2003

Lyman break galaxies and the Ly alpha forest

JA Kollmeier

DH Weinberg

R Dave

N Katz

Accepted by ApJ

Lyman Break Galaxies and the Lyman-alpha Forest

Juna A. Kollmeier[1], David H. Weinberg[1], Romeel Davé,[2], Neal Katz[3]

ABSTRACT

We use hydrodynamic cosmological simulations to predict correlations between Lyα forest absorption and the galaxy distribution at redshift $z \approx 3$. The probability distribution function (PDF) of Lyα flux decrements shifts systematically towards higher values in the vicinity of galaxies, reflecting the overdense environments in which these galaxies reside. The predicted signal remains strong in spectra smoothed over $50 - 200\,\mathrm{km\ s^{-1}}$, allowing tests with moderate resolution quasar spectra. The strong bias of high redshift galaxies towards high density regions imprints a clear signature on the flux PDF, but the predictions are not sensitive to galaxy baryon mass or star formation rate, and they are similar for galaxies and for dark matter halos. The dependence of the flux PDF on galaxy proximity is sensitive to redshift determination errors, with rms errors of $150 - 300\,\mathrm{km\ s^{-1}}$ substantially weakening the predicted trends. On larger scales, the mean galaxy overdensity in a cube of 5 or $10\,h^{-1}$ Mpc (comoving) is strongly correlated with the mean Lyα flux decrement on a line of sight through the cube center. The slope of the correlation is ~ 3 times steeper for galaxies than for dark matter as a result of galaxy bias. The predicted large scale correlation is in qualitative agreement with recently reported observational results. However, observations also show a drop in the average absorption in the immediate vicinity of galaxies, which our models do not predict even if we allow the galaxies or AGNs within them to be ionizing sources. This decreased absorption could be a signature of galaxy feedback on the surrounding intergalactic medium, perhaps via galactic winds. We find that a simplified "wind" model that eliminates neutral hydrogen in spheres around the galaxies can marginally explain the data.

[1]Ohio State University, Dept. of Astronomy, Columbus, OH 43210, jak,dhw@astronomy.ohio-state.edu

[2]University of Arizona, Dept. of Astronomy, Tucson, AZ 85721, rad@as.arizona.edu

[3]University of Massachusetts, Dept. of Physics and Astronomy, Amherst, MA 91003, nsk@kaka.phast.umass.edu

However, because peculiar velocities allow gas at large distances to produce saturated absorption at the galaxy redshift, these winds (or any other feedback mechanism) must extend to comoving radii of $\sim 1.5\,h^{-1}$ Mpc to reproduce the observations. We also discuss the possibility that extended Lyα emission from the target galaxies "fills in" the expected Lyα forest absorption at small angular separations.

1. Introduction

The strong clustering of Lyman-break galaxies (LBGs) at $z \approx 3$, comparable to that of present-day, optically selected galaxies, suggests that they are highly biased tracers of the underlying dark matter distribution (Adelberger et al. 1998, 2002a). This bias appears to arise naturally in semi-analytic models and hydrodynamic numerical simulations (e.g., Baugh et al. 1998; Governato et al. 1998; Katz, Hernquist, & Weinberg 1999; Kauffman et al. 1999; Cen & Ostriker 2000; Benson et al. 2001; Pearce et al. 2001; Yoshikawa et al. 2001; Weinberg et al. 2002a), which predict that the luminous members of the high redshift galaxy population reside in massive halos, which in turn reside in regions of high background density (Kaiser 1984; Mo & Fukugita 1996; Mo & White 1996). However, the bias of Lyman-break galaxies is inferred by comparing their observed clustering to the predicted clustering of dark matter, which depends on the assumed cosmological model. The Lyα forest offers a tracer of structure whose relation to the underlying dark matter distribution appears to be well understood on theoretical grounds, probing the same redshifts as the LBG population. Correlations between LBGs and Lyα forest absorption therefore offer a natural and potentially powerful probe of the relation between high redshift galaxies and the dark matter distribution, and perhaps for the influence of high redshift galaxies on the surrounding intergalactic medium (IGM).

In this paper we use smoothed particle hydrodynamics (SPH) simulations to predict the correlations between $z = 3$ galaxies and Lyα forest absorption. Hydrodynamic simulations are ideal for this purpose, since they simultaneously predict the locations and properties of the galaxies (e.g. Katz, Hernquist, & Weinberg 1999; Nagamine et al. 2001; Weinberg et al. 2002b) and the structure and ionization state of the intergalactic gas that produces the Lyα forest (e.g. Cen et al. 1994; Zhang, Anninos, & Norman 1995; Hernquist et al. 1996; Theuns et al. 1998). However, the basic expectations can be understood in simple terms, using the "Fluctuating Gunn-Peterson Approximation" (Bi & Davidsen 1997; Croft et al. 1997; Rauch et al. 1997; Weinberg, Katz, & Hernquist 1998; Croft et al. 1998), which provides a fairly accurate description of the Lyα forest results from full hydrodynamic calculations. In this approximation, the Lyα flux decrement at wavelength λ is related to the overdensity at a

redshift $z = \lambda/\lambda_{\alpha,\rm rest} - 1$ along the line of sight by

$$D \;=\; 1 - \frac{F}{F_c} = 1 - e^{-\tau} \;=\; 1 - \exp\left[-A(\rho/\bar{\rho})^{\beta}\right] \;, \tag{1}$$

with

$$A = 0.694 \left(\frac{1+z}{4.0}\right)^{6} \left(\frac{\Omega_b h^2}{0.02}\right)^{2} \left(\frac{T_0}{6000\,\rm K}\right)^{-0.7} \left(\frac{h}{0.65}\right)^{-1} \left(\frac{H(z)/H_0}{5.12}\right)^{-1} \left(\frac{\Gamma}{1.5\times10^{12}\ \rm sec^{-1}}\right)^{-1} . \tag{2}$$

Here $h \equiv H_0/100$ km s^{-1} Mpc^{-1}, Ω_b is the baryon density parameter, and Γ is the HI photoionization rate due to the diffuse UV background at redshift z. This approximation assumes that all gas obeys a temperature-density relation $T = T_0(\rho_b/\bar{\rho}_b)^{\alpha}$, which emerges from the interplay between photoionization heating and adiabatic cooling caused by cosmic expansion. For typical reionization histories, one expects $T_0 \sim 5000 - 20,000\,\rm K$ and $\beta = 2 - 0.7\alpha \approx 1.6 - 1.8$ at $z \sim 3$, with the higher T_0 and β values arising if helium reionization occurs close to this redshift (Hui & Gnedin 1997). Strictly speaking, the overdensity in equation (1) is the gas overdensity, but in the moderate density regions that contribute most of the Lyα forest opacity, the gas pressure is low, and the gas and dark matter trace each other fairly well. This approximation ignores the effects of collisional ionization and peculiar velocities.

From equation (1), one can see that the probability distribution function (PDF) of Lyα forest flux decrements, $P(D)$, is closely related to the PDF of the underlying density field, $P(\rho/\bar{\rho})$, with an effective smoothing scale determined by a combination of pressure support and thermal broadening along the line of sight. The flux decrement PDF can therefore provide a diagnostic for the amplitude of mass fluctuations and for departures from Gaussianity in primordial fluctuations (Weinberg et al. 1999a; Croft et al. 1999; McDonald et al. 2000).

The statistical measure that we focus on in this paper is the *conditional* flux decrement PDF, the probability $P(D)$ that a pixel in a Lyα forest spectrum has a flux decrement in the range $D \rightarrow D + dD$, conditioned on the presence of a galaxy at the same redshift close to the line of sight. (Since $D = 1 - F/F_c$, we will use the slightly less cumbersome term "conditional flux PDF," but we treat D as the observable of interest because it is an increasing function of density.) This statistic can be measured by obtaining Lyα forest spectra in fields covered by LBG surveys. The bias of galaxies towards high density regions should reveal itself as a systematic shift of $P(D)$ towards higher flux decrements in pixels close to galaxies, with closer proximity yielding stronger shifts. Conversely, feedback of galaxies on the local IGM via ionization or galactic winds might shift the PDF towards lower flux decrements at small separations.

A complete measurement of the conditional flux PDF will require ambitious, time-consuming surveys on large telescopes, since each LBG-quasar pair contributes only a single point to the distribution. However, characteristics of the conditional PDF like the mean decrement or the fraction of saturated and "transparent" pixels can be measured with fewer data points, and we present predictions for these quantities in addition to the full PDF. We also show that the signature of galaxy proximity should persist in spectra smoothed over $50 - 200\,\mathrm{km\ s^{-1}}$, allowing theoretical predictions to be tested using relatively low resolution quasar spectra. With higher resolution spectra the dependence of the conditional PDF on spectral smoothing offers a further test of the models.

K. Adelberger and collaborators have been carrying out a survey along the lines envisioned here (Adelberger et al. 2002a, hereafter ASSP; Adelberger et al. 2002b, hereafter ASPS). While we have not designed our study primarily for comparison with their results, we will discuss the comparison between our predictions and some of their measurements in §5 below. Inspired by their work, we also consider an additional statistic, the mean galaxy overdensity in $5\,h^{-1}$ and $10\,h^{-1}$ Mpc (comoving) cubes as a function of the mean decrement on a line of sight through the middle of the cube. This statistic characterizes the large scale correlations between galaxies and IGM overdensity, while the conditional flux PDF characterizes correlations on smaller scales. McDonald, Miralda-Escudé, & Cen (2002) have presented predictions for this statistic using dark matter in N-body simulations, and here we show predictions for the biased galaxy population of a hydrodynamic simulation.

There is substantial overlap between our investigation and the recent paper of Croft et al. (2002a). As discussed in §5, our results generally agree well with theirs where we examine similar quantities. Croft et al. (2002a) devote considerably more attention to models that incorporate galactic winds, which they add to their simulations using the post-processing approach of Aguirre et al. (2001). Here we concentrate on the direct predictions of our simulations, which incorporate thermal feedback in the local ISM that is usually radiated away before it can drive a galactic wind. However, we do estimate the possible effects of galaxy photoionization, and of photoionization by recurrent AGN activity associated with galaxies, and we briefly discuss the possible influence of stronger winds. We describe our simulations, galaxy identification, and spectral extraction in the next section. We present our results for the conditional flux PDF in §3 and for the correlation between mean decrement and galaxy counts in cubic cells in §4. In §5, we briefly review our results, then discuss them in relation to the theoretical study of Croft et al. (2002a) and the observational analyses of ASSP, concluding with some remarks about future directions.

2. Method

Our results are derived from SPH simulations of an inflationary cold dark matter model with a cosmological constant (ΛCDM), with cosmological parameters $\Omega_m = 0.4$, $\Omega_\Lambda = 0.6$, $h = 0.65$, $\Omega_b = 0.02h^{-2} = 0.0473$, an inflationary spectral index $n = 0.95$, and a power spectrum normalization $\sigma_8 = 0.80$. These simulations use a parallel implementation of TreeSPH (Hernquist & Katz 1989; Katz, Weinberg & Hernquist 1996, hereafter KWH; Davé et al. 1997), incorporating radiative cooling and star formation as described by KWH. We rely mainly on two simulations, one that uses 128^3 SPH particles and 128^3 dark matter particles in a periodic cube $22.222\,h^{-1}$ Mpc (comoving) on a side, and one that uses 2×144^3 particles in a cube $50\,h^{-1}$ Mpc on a side. We use two additional simulations to investigate numerical resolution effects. We refer to simulations with the nomenclature LXnY, where the box size is $\mathrm{X}\,h^{-1}$ Mpc and the particle number is Y^3 for each species. The main simulations are thus L22n128 and L50n144, and the additional simulations are L11n64 and L11n128. The properties of all of these simulations are summarized in Table 1. L22n128 and L11n64 have the same resolution as simulations that we have previously used to study the Lyα forest (Hernquist et al. 1996; Davé et al. 1999), while L11n128 has a mass resolution a factor of eight higher. The L50n144 simulation was designed primarily for studies of galaxy assembly, galaxy clustering, and the low redshift intergalactic medium (Davé et al. 2001, 2002; Croft et al. 2001; Chen et al. 2002; Fardal et al. 2002; Murali et al. 2002; Weinberg et al. 2002a). It has a substantially larger volume than our other simulations, but its mass resolution is a factor of eight lower than that of L22n128. We analyze all of the simulations at redshift $z = 3$, close to the mean redshift of the main LBG samples.

With the exception of L50n144, all of the simulations were evolved with the photoionizing UV background calculated by Haardt & Madau (1996). At the processing stage, the intensity of the background is adjusted by small factors ($0.75 - 1.16$, depending on the simulation) so that the mean decrement of the Lyα forest spectra extracted from the simulation matches the value $D = 0.36$ found by Press, Rybicki, & Schneider (1993) at $z = 3$ (Bernardi et al. 2003 find a nearly identical value with a much larger quasar sample). The L50n144 simulation was run with no ionizing background (for the reasons discussed by Weinberg, Hernquist & Katz 1997), so its unshocked gas cools adiabatically to unrealistically low temperatures. We therefore modify the temperatures of all gas particles that have $(\rho/\bar{\rho}) < 100$ and $T < 10^4(\rho/\bar{\rho})^{0.6}\,\mathrm{K}$, setting them to $T = 10^4(\rho/\bar{\rho})^{0.6}\,\mathrm{K}$. We do not modify the temperatures of hotter particles. In this way, we mimic the results expected if we had used an ionizing background during evolution, and the tests in §3.1 below indicate that this approach is adequate for our purposes. We again extract spectra with an ionizing background intensity chosen to reproduce the Press, Rybicki, & Schneider (1993) mean flux decrement.

The star formation algorithm is discussed in detail by KWH. In brief, cold gas with physical density $n_H > 0.1\,\mathrm{cm}^{-3}$ and overdensity $\rho/\bar{\rho} > 1000$ is converted into stars on a timescale determined by the local dynamical and cooling times. Feedback is deposited as thermal energy in the surrounding medium, and it is usually radiated away before the interstellar medium becomes hot enough to drive a "galactic wind." The SKID algorithm[1] (Spline-Kernel Interpolated DENMAX; see KWH and Gelb & Bertschinger 1994) is used to identify clumps of stars and cold gas ($T < 30,000\,\mathrm{K}$, $\rho/\bar{\rho} > 1000$) that correspond to "galaxies." Because these systems are extremely overdense, there is no ambiguity in their identification, and their masses are robust to changes in the details of the radiative cooling or star formation treatment (KWH; Weinberg, Hernquist & Katz 1997). We restrict our attention to galaxies with baryonic mass (stars plus cold gas) $M_b \geq 64 m_{\mathrm{SPH}}$, since our comparisons among simulations of different resolution indicate that this population is insensitive to resolution effects.

The simulations produce a population of high redshift galaxies whose clustering, star formation properties, and sub-mm luminosities are in approximate agreement with observations (Katz, Hernquist, & Weinberg 1999; Weinberg et al. 1999b, 2002a,b; Fardal et al. 2002), though the uncertainties in some of these comparisons are substantial. Correlations with the Lyα forest offer the possibility of testing the simulation predictions for the relation between these galaxies and the underlying mass distribution. The simulated galaxies span a wide dynamic range in baryonic mass and star formation rate (SFR), and we want to know whether the predicted correlations with the Lyα forest depend substantially on galaxy properties. Rather than model the galaxy spectral energy distributions in detail, we define galaxy samples above thresholds in baryonic mass or SFR and characterize them by their space density,

$$\Sigma = \frac{N_{\mathrm{gal}}}{\theta^2_{\mathrm{box}} \times (\Delta z)_{\mathrm{box}}} \ , \tag{3}$$

where N_{gal} is the number of galaxies above the threshold in the simulation box, θ_{box} is the projected size of the box in arc-minutes, and $(\Delta z)_{\mathrm{box}}$ is the redshift depth of the box. The units of Σ are thus number per square arc-minute per unit redshift (which we will write $\mathrm{arcmin}^{-2}\,\Delta z^{-1}$), and it is worth emphasizing that "per unit redshift" makes Σ a space density rather than a surface density. To the extent that galaxies' rest-frame UV luminosities are monotonically related to their instantaneous star formation rates, samples selected above an SFR threshold should correspond to observed LBG samples above an optical magnitude limit. Typical LBG spectroscopic samples to $R \sim 25.5$ have $\Sigma \sim 1-3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ (Steidel et al. 1996; Adelberger et al. 1998).

[1] `http://www-hpcc.astro.washington.edu/tools/skid.html`

Lyα forest spectra are extracted using TIPSY[2], as described by Hernquist et al. (1996). Briefly, the optical depth at each position is calculated from the neutral hydrogen mass, neutral mass-weighted velocity, and neutral mass-weighted temperature, derived by integrating through all of the SPH smoothing kernels that overlap that spatial position. We extract 1200 spectra along random lines of sight (400 in each projection) for each simulation. We use "pixels" of $1.5 - 3\,\mathrm{km\ s^{-1}}$, small enough that variations on the thermal broadening scale $\sim 10\,\mathrm{km\ s^{-1}}$ are fully resolved. As discussed below, we compute statistics after smoothing these spectra with Gaussians of 1σ width $10 - 200\,\mathrm{km\ s^{-1}}$, with the smallest of these scales representing the approximate spectral resolution of typical high resolution QSO spectra. We do not add noise to our artificial spectra; in general, noise will act to broaden measured distribution functions but will average out in measurements of mean decrements. To compute the conditional flux PDF statistics discussed in the next section, we search around each sightline for nearby galaxies in the cylindrical volume defined by a given search radius and the depth of the box. This method yields few matches at small angular separations or for rare galaxy classes, so we supplement the random spectra with spectra extracted close to the known positions of the galaxies, with the distribution of projected separations that would be expected for random positioning.

