that falls off more quickly with increasing $k^{-1} [\sigma_0(k) \propto k^{(n_{\text{eff}} + 3)/2}]$ and therefore a steeper column density distribution, assuming high column density corresponds to high $k^{-1}$ (the Press-Schechter argument; see § 7). Equations of state that are closer to isothermal [smaller $\gamma$, where $\gamma$ satisfies $T \propto (1 + \delta)^{-1}$] tend to produce flatter column density distributions. However, within the reasonable range of $\gamma$ (see Hui & Gnedin 1996), its precise value depending upon the reionization history, the effect of changing the equation of state is small. We put forward an approximate expression relating the slope of the column density distribution to $\gamma$ and the amplitude ($\sigma_0$) and slope ($n_{\text{eff}}$) of the power spectrum on small scales (eq. [41]), which describes reasonably well all the models studied in this paper.

Hence the slope of the column density distribution provides a measure of the amplitude and slope of the power spectrum on small scales for a given cosmological model and given temperature-density relation. We applied our techniques to study a class of CHDM models, which are known to have less power on small scales compared with other popular CDM models. We conclude that the CHDM models indeed produce steeper column density distributions compared with the CDM models. In particular, the low Hubble constant ($h = 0.5$) $\Omega_0 = 0.2$ CHDM models, which have the least amount of power on small scales among the models that we studied, have column density distributions that can be made consistent with observations only for $\gamma$ less than the values we consider reasonable. We emphasize, however, that only after a more detailed comparison between theories and observations, including all the
effects of noise and biases of the line identification methods, can any model be considered ruled out by the observed column density distribution.

We therefore conclude that much work still needs to be done, both on the observational and theoretical fronts. The biases of the line identification techniques used for data reduction deserve close study so that the error bars in the observed column density distributions can be better understood and perhaps reduced. Numerical simulations of the CHDM models should be carried out to further test the accuracy of the approximations made in the present work. Detailed comparisons with hydrodynamic simulations will shed light on the appropriate choice of smoothing scales for approximate methods such as that presented here. The effect of a fluctuating radiation field, instead of a uniform one as is assumed here, has to be investigated. Moreover, in terms of constraining models, it is also important to examine other possible statistics. We have shown, for instance, that the column density distribution is relatively independent of peculiar velocities. Are there other statistics that can take advantage of the different peculiar velocity structures predicted by different cosmological models?

In short, the study of the Ly\textsubscript{x} forest has entered an exciting stage. There is a gold mine of information contained in the quasar absorption spectra waiting to be discovered.

We would like to thank Edmund Bertschinger for his encouragement and many insightful suggestions that greatly improved the paper. We are also grateful to Jeremiah Ostriker for his helpful comments. It is a pleasure to thank Thomas Buchert, Lars Hernquist, Fred Rasio, Urö Seljak, and Matias Zaldarriaga for useful discussions and Esther Hu for kindly sending us her data on the column density distribution. We thank Hongguang Bi and Arthur Davidsen for correspondence regarding their preprint, and we gratefully acknowledge useful comments from Edward Wright. We are indebted to our referee, David Weinberg, for many thoughtful comments and interesting discussions, particularly relating to § 7. This work was supported in part by funds provided by the US Department of Energy under cooperative research agreement DE-FC02-94ER40818 and in part by NSF grant AST 93-18185 and NASA grant NAG 5-2816.

APPENDIX A

SMOOTHING AT JEA\t\t\nscale

The effect of gas pressure is to smooth the baryon density field compared to its dark matter counterpart. The length scale below which this becomes important is the Jeans scale. In linear theory, the baryon overdensity obeys the following equation in a dark matter–dominated universe (Bi et al. 1992; Peebles 1993):

$$\frac{\partial^2 \delta_b}{\partial t^2} + 2H \frac{\partial \delta_b}{\partial t} = 4\pi G \bar{\rho}_{DM} \delta_{DM} - \frac{\gamma k_b T}{m_d} k^2 \delta_b ,$$

(A1)

where the tilde denotes functions in Fourier space as before, $H$ is the Hubble constant, $k_b$ is the Boltzmann constant, $G$ is the Newton constant, $\bar{\rho}_{DM}$ is the average dark matter mass density, $T$ is the average temperature of the gas, and $\mu$ is the mean mass of each gas particle (for a fully ionized gas composed of hydrogen and helium with primordial abundances, it is about 0.6 times the proton mass). The relation between the temperature (not its average but its actual value) and $1 + \delta_b$ is described by $\gamma$, with the temperature being proportional to $(1 + \delta_b)^{-1}$.

The Jeans scale is defined in equation (16). For a dark matter–dominated universe, one can replace $\bar{\rho}_{DM}$ by the total mean density of the universe.

For the special case of $T \propto a^{-1}$, making use of an equation for $\delta_{DM}$ which is the same as equation (A1) except for the absence of the temperature term, it can be shown that

$$\tilde{\delta}_b(k) = \frac{\delta_{DM}(k)}{1 + k^2/k_J^2} ,$$

(A2)

if one ignores decaying modes. It expresses in a quantitative way the expectation that the overdensity in baryons is the same as that of dark matter on large scales (low $k$) but is lower on small scales (high $k$). For $T$ with some other power-law dependence on $a$, solutions for equation (A1) are more complicated but the low- and high-$k$ limits are the same: $\tilde{\delta}_b = \delta_{DM}$ for small $k$ and $\tilde{\delta}_b = \delta_{DM} k_J^2/k^2$ for large $k$ (Bi et al. 1992). We note, however, that in practice, the gas temperature is expected to rise from almost zero before reionization to around $10^4$–$10^5$ K after reionization, and so the power-law time dependence of $T$ is probably not realized. The actual linear smoothing scale of the gas should be a little different from the one given above, and one has to be careful about boundary conditions for $\delta_b$ and its derivative at the onset of reionization (see Gnedin & Hui 1997).