3. The Conditional Flux PDF

We begin by examining the physical conditions of the IGM in the vicinity of galaxies. Figure 1, based on the L22n128 simulation, shows the median value, interquartile range, and 10–90% range of the temperature, gas overdensity, neutral hydrogen density, and neutral hydrogen fraction in pixels that lie at the same redshift as a simulated galaxy at transverse separation r_t (in comoving h^{-1} Mpc). For our adopted cosmological parameters, an angular separation of $1'$ at $z = 3$ corresponds to $r_t = 1.19\,h^{-1}$ Mpc (comoving), and to $Hr_t = 152\,\mathrm{km\ s^{-1}}$. As expected, the gas density increases close to galaxies. Typical temperatures are also higher in the vicinity of galaxies. In the median, this temperature increase is mostly accounted for by the increased photoionization heating at higher densities, but the spread to high temperatures is probably an effect of shock heating as gas falls into collapsed regions. In our simulations, supernova heating should have little impact at the separations shown in Figure 1. While higher temperatures counteract the influence of higher densities, the density effect is stronger, and typical *neutral* hydrogen densities increase sharply in the neighborhood of galaxies. The median neutral fraction is roughly constant at $\sim 10^{-5}$, increasing at $r_t \leq 100\,h^{-1}$ kpc.

[2]`http://www-hpcc.astro.washington.edu/tools/tipsy/tipsy.html`

Figure 1 motivates our use of the statistic described in the introduction, the conditional PDF of the flux decrement D in pixels within angular separation $\theta_{\rm max}$ of a galaxy at the same redshift. We will investigate the dependence of this statistic on the angular separation threshold $\theta_{\rm max}$, on the smoothing of the Lyα forest spectrum, on the velocity separation between the pixel redshift and the galaxy redshift, and on the properties of the galaxies themselves. In addition to the complete conditional PDF, we calculate three quantities that summarize some of its properties: the (conditional) mean decrement $\langle D\rangle = \int_0^1 DP(D)dD$, the fraction $f_{\rm sat}$ of "saturated" pixels with $D > 0.9$, and the fraction $f_{\rm trans}$ of "transparent" pixels with $D < 0.55$. Before presenting our main results from the L22n128 and L50n144 simulations, we first test the sensitivity of the predictions to numerical parameters of the simulations.

In general, one can consider the dependence of the flux PDF on both the angular separation $\Delta\theta$ between the galaxy and the QSO sightline and the velocity separation ΔV between the galaxy redshift and the pixel redshift. For the most part, we concentrate on the $\Delta V = 0$ slice of this more general distribution $P(D|\Delta\theta, \Delta V)$, so for each galaxy-sightline pair we consider the single pixel closest to the galaxy redshift, incorporating the effects of peculiar velocity on both the galaxy redshift and the Lyα forest spectrum. Alternatively, one can define the conditional flux PDF in bins of redshift-space separation $\Delta_r = [(d_A\Delta\theta)^2 + (\Delta V/H)^2]^{1/2}$, where d_A is the angular diameter distance and H the Hubble parameter at redshift z. There are three disadvantages to the latter approach: it requires a cosmological model to compute d_A and $H(z)$ (which do not scale the same way with cosmological parameters), peculiar velocities distort the relation between ΔV and distance, and the statistical properties of the measured distributions are more complex because each galaxy-quasar pair contributes correlated values to $P(D)$ in different bins of Δ_r. However, the Δ_r definition has a significant advantage when computing the conditional mean decrement from a small sample of moderate S/N spectra, namely that each spectrum contributes many pixels and the noise averages out. Our definition requires either high S/N per pixel or a large number of galaxy-QSO pairs to obtain a precise measurement of the conditional mean decrement. In practice, we recommend smoothing spectra enough to ensure high S/N per pixel, since we show below that smoothing over $\Delta V \sim 50 - 100\,{\rm km\ s^{-1}}$ does not reduce the discriminating power of the measurements. This way explicit smoothing achieves the noise suppression that is obtained by binning in the Δ_r approach. In §5, we will calculate the mean decrement in bins of Δ_r for comparison to Croft et al.'s (2002a) predictions and ASSP's measurements.

3.1. Dependence on Numerical Parameters

We investigate the influence of mass resolution using the L11n128 and L11n64 simulations, which have the same initial conditions and box size but a factor of eight difference in particle mass. Figure 2 presents the conditional flux PDFs of these simulations, in a format similar to that used in many of our subsequent figures. The left hand panel shows results for spectra smoothed with a Gaussian of 1σ width $R_s = 10\,\mathrm{km\ s^{-1}}$, and the right hand panel for $R_s = 50\,\mathrm{km\ s^{-1}}$. Solid curves show the unconditional PDF, computed from all pixels in all of the randomly positioned spectra. Dotted curves show the conditional PDF for pixels within $2'$ of galaxies at the same redshift. The physical mass threshold for these galaxies is the same in the two simulations, corresponding to $64m_{\rm SPH}$ in L11n64 and to $512m_{\rm SPH}$ in L11n128. Dashed curves show the conditional flux PDF for angular separations $\Delta\theta < 0.5'$. Symbols along the top show the mean decrements computed from these three PDFs. Symbols along the right and left hand vertical axes denote the fractions of saturated and transparent pixels, respectively.

We will discuss the physics behind these conditional PDFs in §3.2 below, though the basic behavior is clearly what one would expect based upon the general discussion in §1, given the trends in Figure 1. The crucial point of Figure 2 is the similarity between results represented by the bold lines/symbols and those represented by the light lines/symbols, which come from the L11n128 and L11n64 simulations, respectively. While there are systematic differences, most notably in the number of transparent pixels for $\Delta\theta < 0.5'$, these differences are quite small, especially compared to the much stronger dependence on the separation itself. Figure 2 implies that the L22n128 simulation, which has the same mass resolution as L11n64 and a larger volume, should yield reliable predictions for galaxies above its $64m_{\rm SPH}$ resolution threshold.

Comparing the L11n64 and L22n128 simulations allows us to test the influence of box size on the conditional flux PDF. The simulations cannot include density perturbations on scales larger than the fundamental mode of the box, and this absence of large scale fluctuations could alter the statistical properties of the Lyα forest or the bias of galaxies with respect to the mass. We computed conditional flux PDFs for galaxy samples (selected on SFR) with a space density of $\Sigma = 3\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$ in the two simulations. We found only small differences in the results (and therefore do not plot them), with a slight shift of the distribution in the smaller box towards an increased number of low decrement pixels and a corresponding dearth of high decrement pixels. The conditional mean decrement is thus slightly depressed in the smaller box, but again this impact is much smaller than the impact of galaxy proximity itself. We conclude that the lack of large scale modes in the L11n64 simulation has only a minimal impact on the conditional flux PDF for galaxies of

this space density, and that the box size of the L22n128 simulation should be adequate for our purposes. The bias of rarer, more massive (or higher SFR) galaxies might be more sensitive to the absence of long wavelength modes, but for rare galaxies one needs a large simulation volume to get adequate statistics in any case.

Figure 3 compares the results of the L50n144 and L22n128 simulations, in the same format as Figure 2. We select galaxies above the same physical mass threshold, which corresponds to $64m_{\rm SPH}$ and $512m_{\rm SPH}$, respectively. The space densities are $\Sigma \approx 3\,{\rm arcmin}^{-2}\,\Delta z^{-1}$. In this comparison, differences could arise from random differences in the structure present, or from systematic effects of box size, mass resolution (which is another factor of eight lower in L50n144), or our *post-facto* imposition of a temperature-density relation in L50n144 (see §2). The good agreement in Figure 3 shows that the combined impact of the systematic effects must be reasonably small. In what follows, therefore, we will use L22n128 to make predictions for abundant galaxy populations with $\Sigma > 3\,{\rm arcmin}^{-2}\,\Delta z^{-1}$ and L50n144 to make predictions for rarer populations, where its larger volume is needed to provide adequate statistics. The differences in Figure 3 are probably an upper limit on the level of systematic uncertainty associated with the numerical parameters of the simulations. Even the noticeable difference in the mean decrement for $\Delta\theta < 2'$ is not highly significant relative to the statistical uncertainties, since there are only 37 galaxies above the mass threshold in the L22n128 simulation. We conclude that the trends predicted by our simulations are robust to the effects of finite resolution and box size and that our quantitative results are only mildly sensitive to these numerical effects.

3.2. Predictions of the Conditional Flux PDF

Figure 4 presents the conditional flux PDF for the full population of resolved galaxies ($M_b \geq 64m_{\rm SPH} = 6.7 \times 10^9 M_\odot$) in the L22n128 simulation. There are 641 galaxies in the box at $z = 3$, and the space density of this population is $\Sigma = 48\,{\rm arcmin}^{-2}\,\Delta z^{-1}$. (As indicated in the note to Table 1, the box is $18.74'$ on a side and has a redshift depth $\Delta z = 0.0379 = 2844\,{\rm km\ s}^{-1}$.) The upper left panel is for spectra smoothed over $10\,{\rm km\ s}^{-1}$, comparable to the spectral resolution of typical Keck HIRES or VLT UVES spectra. The unconditional flux PDF, shown by the solid curve, reflects the probability distribution of gas overdensity $P(\rho/\bar{\rho})$, which depends mainly on the amplitude of matter fluctuations at $z = 3$ and on the assumption that the primordial fluctuations were Gaussian (see, e.g., the discussions by Miralda-Escudé et al. 1996, 1997; Cen 1997; Rauch et al. 1997; Weinberg et al. 1999a; Croft et al. 1999; McDonald et al. 2000). The shape of $P(D)$ is, of course, strongly influenced by the non-linear transformation between density and flux decrement (eq. 1); for

example, the high and low density tails of $P(\rho/\bar{\rho})$ map to flux decrements $D \approx 1$ and $D \approx 0$, respectively, so $P(D)$ peaks at these extremes.

Dot-dashed, dotted, and dashed curves show $P(D)$ for pixels that have a galaxy at the same redshift within $\theta_{\rm max} = 2'$, $1'$, and $0.5'$, respectively (comoving separations of 2.4, 1.2, and $0.6\,h^{-1}$ Mpc). The preferential location of galaxies in overdense regions reveals itself through shifts of the distributions towards higher flux decrements, with corresponding increases in the mean decrement (symbols along the top axis) and the fraction of saturated pixels (symbols along the right axis), and decreases in the fraction of transparent pixels (symbols along the left axis). These shifts become progressively stronger as the proximity to galaxies increases ($\theta_{\rm max}$ decreases). Pixels that contribute to the $\Delta\theta < 0.5'$ histogram also contribute to the $\Delta\theta < 2'$ histogram, but the number of galaxy–spectrum separations is proportional to $(\Delta\theta)^2$, so the conditional PDF for a given $\theta_{\rm max}$ is contributed largely by separations close to $\theta_{\rm max}$. A sphere of radius $2.4\,h^{-1}$ Mpc ($\Delta\theta = 2'$) contains 3.4 galaxies on average, so for the values of $\theta_{\rm max}$ in Figure 4 the galaxy "matched" to any given pixel is not necessarily the one with the smallest 3-d physical separation. For the brighter, lower space density samples defined later, the matched galaxy is usually also the closest one.

The remaining panels of Figure 4 show the unconditional and conditional flux PDFs for spectra smoothed over 50, 100, and 200 km s^{-1}. Smoothing brings the flux decrements in extreme pixels closer to the mean, so for larger R_s (which represents the 1σ width of the Gaussian smoothing filter) the fractions of saturated and transparent pixels drop. With $R_s = 200$ km s^{-1} the unconditional PDF peaks close to the mean decrement, though it remains fairly broad. Roughly speaking, the flux PDF of smoothed spectra reflects the PDF of density smoothed on a similar physical scale, but the smoothing of the spectra is 1-d rather than 3-d, and smoothing does not commute with the non-linear transformation between density and flux.

For our purposes, the important feature of these panels is that the difference between the unconditional and conditional flux PDFs, and between the conditional PDFs for different $\theta_{\rm max}$, persists for smoothing lengths up to 200 km s^{-1}, and is perhaps even enhanced relative to the minimal smoothing case. Since we see a strong signature of galaxy proximity at $\Delta\theta = 2' = 304$ km s^{-1}, it is no surprise that the signature persists when spectra are smoothed over this physical scale; the difference between the unconditional and conditional PDFs reflects the large scale environments in which galaxies reside. Figure 4 shows that the simulation predictions can be tested with moderate-to-low resolution spectra, thus allowing the use of relatively faint AGN as targets for observational studies.

The principal difficulty with lower resolution spectra is the greater difficulty of continuum fitting, which causes systematic underestimation of flux decrements, especially at higher

redshifts where large expanses of unabsorbed continuum are rare. However, the qualitative trends with galaxy proximity should persist even in the presence of continuum fitting errors. With high resolution spectra, the dependence of the conditional flux PDF (or properties like $\langle D \rangle$, $f_{\rm sat}$, and $f_{\rm trans}$) on smoothing scale can be used to test theoretical predictions, in addition to the dependence on galaxy proximity itself. Note that the mean of the unconditional flux PDF does not depend on R_s, since smoothing preserves flux overall, but the mean decrement for conditional PDFs decreases with increasing R_s, since pixels close to galaxies tend to have high decrements that are reduced by smoothing.

Another indicator of the physical scale of correlations is the dependence of the conditional PDF on the velocity separation between the pixel redshift and the galaxy redshift. In Figure 5, long-dashed lines and solid symbols repeat results for the case investigated in Figures 2–4, velocity separation $\Delta V = 0$. Upper and lower panels show $\theta_{\rm max} = 2'$ and $0.5'$, and left and right columns show smoothing lengths $R_s = 10$ and $50\,{\rm km\ s^{-1}}$. Dashed, dotted, and dot-dashed lines show the conditional PDFs computed from pixels that are offset by $|\Delta V| = 100$, 200, and $400\,{\rm km\ s^{-1}}$ from the galaxy redshift, where we have averaged the contributions at $\pm\Delta V$. For $|\Delta V| = 100$ or $200\,{\rm km\ s^{-1}}$, the signature of galaxy proximity on the flux PDF persists, but it is significantly weakened. By $|\Delta V| = 400\,{\rm km\ s^{-1}}$, it has mostly vanished. The dependence on velocity separation provides another potential test of the simulation predictions. Perhaps more significantly, Figure 5 shows that redshift measurement errors of $\sim 100 - 400\,{\rm km\ s^{-1}}$ would significantly affect the shape of the conditional flux PDF. We will return to this point in §3.6 below.

The unconditional flux PDF is weighted by volume, since every volume element has an equal chance of becoming a pixel in a Lyα forest spectrum. Therefore, even if galaxies traced the underlying mass distribution, we would expect the conditional flux PDF to differ from the unconditional PDF because the mass itself is preferentially located in overdense regions. Figure 6 demonstrates, however, that the strong dependence of the flux PDF on galaxy proximity is primarily a result of galaxy bias, not the difference between volume and mass weighting. Light curves, repeated from the corresponding panels of Figure 4, show the unconditional PDFs and the conditional PDFs for galaxy separations $\Delta\theta < 2'$ and $\Delta\theta < 0.5'$. Heavy curves show conditional PDFs in which we use randomly selected dark matter particles in place of galaxies. While these curves are shifted relative to the unconditional PDFs shown by the solid lines, they are not shifted nearly as much as the galaxy conditional PDFs. High-redshift galaxies form in special locations of the large scale matter distribution, and the signature of these special environments is imprinted on the nearby Lyα forest.

3.3. Dependence on Galaxy Properties

In the local universe, galaxies with different stellar populations, star formation rates, morphologies, and masses tend to populate different environments, and the same may be true of high redshift galaxies. Theoretical models predict some dependence, albeit rather weak, of the LBG autocorrelation function on galaxy mass or SFR (e.g., Mo & Fukugita 1996; Katz, Hernquist, & Weinberg 1999), and observations provide at least tentative evidence for stronger clustering of brighter LBGs (e.g., Adelberger et al. 1998; Giavalisco & Dickinson 2001). The dependence of the conditional flux PDF on the properties of the conditioning galaxies is a potential diagnostic for the dependence of LBG properties on large scale environment. Here we will focus on instantaneous SFR for defining different galaxy subsamples, since it is the property most closely tied to rest-frame UV luminosity. In the simulated galaxy population, these SFRs are correlated with galaxy baryonic mass (Weinberg et al. 2002b), so this procedure is approximately, but only approximately, equivalent to selecting on baryonic mass.

In Figure 7, heavy solid lines show the conditional flux PDFs for the full galaxy population of the L22n128 simulation, as in Figure 4, for $\Delta\theta < 2'$ (top) and $\Delta\theta < 0.5'$ (bottom), with spectral smoothing $R_s = 10\,\mathrm{km\ s^{-1}}$ (left) or $50\,\mathrm{km\ s^{-1}}$ (right). Light solid lines show the corresponding results for the full population of resolved galaxies ($M_b > 64 m_{\mathrm{SPH}} = 5.4 \times 10^{10} M_\odot$) in L50n144. The space density of this more massive galaxy population is $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$, with 450 galaxies in the simulation volume. Dotted lines show results for the 150 galaxies with the highest SFRs, and dashed lines show results for the top 50 galaxies. The conditional flux PDF is remarkably insensitive to galaxy SFR over a range of nearly 150 in space density, though for $\Delta\theta < 0.5'$ the more massive, high SFR galaxies do show slightly higher mean decrement and saturated fraction and slightly lower transparent fraction, as expected if they reside in denser environments.

One of the simplest and most widely used models for the spatial clustering of LBGs associates each LBG with a single dark matter halo (e.g., Mo & Fukugita 1996; Adelberger et al. 1998; Moscardini et al. 1998; Bagla 1998). The large scale bias factor of these halos can be computed using the approximations of Mo & White (1996) and Sheth, Mo, & Tormen (2001). Katz, Hernquist, & Weinberg (1999) show that the correlation function of $z = 3$ galaxies in SPH simulations like these is similar to that of dark matter halos of the same space density, even though the galaxy-halo association for a given space density is not truly 1-to-1. Figure 8 presents an analogous comparison for the conditional flux PDF. Light lines and symbols show results for the 450 resolved galaxies in the L50n144 simulation, while heavy lines and symbols show the results when these galaxies are replaced by the 450 most massive dark matter halos in the simulation (identified by a friends-of-friends algorithm with

linking parameter of 0.2). The conditional PDFs for galaxies and halos are very similar, with galaxies showing a slightly stronger preference for regions of high flux decrement, suggesting that they are more likely to occur in halos with higher density environments (where the halos themselves tend to be more massive). This similarity holds for other choices of angular separation, spectral smoothing scale, and population space density, not shown in Figure 8.

A more general model for galaxy bias is the "halo occupation distribution" (HOD) formalism, which characterizes the mean number of galaxies of a given type in halos of virial mass M by an arbitrary function $N_{\rm avg}(M)$. This approach has been widely applied to studies of low redshift clustering (e.g., Jing, Mo, & Börner 1998; Seljak 2000; Ma & Fry 2000; Peacock & Smith 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Cooray 2002; Cooray & Sheth 2002; Marinoni & Hudson 2002; Scranton 2002; Zehavi et al. 2003), and Bullock et al. (2001) and Moustakas & Somerville (2002) have analyzed LBG clustering in this context. If we neglect the finite size of the halos themselves, then the influence of different $N_{\rm avg}(M)$ choices on the conditional flux PDF can be calculated by simply weighting the contribution of each halo by $N_{\rm avg}(M)$. Figure 9 shows the conditional flux PDF based on randomly selected dark matter particles (solid line, as in Figure 6) and on the top 450 halos of the L50n144 simulation, with contributions weighted by $W \propto M^0$ (one galaxy per halo), $W \propto M$ (galaxy number proportional to halo mass), and additional cases $W \propto M^{0.5}$ and $W \propto M^{1.5}$. While the bias of halos with respect to dark matter remains clearly evident, the conditional flux PDF is almost completely independent of the relative weighting of the halos, and, by implication, of the galaxy HOD over the rather wide parameter range considered here. Similar results hold for other angular separations and spectral smoothings. Presumably the more massive halos do tend to reside in denser environments, but saturation removes the sensitivity of the flux decrement to the neutral hydrogen density once it becomes sufficiently high. Thus, the sign of the effect, more absorption when more massive halos are weighted more strongly, is in the expected direction, but the magnitude is tiny.

At one level, the results in Figures 7–9 are disappointing, since they suggest that the conditional flux PDF will not help to distinguish among competing scenarios for the origin of LBG properties. However, the robustness of the predictions to the details of the galaxy formation assumptions makes the conditional flux PDF a good diagnostic for the influence of galaxies on the IGM. Since a wide range of reasonable galaxy bias models make similar predictions, a substantial departure from these predictions would most likely arise from galaxies directly affecting the Lyα forest through photoionization, winds, or other processes. Furthermore, differences in the conditional flux PDF for different galaxy classes — based on luminosity, stellar populations, or emission line profiles, for example — would probably reflect their differing effects on the IGM, not their different clustering properties.

3.4. The Conditional Mean Flux Decrement

The dependence of the mean flux decrement on the galaxy proximity condition provides a convenient summary of our results. The mean decrement does not capture the full information in the PDF; we have tried, for example, scaling the optical depths of the unconditional flux PDF by a constant factor to see whether we can recover the conditional flux PDF simply by matching its mean, and we cannot. However, measurement of the full conditional flux PDF requires many galaxy-quasar pairs, so measurements of the conditional mean decrement are likely to come much sooner (e.g., ASSP).

Figure 10 plots the mean flux decrement as a function of maximum angular separation $\theta_{\rm max}$ for galaxy populations of different space density, summarizing and reinforcing the results in Figures 4 and 7. The spectral smoothing length is $R_s = 10\,{\rm km\ s^{-1}}$. Hexagons and pentagons show results for the full populations of resolved galaxies in the L22n128 and L50n144 simulations, with space densities $\Sigma = 48\,{\rm arcmin^{-2}}\,\Delta z^{-1}$ and $3\,{\rm arcmin^{-2}}\,\Delta z^{-1}$, respectively. Squares and triangles show results for the 150 and 50 "brightest" (i.e., highest SFR) galaxies in L50n144. For $\Delta\theta < \theta_{\rm max} = 4'$ ($4.8\,h^{-1}$ Mpc comoving) the mean decrement is already well above the global mean of 0.36, and it rises steadily with decreasing $\theta_{\rm max}$ as the Lyα spectra probe the denser IGM in the immediate vicinity of the galaxies, with $\langle D\rangle > 0.9$ for $\theta_{\rm max} = 0.25'$. The mean decrement varies only weakly with SFR, with the rare, high SFR galaxies having slightly higher mean decrement. The $\times$'s show, for the full population of L50n144, the conditional mean decrement computed in bins of angular separation $0.8\theta_{\rm max} < \Delta\theta < 1.25\theta_{\rm max}$, instead of all separations $\Delta\theta < \theta_{\rm max}$. This "differential" form of the statistic shows a similar trend, but the increase of $\langle D\rangle$ at a given $\theta_{\rm max}$ is, of course, smaller, since closer separation pairs have been eliminated.

Figure 11 plots the mean decrement against $\theta_{\rm max}$ for random dark matter particles and for halos weighted in various ways, as in Figures 6 and 9. The effect of bias stands out clearly as an offset between the matter and halo results. Results for halos are similar to those for galaxies in Figure 10, though the mean decrements near halos are typically slightly lower. The weighting of halos has only a small impact on the conditional mean decrement, with a modest increase in $\langle D\rangle$ for a stronger weighting of high mass halos.

3.5. Local Photoionization

We have so far assumed that the photoionizing background is uniform, and Croft et al. (2002b) show that this assumption should be adequate for most computations of Lyα forest statistics. However, the conditional flux PDF is derived from the small fraction of pixels

that lie close to galaxies, and if the galaxies make a significant contribution to the ionizing background, then the flux decrement in these pixels may be depressed by the ionizing flux of their nearest galaxy neighbors. With the ionizing background predicted by Haardt & Madau (1996) based on the observed quasar population, these simulations (and similar ones by other groups) already match the observed mean decrement of the Lyα forest given a baryon density $\Omega_b = 0.02h^{-2}$. However, estimates of the background based on the proximity effect usually yield a higher intensity than predicted from the quasar population alone (Scott et al. 2000, and references therein), and observations of LBGs suggest that the escape fraction of ionizing photons is high enough to make them an important contributor to the UV background (Steidel, Pettini, & Adelberger 2001). Within the theoretical and observational uncertainties, there is probably room for galaxies to make a contribution equal to that of quasars without making the predicted mean flux decrement too low at $z \sim 3$ (see Schirber & Bullock 2003 for a recent discussion of these issues).

The intensity of the radiation from an individual galaxy or quasar falls off as $1/r^2$, but it is further attenuated by redshifting and, more importantly for the case of ionizing radiation, by IGM absorption. Haardt & Madau (1996; see also Madau, Haardt, & Rees 1999) estimate that a path length $\Delta z \sim 0.17$ produces an optical depth of $\tau(912\text{Å}) \sim 1$ at $z = 3$, corresponding to $100\,h^{-1}$ Mpc (comoving) for our cosmology. Fardal & Shull (1993) conclude that the effective attenuation length for "average" ionizing photons is ~ 2.4 times that at 912Å, so a reasonable estimate (with probably a factor of two uncertainty) is $r_{\rm att} \sim 240\,h^{-1}$ Mpc (comoving). Here we will make the simple approximation that intensity falls as $1/r^2$ until $r_{\rm max} = r_{\rm att}$ and is sharply truncated beyond $r_{\rm max}$. If a population of galaxies with comoving space density n and mean ionizing luminosity $\langle L \rangle$ contributes a fraction f of the ionizing background intensity I, we then have

$$fI \;=\; \int_0^{r_{\rm max}} 4\pi r^2 dr\, n\, \frac{\langle L \rangle}{4\pi r^2} \;=\; n\langle L \rangle r_{\rm max}\ , \tag{4}$$

implying

$$I \;=\; \frac{n\langle L \rangle r_{\rm max}}{f}\ . \tag{5}$$

The intensity from a galaxy with luminosity L_i is equal to the mean background intensity at an "influence radius" r_i given by

$$\frac{L_i}{4\pi r_i^2} = I \qquad \Longrightarrow \qquad r_i = r_{\rm max} \times \left[\frac{L_i}{\langle L \rangle} \frac{f}{4\pi r_{\rm max}^3 n} \right]^{1/2}\ . \tag{6}$$

Note that r_i decreases as either $r_{\rm max}$ or n increases, since either change makes the contribution of a nearby galaxy less important relative to that of the numerous, distant galaxies.

For $f = 0.5$, $L_i = \langle L \rangle$, $r_{\rm max} = 240\,h^{-1}$ Mpc, and $n = 1.2 \times 10^{-3} h^3 {\rm Mpc}^{-3}$ ($\Sigma = 1\,{\rm arcmin}^{-2}\,\Delta z^{-1}$), equation (6) yields $r_i = 0.37\,h^{-1}$ Mpc (all length scales comoving). For our cosmological model, the corresponding angular and velocity scales are $\Delta\theta = 0.3'$, $\Delta V = 47\,{\rm km\ s}^{-1}$, respectively. This simple estimate suggests that galaxy photoionization could have a significant impact on the conditional flux PDF at separations of $\Delta\theta \sim 0.5'$, though its impact at $\Delta\theta \sim 2'$ is likely to be small. The optical depth of gas at temperature T, with electron density n_e, and at a distance r from a galaxy is reduced relative to the optical depth τ_u of the uniform background case by a factor

$$\tau/\tau_u = \left[1 + F_\Gamma \left(\frac{r_i}{r}\right)^2\right]^{-1} \quad , \tag{7}$$

where

$$\frac{1}{F_\Gamma} = \frac{\Gamma_u + \Gamma_c(T) n_e}{\Gamma_u} \approx 1 + 0.29(\rho/\bar{\rho}) e^{-1.578/T_5} T_5^{1/2} (1 + T_5^{1/2})^{-1} \tag{8}$$

is the ratio of the total (photo + collisional) ionization rate to the uniform photoionization rate Γ_u. The second part of equation (8) incorporates the Haardt & Madau (1996) value of $\Gamma_u = 8.3 \times 10^{-13}\,{\rm s}^{-1}$ and the Cen (1992) expression for collisional ionization at temperature $T = 10^5 T_5\,{\rm K}$ (which is also used in TreeSPH and TIPSY), assuming fully ionized gas to relate n_e to $\rho/\bar{\rho}$.

Figure 12 illustrates the impact of galaxy photoionization on the conditional mean decrement (left panels) and saturated fraction (right panels) for the $\Sigma = 1\,{\rm arcmin}^{-2}\,\Delta z^{-1}$ sample from the L50n144 simulation. We compute the influence radii via equation (6) assuming that each galaxy's ionizing flux is proportional to its SFR and that galaxies of this space density collectively produce a fraction $f = 0.5$ of the ionizing background, with fainter galaxies having no significant ionizing flux. Asterisks show the result of a simplified calculation in which we multiply the optical depths of the extracted spectra by the factor in equation (7) assuming $F_\Gamma = 1$ and a pixel-galaxy distance r corresponding to the transverse separation between the galaxy and the line of sight. As expected from the order-of-magnitude estimate above, the suppression relative to the uniform UV background case is small at $\Delta\theta \geq 1'$, but the mean decrement and saturated pixel fraction are noticeably reduced for $\Delta\theta \leq 0.5'$. Triangles show the results of a complete calculation, in which we put individual galaxy UV sources into the simulation before extracting the Lyα forest spectra with TIPSY, applying equation (7) to each SPH particle with r equal to the galaxy-particle separation and F_Γ computed from the particle's temperature and density. Here the reduction of $\langle D \rangle$ and $f_{\rm sat}$ is much smaller, and the predicted trends no longer turn over at the smallest angular separations.

Why does the approximate calculation drastically overestimate the impact of galaxy photoionization? One possibility is that collisional ionization in the hotter, denser gas within

$r \sim 0.5\,h^{-1}$ Mpc of galaxies dilutes the effect of photoionization by making $F_\Gamma \ll 1$. However, if we repeat the TIPSY calculation with $F_\Gamma = 1$ for all particles, we get a nearly identical result for the conditional mean decrement. Collisional ionization *is* important along some lines of sight (compare equation [8] to Figure 1), but these are cases where dense gas produces heavily saturated absorption, which remains saturated even if the optical depth is reduced by a factor of several. Thus, overestimating photoionization by setting $F_\Gamma = 1$ still leaves a flux decrement $D \approx 1$. The second possibility is that the gas producing absorption at the galaxy redshift is further away than the transverse separation r_t, and is shifted to the galaxy redshift by its peculiar infall velocity. Figure 13 shows the mean line-of-sight distance, weighted by optical depth, of gas that produces absorption within 20 km s^{-1} of the galaxy redshift. We consider only galaxy-spectrum pairs with $\Delta\theta < 0.5'$, implying a transverse separation $r_t < 0.6\,h^{-1}$ Mpc (comoving); results are similar for $\Delta\theta < 0.25'$. Figure 13 shows that absorption for these close pairs arises mainly in gas whose line-of-sight distance substantially exceeds the influence radius r_i and is therefore little affected by galaxy photoionization. When the optical depth is high ($\tau > 10$, left-hand panel), the mean line-of-sight distance is usually smaller, but in these cases even a factor of several reduction in τ does not move the flux decrement significantly below $D = 1$. Thus, even if galaxies contribute a large fraction of the ionizing background, local photoionization has little effect on the conditional mean decrement or saturated fraction because the real-space distance to absorbing gas is larger than the redshift-space distance, except for lines of sight that are heavily saturated.

Steidel et al. (2002) find that $\sim 3\%$ of the galaxies in their LBG sample exhibit detectable AGN activity. Given the apparent ubiquity of supermassive black holes in local galaxies, it seems plausible that most LBGs go through AGN phases, with $f_Q \sim 0.03$ being the "duty cycle", i.e. the fraction of time that these black holes are active at $z \sim 3$. Could photoionization by these low luminosity AGN have a larger impact on the local IGM than the galaxies themselves? If the interval between periods of AGN activity is longer than the time

$$t_{\rm eq} = \frac{n_{\rm HI}^{\rm eq}}{\alpha(T) n_e n_{\rm HII}} = [\Gamma_u + \Gamma_c(T) n_e]^{-1} \approx 4 \times 10^4 F_\Gamma \,{\rm yrs} \qquad (9)$$

that it takes gas to return to its equilibrium neutral hydrogen fraction after being fully ionized,[3] then we are essentially back to the uniform ionization case: the few LBGs that currently host AGN have large "proximity zones," but the rest have their normal complement of associated Lyα optical depth. However, if the process of feeding gas to the central black hole is sufficiently stochastic, then AGN might "flicker" on and off on timescales shorter than

[3]We assume that the equilibrium state is already highly ionized, so that the addition of $n_{\rm HI}^{\rm eq}$ hydrogen ions does not significantly increase $n_{\rm HII}$ or n_e. The second equality follows from the equilibrium condition $\alpha(T) n_e n_{\rm HII} = \Gamma_u n_{\rm HI}^{\rm eq} + \Gamma_c(T) n_e n_{\rm HI}^{\rm eq}$.

$t_{\rm eq}$, with individual activity cycles lasting $t_{\rm act} \lesssim f_Q t_{\rm eq} \lesssim 10^3$ years. In this case, gas around a large fraction of LBGs could be out of ionization equilibrium, with the neutral hydrogen fraction depressed by the memory of the most recent AGN outburst.