APPENDIX B

THERMAL AND IONIZATION EVOLUTION

The evolution of temperature is governed by

$$\frac{d T}{d t} = -2 HT + \frac{2 T}{3(1 + \delta_b)} \frac{d \delta_b}{d t} - \frac{T}{3} \frac{d}{d t} \sum_i \bar{X}_i + \frac{2}{3k_B n_b} \frac{d Q}{d t} ,$$

(B1)

where $d/dt$ is the Lagrangian derivative following each fluid element, $n_b$ is the proper number density of all gas particles, and $T$ is the temperature which depends on both space and time. The symbol $\bar{X}_i$ is defined by $n_i \equiv (1 + \delta_b) \bar{X}_i \rho_i/m_p$, where $n_i$ is the
proper number density of the species \( i \), \( \tilde{\rho}_i \) is the mean mass density of baryons at the time of interest, \( m_p \) is the mass of the proton, and \( \delta_i \) is the overdensity as in equation (8). The neutral fraction of hydrogen, \( X_{HI} \) (distinct from \( \tilde{X}_{HI} \)) as in equation (9), is then \( \tilde{X}_{HI} / (X_{HI} + \tilde{X}_{HI}) \). Note that \( \tilde{X}_i \) is a function of space and time in general.

The first two terms on the right-hand side of equation (B1) take care of adiabatic cooling or heating. The third accounts for the change of internal energy due to the change in the number of particles. The last term, \( dQ/dt \), is the heat gain (or negative heat loss) per unit volume by the gas particles from the surrounding radiation field. At a redshift of 2 to 4 and for densities of our interest, the main source of heat gain is photoionization and the main source of heat loss is through the recombinantion of ionized hydrogen and the free electron. At higher redshifts, other processes become important, such as Compton cooling. More discussion on these processes will be presented in Hui & Gnedin (1996).

For \( \tilde{X}_i = \tilde{X}_{HI} \), \( \tilde{X}_u \) is the photoionization rate. It is given by

\[
P_h = 2 \int_{v_{HI}}^\infty J_v \, \sigma_{HI} \frac{dv}{h v},
\]

where \( h \) is the Planck constant, \( 2\pi \hbar v_{HI} = 13.6 \) eV, \( \sigma_{HI} \) is the cross section for photoionization as a function of the frequency \( v \), and \( J_v \) is the specific intensity. The photoionization rate \( P_h \) depends on the normalization as well as spectrum of \( J_v \). The specific intensity \( J_v \) is generally taken to have a power-law spectrum, \( v^{-\alpha} \) to \( v^{-\alpha.5} \), for frequencies just above \( v_{HI} \). The spectrum at higher frequencies is less important for the photoionization rate of hydrogen. A convenient way to hide our ignorance of the spectrum is to define \( J_{HI} \) as in equation (18).

For \( \tilde{X}_i = \tilde{X}_{HI} \), \( R \) is the recombinant rate of ionized hydrogen and the free electron (\( \tilde{X}_u = \tilde{X}_u \) and \( \tilde{X}_h = \tilde{X}_{HI} \) in eq. [B2]):

\[
R \sim 4 \times 10^{-13} \frac{T}{10^4 K}^{-0.7} \text{ cm}^3 \text{s}^{-1}.
\]
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press)
Petitjean, P., Mücket, J. P., & Kates, R. E. 1995, A&A, 295, L9
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
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Rees, M. 1986, MNRAS, 218, 25P
Reisenegger, A., & Miralda-Escudé, J. 1995, ApJ, 449, 476
Rugers, M., & Hogan, C. 1996, ApJ, 495, L1
Rybicki, G. B., & Lightman, A. P. 1979, Radiative Processes in Astrophysics (New York: Wiley)
Shandarin, S. F., & Zeldovich, Ya. B. 1989, Rev. Mod. Phys., 61, 185
Songaila, A., & Cowie, L. L. 1996, AJ, 112, 335
Spitzer, L. 1978, Physical Processes in the Interstellar Medium (New York: Wiley)
Tytler, D. 1987, ApJ, 321, 49
---. 1992, in Statistical Challenges in Modern Astronomy, ed. E. D. Feigelson & G. J. Babu (New York: Springer)
Tytler, D., & Burles, S. 1996, preprint (astro-ph/9606110)
Weinberg, D. H., Hernquist, L., Katz, N., & Miralda-Escudé, J. 1996, in Cold Gas at High Redshift, ed. M. N. Bremer, P. P. van der Werf, H. J. A. Röttgering, & C. L. Carilli (Dordrecht: Kluwer), 93
Zeldovich, Ya. B. 1970, A&A, 5, 84
Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57
Zhang, Y., Anninos, P., Norman, M. L., & Meiksin, A. 1997, ApJ, 485, 496