Suppose we assume that the AGN associated with the galaxies under study collectively produce 50% of the ionizing background. Since the time-averaged flux of this population equals, by assumption, the flux that we previously ascribed to the galaxies themselves, the photoionization rate at distance r during an active phase must exceed the previous value by a factor f_Q^{-1}, i.e., $\Gamma_Q = \Gamma_u (r_i/r)^2 f_Q^{-1}$. During an active phase, the neutral hydrogen density at distance r drops exponentially in time, $n_{\rm HI} \approx n_{\rm HI}^{\rm eq} \exp(-\Gamma_Q t)$, until it reaches a new equilibrium value with Lyα optical depth

$$\tau/\tau_u = \left[1 + f_Q^{-1} F_\Gamma \left(\frac{r_i}{r}\right)^2\right]^{-1} \quad . \tag{10}$$

However, at most radii the duration of the active phase will not be long enough to achieve this equilibrium. If we assume that the activity during a recombination interval $t_{\rm eq}$ occurs in a single "outburst" of duration $t_{\rm act} = f_Q t_{\rm eq}$, then the optical depth at the end of this phase is

$$\tau/\tau_u \approx \exp(-\Gamma_Q t_{\rm act}) = \exp\left[-F_\Gamma \left(\frac{r_i}{r}\right)^2\right] \ , \tag{11}$$

except that it never falls below the equilibrium value of equation (10). Note that the duty cycle f_Q cancels out of equation (11) because we have fixed the time-averaged flux of the AGN population relative to Γ_u. If the AGN are more luminous while they are on, then they must be active for less time, producing the same number of ionizing photons. We see from equation (11) that the scale over which recurrent AGN activity can affect the conditional mean flux decrement is the same influence radius that we found previously, but that departures from photoionization equilibrium allow the impact within this radius to be much stronger.

Since we have already found that much of the absorption for small $\Delta\theta$ comes from gas beyond the influence radius, we can guess from this analysis that local AGN photoionization will have little impact on the conditional mean decrement. Filled circles in Figure 12 confirm this expectation, showing a calculation based on equation (11), with sources embedded in the TIPSY spectral extraction and the same influence radii used previously. The minimum optical depth is given by equation (10) with $f_Q^{-1} = 30$. While the absorption is slightly suppressed relative to the galaxy photoionization case, the mean decrement and saturated fraction continue to increase as $\theta_{\rm max}$ decreases. These results can be considered a conservative upper limit on the effects of recurrent AGN activity, since we have ignored the partial return to equilibrium with the uniform background that will occur following any given active cycle.

If we attributed 100% of the ionizing background to the AGN associated with the observed LBGs, then the influence radii would be larger by $2^{1/2}$, but the impact on the conditional mean decrement would be only slightly stronger.

We conclude that local photoionization by galaxies or by the AGN that they host will not reverse the trend of increasing Lyα forest absorption with decreasing angular separation. Figure 13 suggests that *any* feedback mechanism must have a strong influence out to comoving distances $\sim 1\,h^{-1}$ Mpc or more if it is to effect such a reversal. We have examined simple models in which we completely eliminate neutral hydrogen out to a fixed radius around target galaxies or, alternatively, reduce the neutral fraction within this radius by a factor of three. If we eliminate all neutral hydrogen to $r = 1\,h^{-1}$ Mpc, then the conditional mean decrement drops to $\langle D \rangle \approx 0.6$ for $\theta_{\rm max} = 0.25'$ and $0.5'$, and the saturated pixel fraction falls to $f_{\rm sat} \approx 0.4$. However, a factor of three reduction out to $1\,h^{-1}$ Mpc, or complete elimination out to $0.5\,h^{-1}$ Mpc, has only small impact, comparable to that of the AGN ionization model shown in Figure 12. Elimination of neutral hydrogen out to $0.75\,h^{-1}$ Mpc has an intermediate effect, reducing $\langle D \rangle$ to ~ 0.75 and $f_{\rm sat}$ to ~ 0.5.

3.6. Redshift Errors

We showed in Figure 5 that the signature of the overdense environments of galaxies in the conditional flux PDF is substantially weakened if one looks at pixels $\sim 200\,{\rm km\ s^{-1}}$ away from the galaxy redshift instead of at the galaxy redshift itself. (The angular separation corresponding to $Hr = 200\,{\rm km\ s^{-1}}$ is $1.3'$.) Since the redshifts of LBGs are usually estimated from nebular emission or absorption lines, which could be offset relative to the mean systemic velocities, this result suggests that redshift measurement errors could have a significant impact on practical determinations of the conditional flux PDF. To address this point quantitatively, we show in Figure 14 the conditional mean decrement and saturated fraction when pixel redshifts are drawn from a Gaussian of mean zero and dispersion $\sigma_{\Delta V} = 150\,{\rm km\ s^{-1}}$ (squares) or $300\,{\rm km\ s^{-1}}$ (triangles). We use the $\Sigma = 1\,{\rm arcmin^{-2}}\,\Delta z^{-1}$ sample of the L50n144 simulation, with no galaxy photoionization. Redshift errors of $\sigma_{\Delta V} = 150\,{\rm km\ s^{-1}}$, similar to those estimated by ASSP, depress $\langle D \rangle$ and $f_{\rm sat}$ by ~ 0.05 at $\theta_{\rm max} = 2'$ and by $\sim 0.1 - 0.2$ at $\theta_{\rm max} \leq 1'$. Errors of $\sigma_{\Delta V} = 300\,{\rm km\ s^{-1}}$ depress $\langle D \rangle$ and $f_{\rm sat}$ more severely, and they largely (though not entirely) remove the trend of increasing absorption with decreasing $\theta_{\rm max}$. Redshift errors also reduce the already small differences between models with and without galaxy photoionization, since the absorption is now usually measured at a pixel further away from the galaxy.

4. Large Scale Correlations

Inspired by the observational work of ASSP and ASPS, we also examine a statistical measure that characterizes LBG-Lyα forest correlations on larger scales than those examined in §3. We divide each of the 1200 random spectra through the L50n144 simulation into L/S segments of comoving length S, where $L = 50\,h^{-1}$ Mpc is the box size. For each segment i, we compute the mean flux decrement $D_{S,i}$ along the segment, and we count the number of galaxies N_i in a cube of side length S centered on the spectral segment, making use of the simulation's periodic boundaries as necessary. We include galaxy peculiar velocities when computing positions along the line of sight, again utilizing the periodic boundaries when needed. We then consider all segments in a bin of D_S values and compute the mean value of the galaxy density contrast $\delta_{\rm gal,}{}_i = N_i/\bar{N} - 1$, where $\bar{N} = S^3 n$ is the average number of galaxies in a cubic cell of size S. We use the full resolved galaxy population of the simulation, for which $\bar{N} = 3.6(S/10\,h^{-1}\ {\rm Mpc})^3$. Poisson noise fluctuations in $\delta_{\rm gal}$ are large, and if one plots the mean value of D_S in bins of $\delta_{\rm gal}$, then a "Malmquist" type bias effectively erases any trend (there are more cells of intrinsically low density contrast that scatter to high $\delta_{\rm gal}$ than vice versa). However, D_S is not a noisy quantity, and if one has many cells, then Poisson fluctuations in $\delta_{\rm gal}$ average to zero without bias. A plot of $\delta_{\rm gal}$ in bins of D_S therefore reveals the preferential location of galaxies in overdense environments.

The solid points in the upper panels of Figure 15 show $\delta_{\rm gal}$ vs. D_S for $S = 10\,h^{-1}$ Mpc (left) and $5\,h^{-1}$ Mpc (right). Error bars represent the 1σ error on the mean, i.e., the dispersion of the $N_{\rm seg}$ values of $\delta_{\rm gal}$ in the D_S bin divided by $(N_{\rm seg} - 1)^{1/2}$. Smooth curves show the cumulative distribution of D_S values. Only a small number of segments have $D_S > 0.6$ for $S = 10\,h^{-1}$ Mpc, so the error bars on $\delta_{\rm gal}$ become large for high D_S values. Nonetheless, there is a clear, strong trend of increasing $\delta_{\rm gal}$ with increasing D_S up to $D_S \approx 0.65$, where the mean galaxy overdensity is $1 + \delta_{\rm gal} \approx 2.5$. For $S = 5\,h^{-1}$ Mpc, there are more high decrement segments and the trend of $\delta_{\rm gal}$ with D_S continues to $D_S \approx 1$.

The open points in Figure 15 show the results using dark matter particles instead of galaxies. While there is a steady trend of $\delta_{\rm dm}$ with D_S, the bias of the galaxy population is strikingly evident in the much steeper slope of the $\delta_{\rm gal} - D_S$ trend. The suppression of high redshift galaxy formation in underdense regions is especially clear, with mean density contrasts $\delta_{\rm gal}$ approaching -1 for $D_S < 0.2$. In the $S = 5\,h^{-1}$ Mpc plot, the physical requirement that $\delta \geq -1$ forces the galaxy correlation to be concave upward at low D_S. Peculiar velocities make the $\delta_{\rm dm} - D_S$ correlation significantly steeper in redshift space than in real space, as coherent flows amplify density contrasts (Sargent & Turner 1977; Kaiser 1987). However, the $\delta_{\rm gal} - D_S$ correlation is minimally affected by peculiar velocities owing to the high bias (and correspondingly low $\Omega_m^{0.6}/b$) of the galaxy population.

The rms fluctuations of D_S and density contrast are $(\sigma_D, \sigma_{\rm dm}, \sigma_{\rm gal}) = (0.13, 0.35, 1.01)$ for $S = 10\,h^{-1}$ Mpc and $(0.17, 0.56, 1.93)$ for $S = 5\,h^{-1}$ Mpc. Here we have subtracted a shot noise contribution to the galaxy dispersion assuming Poisson statistics, i.e., we define $\sigma_{\rm gal} = \left[\langle (N-\bar{N})^2\rangle - \bar{N}\right]^{1/2}/\bar{N}$, where $\langle ...\rangle$ represents an average over all cells and the division by $\bar{N}$ converts from a dispersion of N to a dispersion of $\delta_{\rm gal}$. We also subtract shot noise when computing $\sigma_{\rm dm}$, but since the number of dark matter particles is large, the correction is negligible.

The ratio of correlation slopes between $\delta_{\rm gal} - D_S$ and $\delta_{\rm dm} - D_S$ is roughly the ratio of $\sigma_{\rm gal}$ to $\sigma_{\rm dm}$, as we show in the bottom panels of Figure 15, which plot $\delta_{\rm gal}/\sigma_{\rm gal}$ or $\delta_{\rm dm}/\sigma_{\rm dm}$ against $(D_S - \bar{D})/\sigma_D$, where $\bar{D} = 0.36$ is the unconditional mean decrement. In these normalized variables, the galaxy and dark matter correlations are similar, though the constraint $\delta \geq -1$ forces the galaxy correlation to become shallower at low D_S, while the dark matter trend remains linear down to essentially zero flux decrement. Thus, linear bias appears to be a good but not perfect approximation for this statistic. Our results for dark matter agree well with those of McDonald et al. (2002), who examine the unnormalized and normalized correlations of $\delta_{\rm dm}$ with D_S using the hydro-pm approximation of Gnedin & Hui (1998). For $10\,h^{-1}$ Mpc cubes, they find a correlation with a slope of 0.6 in normalized variables, bending towards shallower slopes for $(D_{10} - \bar{D})/\sigma_D \gtrsim 1$ ($\tilde{\delta}_F \gtrsim 1$ in their notation). The box diagonals of the lower panels of Figure 15 have a slope of 0.6, so one can see by visual inspection that we obtain a similar slope and turnover, and that the slope for $5\,h^{-1}$ Mpc is nearly the same as for $10\,h^{-1}$ Mpc. McDonald et al. (2002) show that variations in the Lyα forest assumptions — e.g., the temperature-density relation, the mean decrement, or the flux power spectrum — have only a small effect on the $\delta_{\rm dm} - D_S$ correlation when it is plotted in normalized variables, so we expect the same would hold true for the $\delta_{\rm gal} - D_S$ correlations shown here.

As a quantitative measure of correlation strength, we compute the cross-correlation coefficients $r_{\rm gal,D} \equiv \langle \delta_{\rm gal}\delta_D\rangle/(\sigma_{\rm gal}\sigma_D)$ and $r_{\rm dm,D} \equiv \langle \delta_{\rm dm}\delta_D\rangle/(\sigma_{\rm dm}\sigma_D)$, where $\delta_D \equiv D_S - \bar{D}$, and the averages are over all cells/segments. At $10\,h^{-1}$ Mpc, the correlation coefficients are 0.49 and 0.66 for galaxies and dark matter, respectively. The coefficients for dark matter are substantially higher in redshift space than they would be in real space, while for galaxies they are nearly identical in redshift space and real space. The cross-correlation between galaxies and dark matter is high, $r_{\rm gal,dm} \approx 0.95$, and would again be lower in real space. However, $r_{\rm gal,D}$ is lower than $r_{\rm dm,D}$ for both cube sizes, perhaps because the curvature of the $\delta_{\rm gal} - \delta_{\rm dm}$ relation at low δ_D suppresses the contribution of these cells to the correlation coefficient. As emphasized by ASPS, the correlation $r_{\rm dm,D}$ is necessarily below unity because the flux decrement D_S gives a (non-linearly) weighted average of the density along the line of sight through the cube center, and this quantity is imperfectly correlated with the mean

density of the cube as a whole.

Figure 16 resembles Figure 15, but here the $\delta - D_S$ correlation for galaxies is compared to that for the 450 most massive friends-of-friends halos, with the same weighting schemes used in Figures 8 and 9. The correlation is insensitive to halo weighting except at high D_S, since massive halos are rare in lower density regions and assigning them higher weights makes little difference. At high D_S, the contrast is usually stronger when massive halos receive more weight, as one would expect, but our results in this regime are fairly noisy. The rms density contrast σ *is* sensitive to the relative halo weighting, so in normalized variables (bottom panels) the correlation is actually shallower when massive halos are more strongly weighted, and correlation coefficients are correspondingly lower. The galaxy correlations most closely resemble those for equally weighted halos or halos with weight $\propto M^{0.5}$, and both the rms fluctuations of galaxies and their correlations with Lyα flux decrement are lower than they would be if the number of galaxies were proportional to halo mass.

5. Discussion

The conditional flux PDF considered in §3 and the $\delta_{\rm gal} - D_S$ correlation considered in §4 offer empirical tools with which to test the prediction that high redshift galaxy formation is strongly biased, and to search for the influence of star-forming galaxies on the surrounding IGM. The predicted flux PDF shifts systematically towards higher absorption in pixels close to galaxies — the smaller the angular separation, the higher the mean decrement and saturated fraction and the lower the transparent fraction. The shift of the flux PDF remains strong in spectra smoothed over 50, 100, or 200 km s^{-1}, allowing further tests of the predicted scale of correlations and demonstrating that moderate resolution spectra of relatively faint AGN may profitably be used in such investigations. The influence of galaxy proximity is much stronger than that of random mass elements (dark matter particles), a clear signature of the preferential formation of galaxies in overdense environments.

The conditional flux PDF is only weakly sensitive to the baryonic mass or SFR of the sample galaxies over the range that we have investigated, spanning more than two orders of magnitude in space density. It is also insensitive to the occupation distribution of dark matter halos. While this insensitivity reduces the utility of this statistic as a diagnostic of galaxy formation, it increases its power as a diagnostic of galaxy feedback, since the no-feedback prediction is robust. We have investigated the potential impact of one feedback process, galaxy photoionization, and find that it has only a small impact on the conditional mean flux decrement, even if the bright galaxies used in the conditional PDF measurement contribute a large fraction of the ionizing background. The local flux from these galaxies is

comparable to the uniform background at "influence radii" $r_i \sim 0.4\,h^{-1}$ Mpc (comoving), corresponding to angular separations $\Delta\theta \sim 0.3'$, and in the absence of peculiar velocities this flux would drive down the average absorption at smaller separations. However, the gas producing absorption at the galaxy redshift is typically $\gtrsim 1\,h^{-1}$ Mpc away along the line of sight, where it is minimally affected. Recurrent AGN activity, with associated departures from photoionization equilibrium, can have a stronger impact within r_i, but it does not substantially increase the typical influence radius, so it still has little effect on the flux decrement close to galaxies. The predicted dependence of $\langle D \rangle$ on $\Delta\theta$ *is* sensitive to redshift measurement errors, with rms errors $\sigma_{\Delta V} \gtrsim 150\,\mathrm{km\ s^{-1}}$ significantly depressing the absorption signature.

There is significant overlap between our study of the conditional flux PDF and the recent paper by Croft et al. (2002a), who independently investigate the conditional mean decrement statistic and other measures of LBG-Lyα forest correlations, using similar numerical techniques. As discussed in §3, Croft et al. (2002a) and ASSP define the conditional mean decrement in a slightly different way from us, averaging over all galaxy-pixel pairs in bins of redshift-space separation Δ_r instead of considering only pixels at the galaxy redshift. Averaging in redshift-space separation bins instead of angular separation bins at $\Delta V = 0$ (like the "differential" points of Figures 10 and 14) has a small but noticeable impact, yielding slightly less absorption at a given comoving distance. To facilitate comparison to the Croft et al. (2002a) and ASSP results, we have calculated their version of the mean decrement statistic for the 150 highest SFR galaxies in the L50n144 simulation ($\Sigma = 1\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$). Results are shown in Figure 17, assuming rms redshift determination errors of 0, 150, and $300\,\mathrm{km\ s^{-1}}$ (triangles, squares, and pentagons, respectively).

Croft et al. (2002a) adopt slightly different cosmological parameters, with $\Omega_m = 0.3$, $\sigma_8 = 0.9$ instead of $\Omega_m = 0.4$, $\sigma_8 = 0.8$. Relative to our L50n144 run, their impressive, 2×300^3 particle simulation has a volume 3.3 times smaller but mass resolution 30 times higher. Solid and dotted curves in the left hand panel of Figure 17 show their results for galaxies with $M_b > 10^{10} M_\odot$ and $M_b > 10^{11} M_\odot$ (Croft et al., fig. 8), kindly provided by R. Croft. Given the differences in cosmological and numerical parameters, the two independent calculations agree remarkably well. The comoving space densities of the two Croft et al. samples, 0.028 and 0.0008 $h^3\,\mathrm{Mpc^{-3}}$, correspond to $\Sigma = 24\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$ and $\Sigma = 0.67\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$ for our cosmology, so they are comparable to our $\Sigma = 48\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$ and $\Sigma = 1\,\mathrm{arcmin^{-2}}\,\Delta z^{-1}$ samples, respectively. As shown in Figure 10, we find no significant difference in the conditional mean decrement for these two samples, while Croft et al. find stronger absorption for the more massive galaxies. Since their $M_b > 10^{11} M_\odot$ sample comprises only 30 galaxies, this difference in the results does not seem too serious at present. Croft et al. also consider higher density samples, with mass resolution thresholds below the

resolution limit of our L22n128 simulation, and they find a continuing though modest trend of decreasing conditional mean decrement with decreasing galaxy mass threshold. In agreement with Croft et al., we find that redshift determination errors of $\sigma_{\Delta V} \sim 150 - 300\,\mathrm{km\ s^{-1}}$ have a substantial impact on the predicted trend of $\langle D \rangle$ with galaxy separation, and that galaxy photoionization does not have a large impact.

ASSP have investigated LBG-Lyα forest correlations observationally, carrying out an LBG spectroscopic survey in fields around six quasar lines of sight. Their estimates of the conditional mean decrement, kindly provided by K. Adelberger, are shown by the filled points in Figure 17. These are somewhat different from the points shown by Croft et al. (2002a) because of changes to ASSP's analysis procedures, in particular a scaling of optical depths with redshift that should make the results more directly comparable to predictions like the ones presented here. The predicted and observed trends agree reasonably well down to comoving separations $\sim 3\,h^{-1}$ Mpc, with a smooth increase of ~ 0.1 in $\langle D \rangle$. However, ASSP's trend flattens towards smaller separations, while our predicted trend rises and steepens. Most strikingly, ASSP find a *decrease* in $\langle D \rangle$ at separations $r < 1\,h^{-1}$ Mpc, where the predicted absorption is strongest. If the rms redshift errors are $\sigma_{\Delta V} \sim 300\,\mathrm{km\ s^{-1}}$ rather than the $\sim 150\,\mathrm{km\ s^{-1}}$ that ASSP estimate, then the discrepancy between the predictions and observations is reduced, but the two data points at $r < 1\,h^{-1}$ Mpc remain in serious conflict.

ASSP suggest that the flattening and eventual decline of the mean decrement in close proximity to LBGs is a signature of galaxy feedback on the surrounding IGM, perhaps caused by the galactic scale winds suggested by the nebular line profiles of LBGs (Pettini et al. 2001, 2002). Croft et al. (2002a) consider a variety of galaxy feedback models and show that they can roughly reproduce the ASSP mean decrement results if $\sim 10\%$ of the available supernova energy is converted to kinetic energy of a coherent, galactic scale outflow that sweeps up all of the IGM in its path. While this wind model is *ad hoc* and idealized, it does show that it is energetically possible for supernova feedback to influence the conditional mean decrement out distances of a few h^{-1} Mpc (comoving). The range of the effect is larger than that of individual wind bubbles because the galaxies themselves are correlated, and a line of sight $\sim 2\,h^{-1}$ Mpc from one galaxy may pass through the wind-blown bubble of another. We have shown (in agreement with Croft et al. 2002a and ASSP) that galaxy photoionization cannot explain the observed downturn in absorption, and we have shown that this remains the case even if one considers departures from ionization equilibrium that could be caused by recurrent AGN activity.

The results of §3.5 (Figure 13 in particular) have an important implication for wind models: peculiar velocities allow gas at comoving distances $1 - 2\,h^{-1}$ Mpc to produce strong

absorption at redshift-space separations $0 - 0.5\,h^{-1}$ Mpc, so the required wind radius R_w is larger than one would first guess from the scale of the observed downturn. Figure 18 demonstrates this point explicitly, using the $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ sample of the L22n128 sample (the top 40 galaxies in the box). Squares, pentagons, and hexagons show the conditional mean decrement computed after eliminating all neutral hydrogen within a comoving radius $R_w = 0.75$, 1.0, or $1.5\,h^{-1}$ Mpc around each of the target galaxies (see Kollmeier et al. 2002 for further discussion). Only the $R_w = 1.5\,h^{-1}$ Mpc case produces a downturn comparable to the observed one, and even it does not reproduce ASSP's innermost data point (as we discuss further below). Stars show a case in which we put winds around all 641 resolved galaxies in the box and scale the wind volume in proportion to each galaxy's baryon mass, $R_w \propto M_b^{1/3}$. This model has $R_w \approx 1\,h^{-1}$ Mpc for the 40th-ranked galaxy in the box and a maximum $R_w \approx 2\,h^{-1}$ Mpc. The impact is similar to, though slightly weaker than, that of the constant-radius $R_w = 1.5\,h^{-1}$ Mpc model. All of these models are, of course, highly simplified, but they show that galactic winds must efficiently eliminate neutral hydrogen out to large distances to explain ASSP's results. The energetic requirements for such winds are not impossible to meet (see Croft et al. 2002a), but they do require sustained outflow speeds of several hundred km s^{-1} for $t \gtrsim 1$ Gyr and effective entrainment of surrounding gas. Winds with large filling factor can have a noticeable impact on the flux power spectrum of the Lyα forest, though for $R_w \leq 1\,h^{-1}$ Mpc the effect is small (Weinberg et al. 2003).

The innermost ASSP data point is based on only three galaxy-QSO pairs, and the error estimation procedure is somewhat *ad hoc* (see ASSP), so it is natural to ask whether the discrepancy with the data is statistically significant. To address this point, we randomly selected 3-tuples of simulated galaxy-sightline pairs with $\Delta\theta < 0.4'$ and computed the mean decrement in the redshift separation bin $\Delta_r = 0 - 0.5\,h^{-1}$ Mpc, repeating the experiment 500 times. Figure 19 shows the cumulative distribution of these 3-tuple mean decrements, for our standard model with no redshift measurement errors (solid line), for redshift errors $\sigma_{\Delta V} = 300\,\mathrm{km\ s^{-1}}$ (dashed line), and for the $R_w = 1.5\,h^{-1}$ Mpc "wind" model from Figure 18. Vertical lines mark the unconditional mean flux decrement, $\langle D \rangle = 0.36$, and the ASSP measurement for this separation, $\langle D \rangle = 0.11$. With $\sigma_{\Delta V} = 300\,\mathrm{km\ s^{-1}}$, the 3-tuple mean decrement is below 0.36 only $\sim 5\%$ of the time, and none of our 500 trials have a mean decrement as low as the ASSP value. Even for the $R_w = 1.5\,h^{-1}$ Mpc wind model, only $\sim 5\%$ of the trials have a 3-tuple mean decrement as low as ASSP's. We did not incorporate redshift errors in the wind model, but (in contrast to the no-wind case) they tend to raise the predicted mean decrement by moving the measurement point further from the influence of the galaxy. We have only 150 simulated galaxies contributing to the solid and dotted curves and only 40 contributing to the dashed curve, so our 3-tuples are not independent and we may therefore underestimate the incidence of rare, low absorption cases. Nonetheless,

it appears that the ASSP results cannot easily be explained away by appealing to small number statistics.

Given the stringent demands on wind models, it is worth considering alternative explanations. One possibility is that the IGM at these distances from galaxies is multi-phase, with processes that the simulations cannot resolve causing the geometrical cross section of the neutral phase to drop sharply. However, the typical physical conditions at these distances from galaxies are not extreme (see Figure 1), so it would be difficult to reconcile such an explanation with the overall success of simulations in reproducing (unconditional) statistical properties of the Lyα forest.

A final possibility worth considering is that extended Lyα *emission* from the target galaxies "fills in" the Lyα forest at the redshift of emission. Extended Lyα "blobs" have been observed to be associated with LBG's (Steidel et al. 2000) having angular extents $\sim 15''$ and AB apparent magnitudes 21.02 and 21.14 in an 80Å Lyα band. Cooling radiation from gas settling into massive galaxies at $z = 3$ naturally produces Lyα flux of this order, and numerical simulations predict that such blobs should be present around typical LBGs (Fardal et al. 2001). Thus, a galaxy's Lyα cooling radiation could plausibly replace the absorbed continuum flux of a 21st-magnitude background quasar. There are two substantial and partly cancelling corrections to this estimate. First, the Lyα emission extends over $\sim 1000\,\mathrm{km\ s^{-1}}$ (16 Å) rather than the 80Å bandpass used by Steidel et al. (2000), so the flux density at the Lyα wavelength is a factor of five higher (-1.75 mag). Second, at a separation $\Delta\theta \sim 15 - 20$" from a target galaxy, the ~ 0.8" slits used by ASSP should intercept only $\sim 0.8/2\Delta\theta \sim 0.02 - 0.03$ of the galaxy's extended Lyα flux ($\sim +4$ mag), assuming constant surface brightness out to $\Delta\theta$. The three quasars that contribute to ASSP's innermost data point have G-band AB magnitudes of 20.1, 21.6, and 23.4, so with the +2.25 mag net correction, it appears that Lyα emission could replace the absorbed quasar flux only for the faintest of the three quasar targets, unless the galaxies in question are even brighter than the Steidel et al. (2000) blobs. Furthermore, a fourth pair involving the $G = 17.8$ quasar Q0302-0019 shows no sign of absorption near the galaxy redshift, and in this case the quasar is clearly too bright for galaxy emission to compete with it. (This pair and two others are dropped from ASSP's $\langle D \rangle$ calculation because of possible Lyβ contamination.) At this point, therefore, the Lyα emission explanation seems unlikely, but future observations that have more close pairs involving bright quasars will be able to test it conclusively.

On the larger scales probed by the $\delta_{\mathrm{gal}} - D_S$ correlation, our predictions are in qualitative agreement with the observational results reported by ASSP and ASPS. As the Fluctuating Gunn-Peterson Approximation suggests, higher mean flux decrements tend to arise in cells of higher mass density, and these in turn have higher average galaxy overdensity. The limited

size of our largest volume simulation, a single $50\,h^{-1}$ Mpc cube, leaves statistical and systematic uncertainties in our predictions. Nonetheless, our results for the $\delta_{\rm dm} - D_S$ correlation at $S = 10\,h^{-1}$ Mpc are similar to those found by McDonald et al. (2002) using hydro-PM simulations (with $40\,h^{-1}$ Mpc cubes but multiple realizations, and significantly higher mass resolution). The $\delta_{\rm gal} - D_S$ correlation is much steeper than the $\delta_{\rm dm} - D_S$ correlation because of the strong bias of the high redshift galaxy population, and the suppression of galaxy formation in underdense environments is especially clear. The ratio of correlation slopes is approximately equal to the ratio of rms fluctuation amplitudes, $\sigma_{\rm gal}/\sigma_{\rm dm}$, but a "linear bias" model does not fully capture the behavior of the $\delta_{\rm gal} - D_S$ correlation, in particular the curvature at low D_S imposed by the constraint that $\delta_{\rm gal} \geq -1$.

Our $50\,h^{-1}$ Mpc simulation predicts a correlation coefficient $r_{\rm gal,D} = 0.49$ between galaxy density contrast and averaged flux decrement in $5\,h^{-1}$ Mpc or $10\,h^{-1}$ Mpc cells. This is smaller than the value $r = 0.68 \pm 0.06$ that ASPS measure for $S \approx 10\,h^{-1}$ Mpc, with a 1σ error bar derived from the dispersion among five lines of sight. However, we cannot assign a reliable "theoretical error bar" to our predicted $r_{\rm gal,D}$ without modeling ASPS's procedures in greater detail. Even for an ideal, volume-limited sample, there are subtle issues associated with shot noise subtraction and estimator biases (Hui & Sheth, in preparation), and corrections for redshift selection functions or sample incompleteness could also have a significant impact on estimates of r and its uncertainty. Accurate assessment of the discrepancy between predicted and observed correlation coefficients probably requires larger simulation volumes, which would allow creation of multiple independent "surveys" comparable to the observed one. Fortunately, we find that the $\delta - D_S$ correlations for galaxies are similar to those for halos of the same space density, so it should be possible to carry out such assessments using large N-body simulations and the PM or hydro-PM approximation (Croft et al. 1998; Gnedin & Hui 1998) for the Lyα forest. The absolute trend (i.e., not normalized by standard deviations) of $\delta_{\rm gal}$ with D_S is the best test for the strong bias of high redshift galaxies predicted by theoretical models.

As the size of LBG spectroscopic samples grows, moderate resolution spectra of fainter quasars can greatly increase the number of lines of sight available for measuring LBG-Lyα forest correlations. These larger samples will shrink the error bars of the existing measurements, and they will eventually allow measurement of the full conditional flux PDF for different subsets of the LBG population. The potential impact of redshift determination errors can be investigated empirically by comparing results for galaxies whose redshifts are estimated by different methods (e.g., emission vs. absorption lines). Models with galactic winds or other forms of feedback can be tested using the full PDF rather than the mean decrement alone. Variations in feedback from one galaxy to another should widen the range of flux decrement values in addition to reducing the mean, and winds that sweep up the

IGM should produce many lines of sight with essentially zero flux decrement. Perhaps the best diagnostic of galaxy feedback will be separate measurements for galaxy subsets with high and low star formation rates, or older and younger stellar populations, since one would expect the young/high SFR galaxies to have had more vigorous activity in the recent past. Systematic differences in the activity of different galaxy populations should produce systematic differences in the corresponding conditional flux PDFs, and our results show that such differences would be unlikely to arise from differences in the spatial clustering of these populations. A full accounting of the relation between LBGs and the Lyα forest will test our basic expectations for the way that dark matter shapes the large scale distribution of high redshift galaxies, and it will inform our picture of the ways that galaxies can affect the high redshift IGM.

We thank Kurt Adelberger, Rupert Croft, Mark Fardal, Patrick McDonald, Jordi Miralda-Escudé, Max Pettini, Alice Shapley, and Chuck Steidel for helpful discussions of these issues and of their results. We also thank Kurt Adelberger and Rupert Croft for providing results shown in Figure 17. JAK acknowledges the support of a university fellowship at Ohio State, and DHW acknowledges the hospitality and support of the Institut d'Astrophysique de Paris and the French CNRS during much of this work. This project was also supported by NASA grant NAG5-3525, NSF grant AST-0098515, and Hubble Fellowship Grant HST-HF-01128.01-A from the Space Telescope Science Institute.

REFERENCES

Adelberger, K. L., Steidel, C. C., Giavalisco, M., Dickinson, M., Pettini, M., & Kellogg, M. 1998, ApJ, 505, 18

Adelberger, K.L., Steidel, C.C., Shapley, A.E., Pettini, M. 2002a, ApJ, 584,45 (ASSP)

Adelberger, K.L., Steidel, C.C., Pettini, M., & Shapley, A. E. 2002b, in preparation (ASPS)

Aguirre, A., Hernquist, L., Schaye, J., Weinberg, D. H., Katz, N., & Gardner, J. 2001, ApJ, 561, 521

Bagla, J.S. 1998, MNRAS, 297, 251

Baugh, C. M., Cole, S., Frenk, C. S. & Lacey, C. G. 1998, ApJ, 498, 504

Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2001, MNRAS, 327, 1041

Berlind, A. A. & Weinberg, D.H. 2002, ApJ, 575, 587

Bernardi, M., et al. 2003, AJ, 125, 32

Bi, H. & Davidsen, A.F. 1997,ApJ, 479, 523

Bullock, J.S., Wechsler, R.H., & Somerville, R.S. 2001, MNRAS, 329, 246B

Cen, R. 1992, ApJS, 78, 341

Cen, R. 1997, ApJ, 479, L85

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

Cen, R. & Ostriker, J. P. 2000, ApJ, 538, 83

Chen, X., Weinberg, D. H., Katz, N., & Dave', R. 2002, ApJ, submitted, astro-ph/0203319

Cooray, A. 2002, ApJ, 576, L105

Cooray, A. & Sheth, R. 2002, Phys Rep, 372, 1

Croft, R.A.C., Weinberg, D.H., Katz, N., Hernquist, L., 1997, ApJ, 488, 532

Croft, R. A. C., Weinberg, D.H., Katz, N., & Hernquist, L. 1998, ApJ, 495, 44

Croft, R. A. C., Weinberg, D.H., Pettini, M., Katz, N., & Hernquist, L. 1999, ApJ, 520, 1

Croft, R. A. C., Di Matteo, T., Davé, R., Hernquist, L., Katz, N., Fardal, M., & Weinberg, D. H. 2001, ApJ, 557, 67

Croft, R.A.C., Hernquist, L., Springel, V., Westover, M., & White, M. 2002a, ApJ, 580, 634

Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., Tytler, D. 2002b, ApJ, 581, 20

Davé, R., Dubinski, J., & Hernquist, L. 1997, New Astronomy, 2, 277

Davé, R., Hernquist, L., Katz, N., & Weinberg, D.H. 1999, ApJ, 511, 521

Davé, R., Cen, R., Ostriker, J. P., Bryan, G. L., Hernquist, L., Katz, N., Weinberg, D. H., Norman, M. L., & O'Shea, B. 2000, ApJ, 552, 473

Davé, R., Katz, N., & Weinberg, D. H. 2002, ApJ, 579, 23

Fardal, M. A., Katz, N., Weinberg, D. H., Davé, R., & Hernquist, L. 2002, ApJ, in press, astro-ph/0107290

Fardal, M. A., Katz, N., Gardner, J.P., Hernquist, L., Weinberg, D. H., & Davé, R. 2001, ApJ, 562, 605

Fardal, M. A. & Shull, J. M. 1993, ApJ, 415, 524

Gelb, J. M., & Bertschinger, E. 1994, ApJ, 436, 467

Giavalisco, M. & Dickinson, M. 2001, ApJ, 550, 177

Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44

Governato, F., Baugh, C.M., Frenk, C.S., Cole, S., Lacey, C.G., Quinn, T.R. & Stadel, J. 1998, Nature, 392, 359

Haardt, F. & Madau, P. 1996, ApJ, 461, 20

Hernquist, L. & Katz, N. 1989, ApJS, 70, 419

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

Hui, L. & Gnedin, N.Y. 1997, MNRAS, 292, 27

Jing, Y. P., Mo, H. J., & Börner, G. 1998, ApJ, 494, 1

Kaiser, N. 1984, ApJ, 284, L9

Kaiser, N. 1987, MNRAS, 227, 1

Katz, N., Weinberg D.H. & Hernquist, L. 1996, ApJS, 105, 19 (KWH)

Katz, N., Hernquist, L., & Weinberg, D. H. 1999, ApJ, 523, 463

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 307, 529

Kollmeier, J.A., Weinberg, D.H., Davé, R., & Katz, N. 2003, in "The Emergence of Cosmic Structure" eds. S. Holt & C. Reynolds, AIP Conference Proceedings, New York, p.191, astro-ph/0212355

Ma, C., & Fry, J. N. 2000, ApJ, 543, 503

Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648

Marinoni, C., & Hudson, M. J. 2002, ApJ, 569, 101

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1

McDonald, P., Miralda-Escudé, J., & Cen, R. 2002, ApJ, 580,42

Miralda-Escudé J., Cen R., Ostriker, J.P., & Rauch, M. 1996, ApJ, 471, 582

Miralda-Escudé, J., Rauch,. M., Sargent, W.L.W., Barlow, T.A., Weinberg, D.H., Hernquist, L., Katz, N., Cen, R. & Ostriker, J.P. 1998, in Proc. of the 13th IAP Colloquium, Structure and Evolution of the IGM from QSO Absorption Line Systems, eds. P. Petitjean & S. Charlot, (Paris: Nouvelles Frontières), p. 155, astro-ph/9710230

Mo, H.J. & Fukugita, M. 1996, ApJ, 467, L9

Mo, H.J., & White S.D.M. 1996, MNRAS, 282, 1096

Moscardini, L., Coles, P., Lucchin, F. & Matarrese, S. 1998, MNRAS, 299, 95

Moustakas, L. A. & Somerville, R. S. 2002, ApJ, in press, astro-ph/0110548

Murali, C., Katz, N., Hernquist, L., Weinberg, D. H., & Davé, R. 2002, ApJ, 571, 1

Nagamine, K., Fukugita, M., Cen, R., Ostriker, J.P., 2001, MNRAS, 327L, 10N

Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144

Pearce, F. R., Jenkins, A., Frenk, C. S., White, S. D. M., Thomas, P. A., Couchman, H. M. P., Peacock, J. A., & Efstathiou, G. 2001, MNRAS, 326, 649

Pettini, M., Shapley, A. E., Steidel, C. C., Cuby, J., Dickinson, M., Moorwood, A. F. M., Adelberger, K. L., & Giavalisco, M. 2001, ApJ, 554, 981

Pettini, M., Rix, S. A., Steidel, C. C., Adelberger, K. L., Hunt, M. P., & Shapley, A. E. 2002, ApJ, 569, 742

Press, W.H., Rybicki, G.B. & Schneider, D.P. 1993, ApJ, 414, 64

Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., & Ostriker, J. P., 1997, ApJ, 489, 7

Sargent, W. W., & Turner, E. L. 1977, ApJ, 212, L3

Schirber, M., & Bullock, J. S. 2003, ApJ, 584, 110

Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20

Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P. 2000, ApJS, 130, 67

Scranton, R. 2002, MNRAS, 332, 697

Seljak, U. 2000, MNRAS, 318, 203

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17

Steidel, C.C., Adelberger, K.L, Shapley, A., Pettini, M., Dickinson, M., & Giavalisco, M. . 2000, ApJ, 532, 170

Steidel, C., Hunt, M.,Shapley, A., Adelberger, K., Pettini, M., Dickinson, M., & Giavalisco, M. 2002, ApJ, 576, 653

Steidel, C.C., Pettini, M., & Adelberger, K. L. 2001, ApJ, 546, 665

Theuns, T., Leonard, A.,Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, MNRAS, 301, 478

Weinberg, D. H., et al. 1999a, in Evolution of Large Scale Structure: From Recombination to Garching, eds. A.J. Banday, R. K. Sheth, & L. N. Da Costa, (Twin Press: Vledder NL), p. 346, astro-ph/9810142

Weinberg, D. H., Davé, R., Gardner, J. P., Hernquist, L., & Katz, N. 1999b, in Photometric Redshifts and High Redshift Galaxies, eds. R. Weymann, L. Storrie-Lombard, M. Sawicki, & R. Brunner, ASP Conference Series, San Francisco, p. 341, astro-ph/9908133

Weinberg, D.H., Hernquist, L. & Katz, N. 1997, ApJ, 477, 8

Weinberg, D. H., Katz, N., & Hernquist, L. 1998, in Origins, eds. J. M. Shull, C. E. Woodward, & H. Thronson, (ASP Conference Series: San Francisco), p. 21, astro-ph/9708213

Weinberg, D.H., Davé, R., Katz, N., & Hernquist, L. 2002a, in preparation

Weinberg, D.H., Hernquist, L., Katz, N., 2002b, ApJ, 571, 15

Weinberg, D.H., Davé, R., Katz, N., & Kollmeier, J.A. 2003, in "The Emergence of Cosmic Structure" eds. S. Holt & C. Reynolds, AIP Conference Proceedings, New York, p.157, astro-ph/0301186

Yoshikawa, K., Taruya, A., Jing, Y. P., Suto, Y. 2001, ApJ, 558, 520

Zehavi, I., Weinberg, D.H., Zheng, Z., Berlind, A.A., Frieman, J.A., Scoccimarro, R., Sheth, R.K., Blanton, M.R., Tegmark, M., Mo, H.J., et al. 2003, astro-ph/0301280

Zhang, Y., Anninos, P., & Norman, M.L. 1995, ApJ, 453, L57

Table 1. Simulation Parameters[a]

Name	$L_{\rm box}$ (h^{-1} Mpc)	N	$\epsilon_{\rm grav}$ (h^{-1} kpc)	$64m_{dark}$ ($M_\odot$)	$64m_{sph}$ ($M_\odot$)
L11n64	11.11	64^3	3.5	4.9×10^{10}	6.7×10^9
L11n128	11.11	128^3	1.75	6.2×10^9	8.3×10^8
L22n128	22.22	128^3	3.5	4.9×10^{10}	6.7×10^9
L50n144	50.00	144^3	7	4.0×10^{11}	5.4×10^{10}

[a] Length units are comoving. In all cases, the adopted cosmology is $\Omega_m = 0.4$, $\Omega_\Lambda = 0.6$, $h = 0.65$, $\Omega_b = 0.02h^{-2}$, $n = 0.95$, $\sigma_8 = 0.80$, and the output redshift is $z = 3$. The angular size, velocity depth, and redshift depth of the $22.222\,h^{-1}$ Mpc cube are $18.74'$, $2844\,{\rm km\ s^{-1}}$, and 0.0379, and they scale in proportion to $L_{\rm box}$ for the other simulations. $\epsilon_{\rm grav}$ is the gravitational softening length (equivalent Plummer softening), setting the approximate spatial resolution.

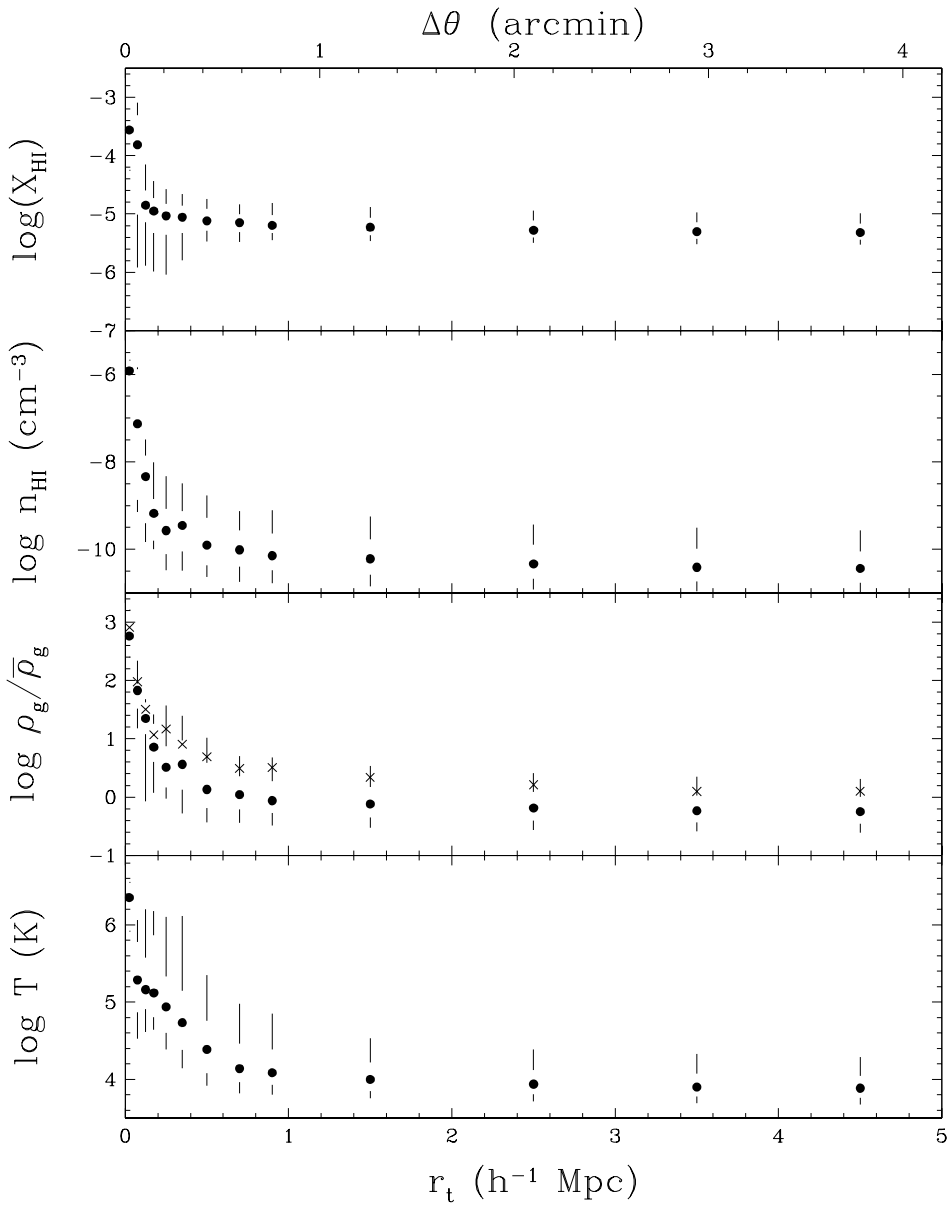

Fig. 1.— Neutral hydrogen fraction, neutral hydrogen density, gas overdensity, and gas temperature (top to bottom) as a function of comoving transverse separation from galaxies, computed using all resolved galaxies in the L22n128 simulation. Filled circles show median values for pixels at the galaxy redshift and this transverse separation, while the inner and outer edges of the vertical segments mark the 25–75% and 10–90% ranges, respectively. In the gas density panel, ×'s show the mean overdensity, which is always higher than the median because of the skewed distribution of $\rho_g/\bar{\rho}_g$. For our cosmology, a transverse separation of $1\,h^{-1}$ Mpc corresponds to $0.843'$. Scales along the top and bottom axes are marked in arc-minutes and comoving h^{-1} Mpc, respectively.

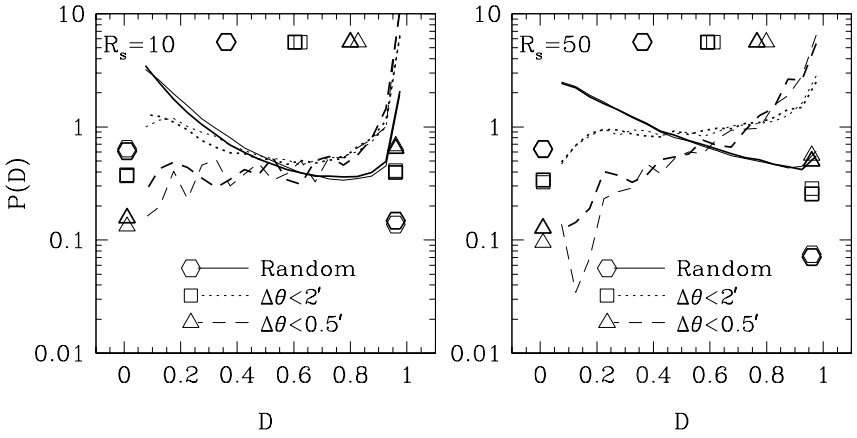

Fig. 2.— Mass resolution effects on the conditional and unconditional flux PDFs, for spectra smoothed over 10 km s^{-1} (left) or 50 km s^{-1} (right). Heavy curves and symbols are computed for galaxies with baryonic mass $M > 512 m_{\rm SPH}$ in the L11n128 simulation, while light curves and symbols are computed for galaxies with $M > 64 m_{\rm SPH}$ in the L11n64 simulation. If there were no resolution effects at all on the galaxy populations or the spectra, then the results would agree perfectly. Solid curves show the unconditional PDF, computed using all pixels along random lines of sight through the simulation box. Dotted and dashed curves show $P(D)$ for pixels at the same redshift as a galaxy with transverse separation $\Delta\theta < 2'$ or $\Delta\theta < 0.5'$, respectively. Symbols along the top, right, and left axes show, respectively, the mean flux decrement, the fraction of "saturated" pixels (with $D > 0.9$), and the fraction of "transparent" pixels (with $D < 0.55$), computed from the corresponding PDF.

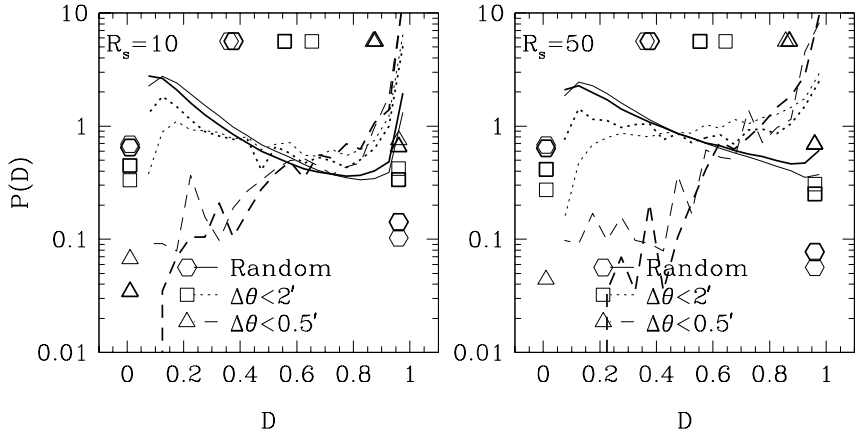

Fig. 3.— Comparison of the unconditional and conditional flux PDFs from the L22n128 and L50n144 simulations, in the same format as Fig. 2. Heavy curves and symbols show results using the 37 galaxies with $M > 512 m_{\rm SPH}$ in L22n128, and light curves and symbols show results using the 450 galaxies with $M > 64 m_{\rm SPH}$ (the same physical mass threshold) in L50n144. Differences can arise from systematic effects of resolution, box size, or random differences in the structures present in the two simulations. In the right hand panel, the heavy triangle representing the saturated fraction of L22n128 is superposed on the light triangle, and the heavy triangle representing the transparent fraction is off the bottom of the plot.

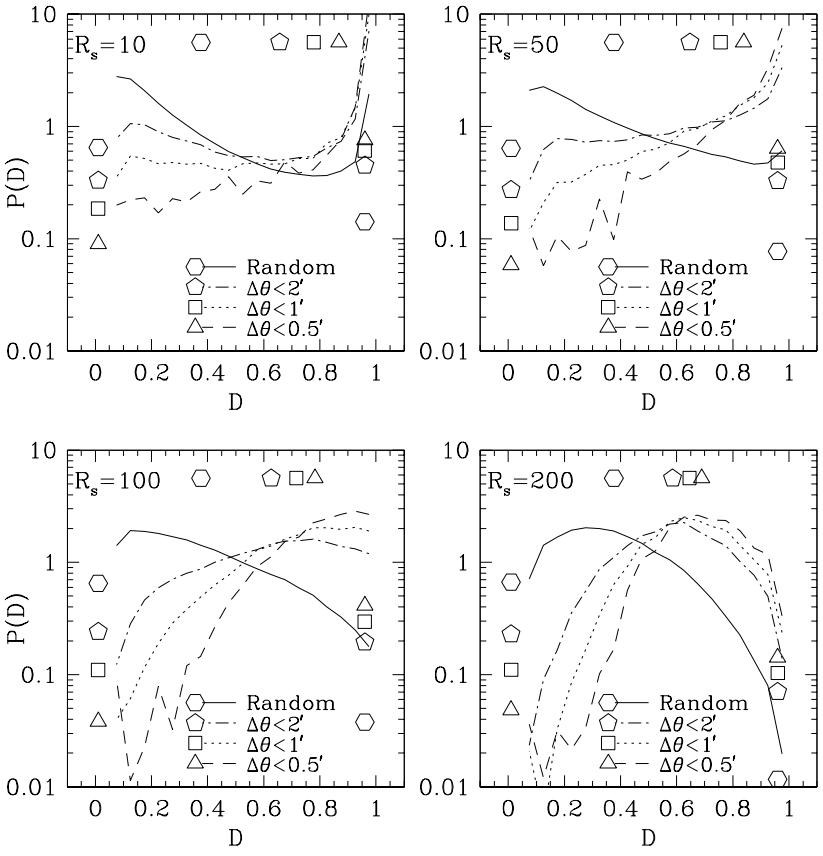

Fig. 4.— The unconditional and conditional flux PDFs computed using all resolved galaxies ($M > 64 m_{\rm SPH}$) in the L22n128 simulation. The four panels show results for spectral smoothing lengths of 10, 50, 100, and 200 km s^{-1} (Gaussian 1σ width), respectively. The format is similar to Figs. 2 and 3: solid curves show $P(D)$ for all pixels in random lines of sight, while dot-dashed, dotted, and dashed curves show $P(D)$ for pixels at the same redshift as galaxies with transverse separations $\Delta\theta < 2'$, $1'$, and $0.5'$, respectively. Symbols along the top, right, and left axes show the mean decrement, saturated fraction, and transparent fraction.

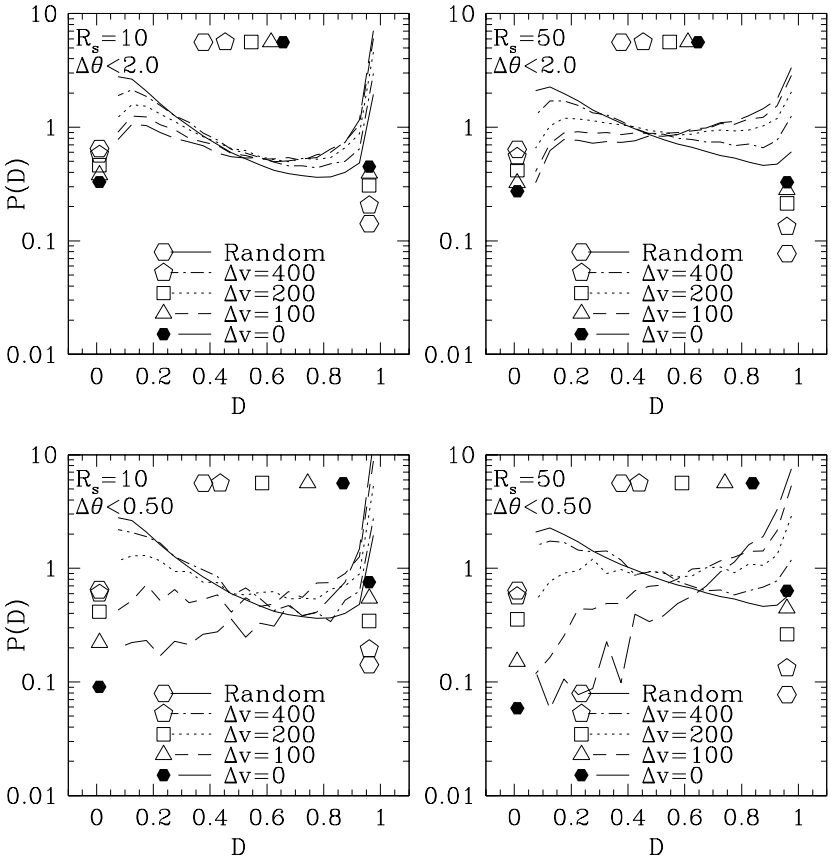

Fig. 5.— Dependence of the conditional flux PDF on velocity offset ΔV, computed using all resolved galaxies in the L22n128 simulation. Top and bottom panels show angular proximity conditions $\Delta\theta < 2'$ and $\Delta\theta < 0.5'$, respectively, with spectral smoothing of 10 km s^{-1} (left) or 50 km s^{-1} (right). In each panel, solid curves show the unconditional flux PDF, and long-dashed curves show the conditional PDF for pixels at the same redshift as a galaxy, as in Fig. 4. Dashed, dotted, and dot-dashed curves show conditional PDFs in pixels offset from the galaxy systemic velocity by $|\Delta V| = 100$, 200, or 400 km s^{-1}, respectively. Symbols show the corresponding mean decrements, saturated fractions, and transparent fractions.

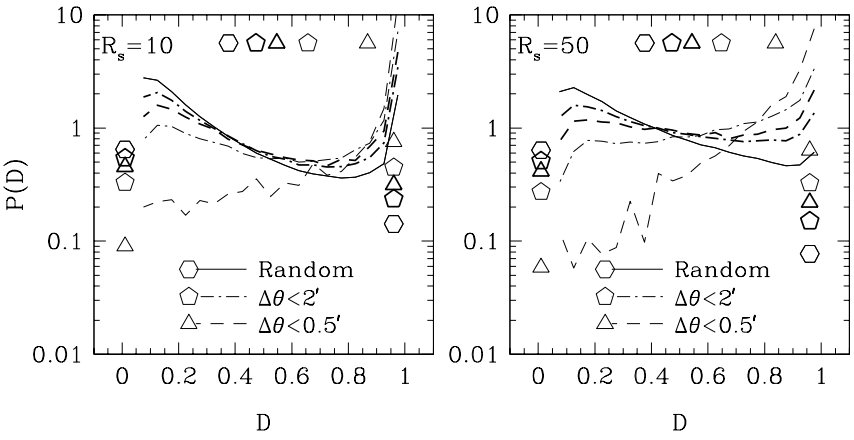

Fig. 6.— The impact of galaxy bias on the conditional flux PDF, for spectral smoothing lengths of 10 km s^{-1} (left) and 50 km s^{-1} (right). Light solid, dot-dashed, and dashed curves, repeated from the corresponding panels of Fig. 4, show the unconditional flux PDF and the conditional flux PDFs for $\Delta\theta < 2'$ and $\Delta\theta < 0.5'$ about resolved galaxies in the L22n128 simulation. Heavy curves show the corresponding conditional flux PDFs using randomly selected dark matter particles in place of galaxies. Light and heavy symbols are based on the corresponding flux PDFs. While the conditional flux PDF about randomly selected mass elements differs from the unconditional flux PDF (which is effectively volume weighted), the deviations of the galaxy conditional PDFs are much stronger, a clear signature of the preferential formation of high redshift galaxies in overdense large scale environments.

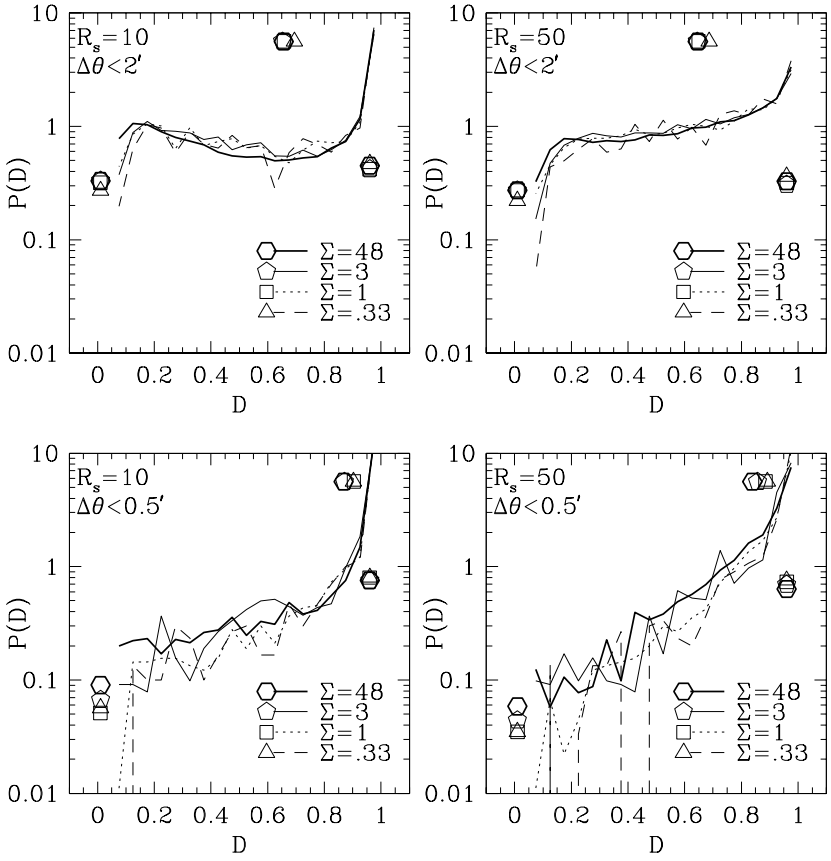

Fig. 7.— Dependence of the conditional flux PDF on galaxy mass and SFR. In each panel, heavy solid lines show the conditional PDF based on all resolved galaxies in the L22n128 simulation, as in Fig. 4. The baryonic mass threshold is $M \geq 6.7 \times 10^9 M_\odot$, and the space density of the population is $\Sigma = 48\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$. Light solid lines show results for the resolved galaxy population of L50n144, with a mass threshold of $5.4 \times 10^{10} M_\odot$ and a space density of $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$. Dotted and dashed lines are for galaxies in L50n144 with SFR $> 108 M_\odot \mathrm{yr}^{-1}$ and SFR $> 180 M_\odot \mathrm{yr}^{-1}$, respectively, with population space densities $\Sigma = 1\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ and $\Sigma = 0.33\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$.

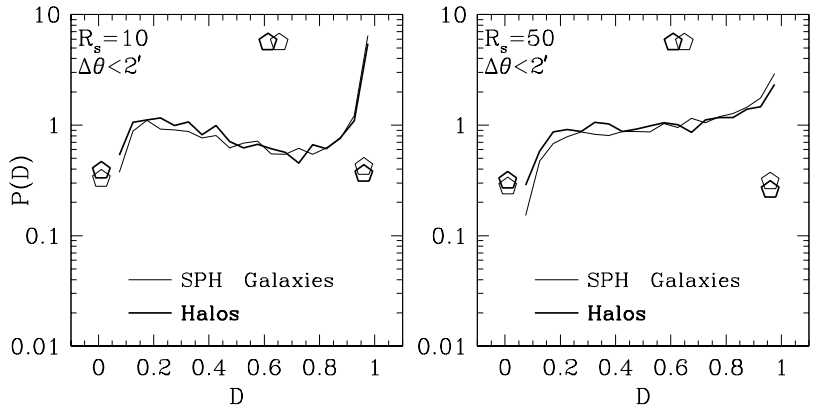

Fig. 8.— Comparison of halo and galaxy conditional flux PDFs. Light lines and symbols, repeated from the corresponding panels of Fig. 7, show results for the resolved galaxies in L50n144, with $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$. Heavy lines and symbols show results when these galaxies are replaced by friends-of-friends halos identified from the simulation's dark matter particle distribution. Halos are selected above a (dark matter) mass threshold of $5.7 \times 10^{11} M_\odot$, yielding 450 halos in the simulation volume and $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$.

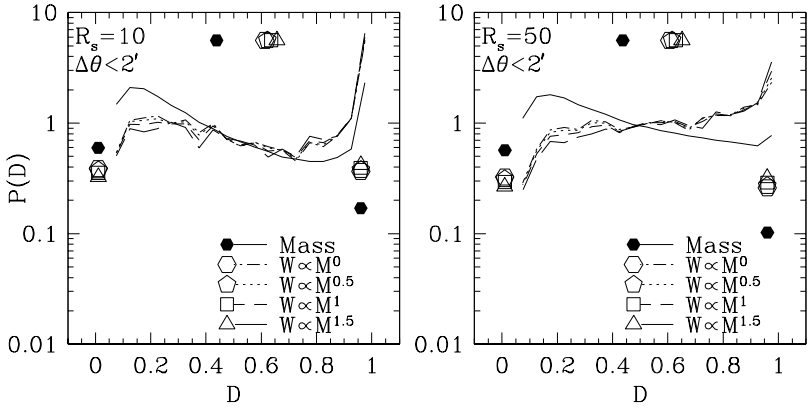

Fig. 9.— Influence of halo weighting on the conditional flux PDF. Dot-dashed lines, repeated from Fig. 8, show results in which each halo above the mass threshold contributes equally. Solid lines and filled symbols show results for randomly selected dark matter particles. Other lines show results in which the contribution of each halo-pixel pair to the conditional PDF is weighted by M^α, with $\alpha = 0.5$ (dotted), $\alpha = 1$ (short-dashed), or $\alpha = 1.5$ (long-dashed). While the overall bias of halos makes the conditional PDF of halos substantially different from that of the mass, stronger weighting of high mass halos has only a small impact on this statistic.

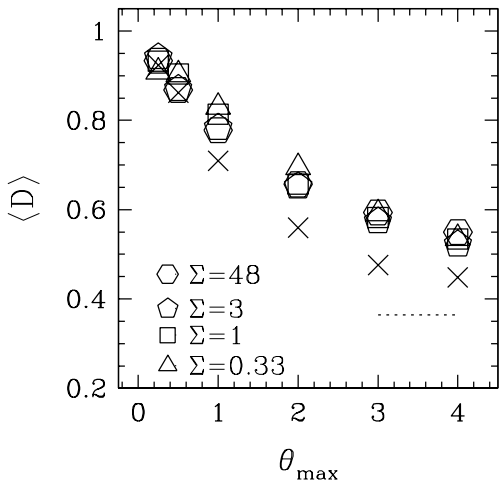

Fig. 10.— The conditional mean flux decrement, for spectral smoothing of 10 km s^{-1}. Open symbols show the mean decrement of pixels at the redshift of a galaxy with angular separation $\Delta\theta < \theta_{\rm max}$ arc-minutes. Results are shown for the resolved galaxy population of L22n128 ($\Sigma = 48\,{\rm arcmin}^{-2}\,\Delta z^{-1}$), the resolved galaxy population of L50n144 ($\Sigma = 3\,{\rm arcmin}^{-2}\,\Delta z^{-1}$), and SFR-thresholded samples of the L50n144 population with $\Sigma = 1\,{\rm arcmin}^{-2}\,\Delta z^{-1}$ and $\Sigma = 0.33\,{\rm arcmin}^{-2}\,\Delta z^{-1}$. The $\times$'s show a "differential" form of this statistic, with $\langle D\rangle$ computed for separations in the range $0.8\theta_{\rm max} \le \Delta\theta < 1.25\theta_{\rm max}$, for the $\Sigma = 3\,{\rm arcmin}^{-2}\,\Delta z^{-1}$ case only. The horizontal dotted line segment marks the unconditional mean decrement, $\langle D\rangle = 0.36$. For our cosmology, $1'$ corresponds to $1.19\,h^{-1}$ Mpc (comoving) and $Hr_t = 152$ km s^{-1} at $z = 3$.

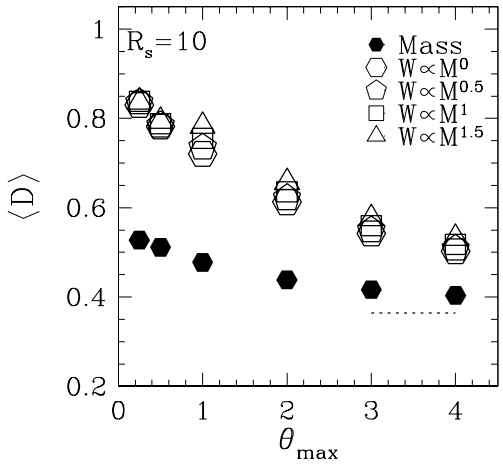

Fig. 11.— The conditional mean flux decrement (cumulative form) for randomly selected dark matter particles (filled symbols) and for dark matter halos with different relative weightings (open symbols, with weights as marked in the legend), computed from the L50n144 simulation. The dotted horizontal line segment shows the unconditional mean decrement $\langle D \rangle = 0.36$.

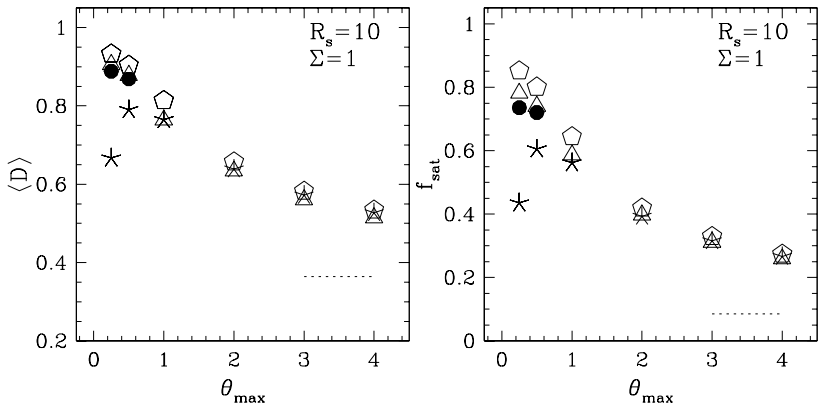

Fig. 12.— Potential impact of galaxy photoionization on the conditional mean flux decrement (left) and the saturated pixel fraction (right), for the $\Sigma = 1\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ sample of the L50n144 simulation and a spectral smoothing of $10\,\mathrm{km\ s}^{-1}$. We assume that these galaxies produce a fraction $f = 0.5$ of the UV background at $z = 3$, with each galaxy's photoionizing flux proportional to its SFR. Pentagons show results in the absence of galaxy photoionization, asterisks show the results of the approximate calculation described in the text, and triangles show the results of a complete calculation with the UV sources embedded in the simulation before spectra are extracted. The complete calculation yields a much weaker effect because peculiar velocities allow relatively distant gas to produce absorption at the galaxy redshift (see Fig. 13). Filled circles show results for the recurrent AGN model described in the text, for $\theta_{\rm max} = 0.25'$ and $\theta_{\rm max} = 0.5'$ only. Dotted horizontal segments show the mean decrement $\langle D \rangle = 0.36$ or saturated fraction $f_{\rm sat} = 0.086$ for the unconditional flux PDF.

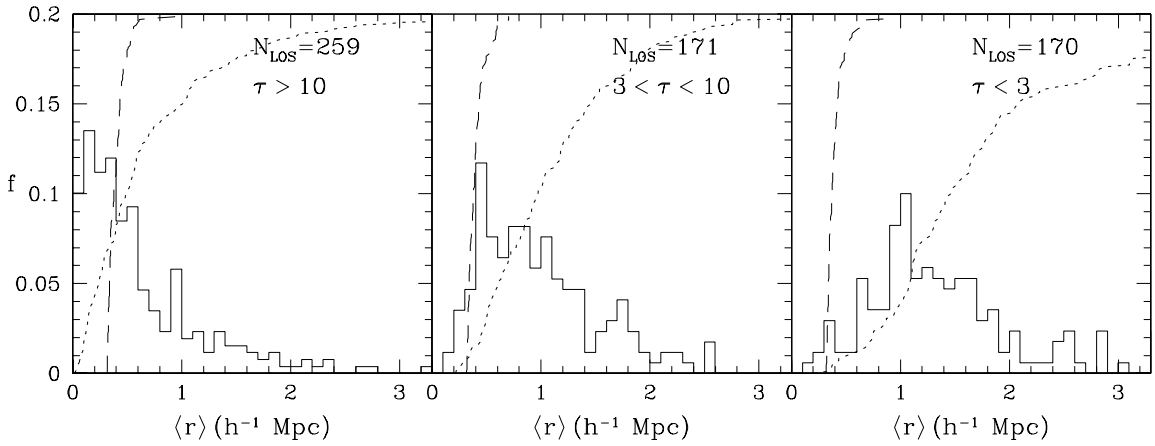

Fig. 13.— Distribution of the mean line-of-sight distance, weighted by optical depth, of gas producing absorption within 20 km s^{-1} of the galaxy redshift, calculated for lines of sight within $0.5'$ of galaxies in the $\Sigma = 1\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ sample of the L50n144 simulation, assuming a uniform ionizing background. Left, middle, and right panels show three different ranges of the optical depth at the galaxy redshift, with the number of lines of sight as indicated. In each panel, histograms show the differential distribution and dotted lines the cumulative distribution (on a scale where the top of the panel is unity). Dashed lines show the cumulative distribution of photoionization influence radii r_i for the 150 galaxies. Units are comoving h^{-1} Mpc.

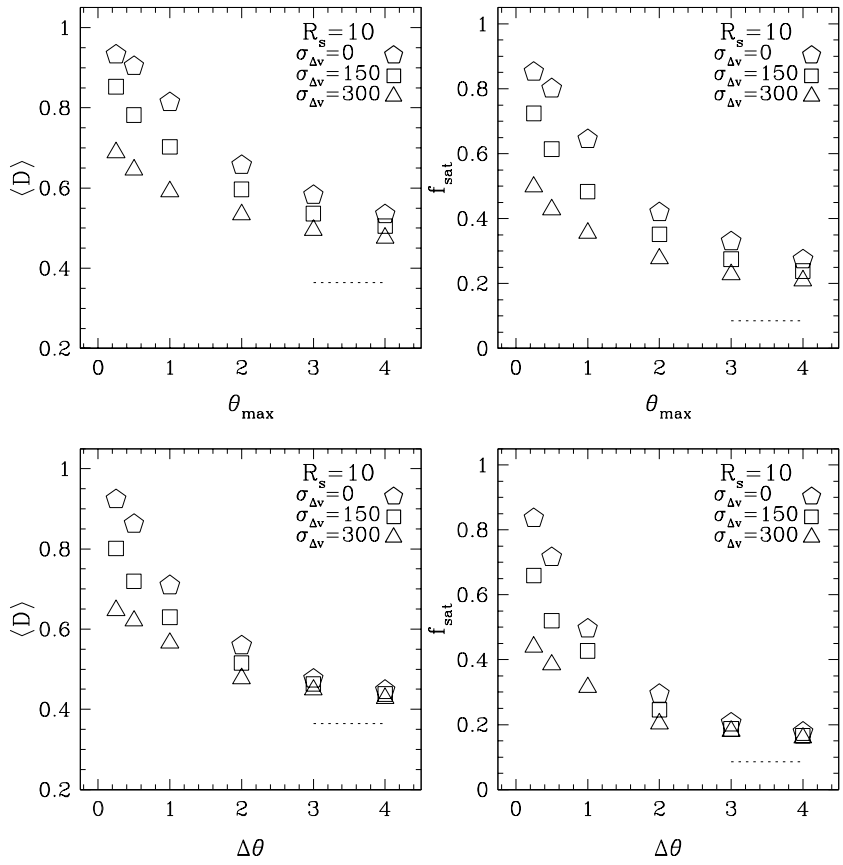

Fig. 14.— Impact of redshift determination errors on the conditional mean decrement (left) and saturated pixel fraction (right). Upper panels show our standard, cumulative form of these statistics, computed for all galaxy-pixel pairs with angular separation $\Delta\theta < \theta_{\rm max}$, and lower panels show the differential form, computed for pairs with $0.8\theta_{\rm max} < \Delta\theta < 1.25\theta_{\rm max}$. In each panel, pentagons show results for the $\Sigma = 1\,{\rm arcmin}^{-2}\,\Delta z^{-1}$ sample of the L50n144 galaxy population in the absence of redshift errors. Squares and triangles show the corresponding results when estimated galaxy redshifts are drawn from a Gaussian distribution of 1σ dispersion $150\,{\rm km\ s^{-1}}$ or $300\,{\rm km\ s^{-1}}$, respectively, centered on the true galaxy redshifts.

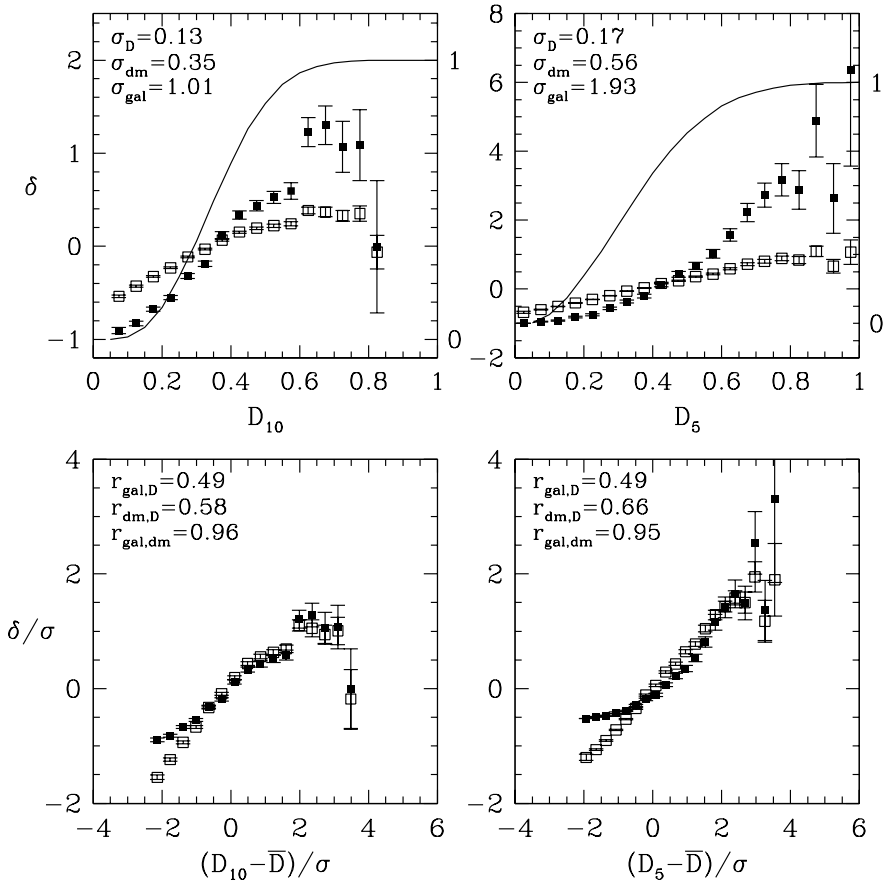

Fig. 15.— Large scale correlations of galaxy and dark matter overdensity with Lyα flux decrement, in cubes of comoving size $10\,h^{-1}$ Mpc (left) and $5\,h^{-1}$ Mpc (right). The mean decrement along a segment of Lyα forest spectrum traversing the center of the cube is denoted D_{10} or D_5, and filled (open) squares in the top panel show the mean value of the galaxy (dark matter) density contrast in cubes with a specified value of D_{10} or D_5. Plots in the lower panels are normalized by subtracting the mean and dividing by the dispersion of the corresponding variable. Curves in the upper panels show the cumulative distribution of D_{10} and D_5, with values zero and one marked $\bar{\text{o}}$n the right axis. Upper panels list dispersions of the decrement, dark matter contrast, and galaxy contrast (with shot noise subtracted), and lower panels list cross-correlation coefficients. All error bars show 1σ uncertainty in the mean; the dispersion of δ values in a given bin of D_{10} or D_5 is much larger.

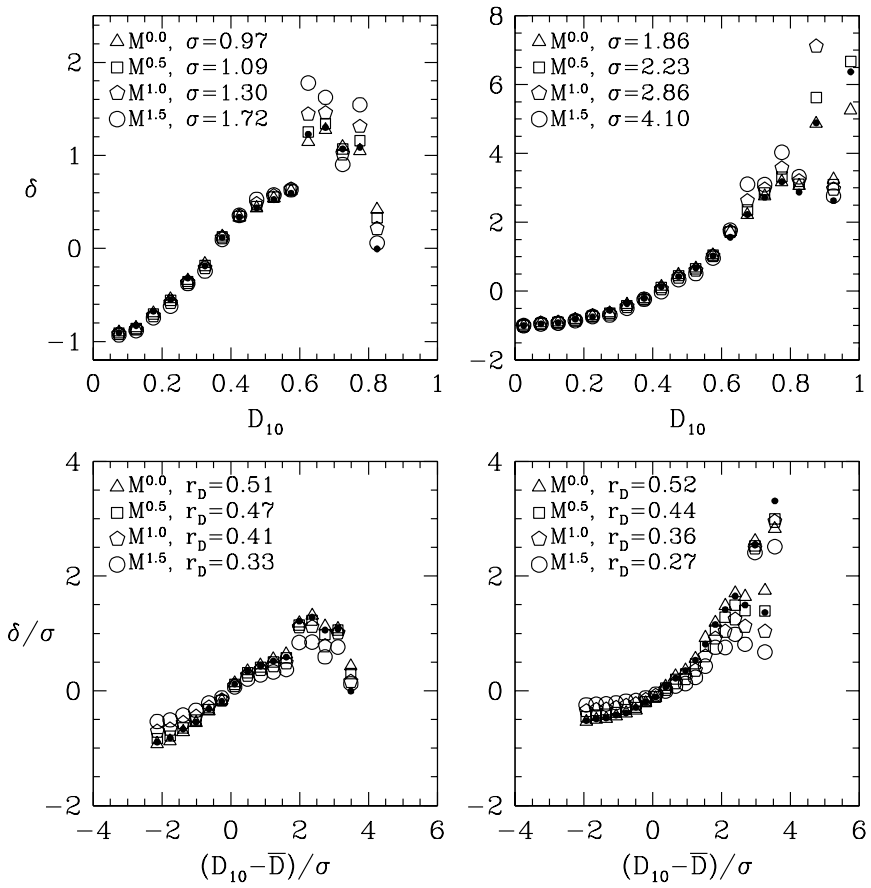

Fig. 16.— Similar to Fig. 15, but for dark matter halos. Triangles show the case in which each of the 450 most massive halos contributes equally. Squares, pentagons, and circles show mass-weighted cases, with the contribution of each halo proportional to $M^{0.5}$, M, and $M^{1.5}$, respectively. Solid dots show the results for galaxies, repeated from Fig. 15.

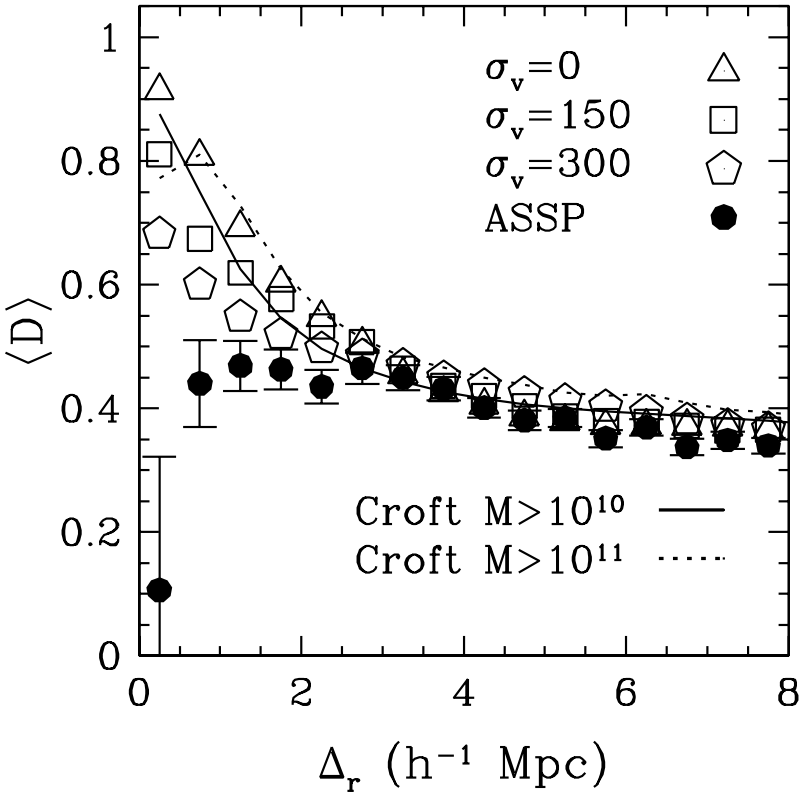

Fig. 17.— The conditional mean flux decrement, now defined by an average over all galaxy-pixel pairs in bins of comoving redshift-space separation. Open symbols show our results for the $\Sigma = 1\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ sample of the L50n144 simulation, with spectral smoothing length $R_s = 10\,\mathrm{km\ s}^{-1}$ and no redshift errors (triangles), rms redshift errors of $150\,\mathrm{km\ s}^{-1}$ (squares) and $300\,\mathrm{km\ s}^{-1}$ (pentagons). Solid and dotted curves show Croft et al.'s (2002a) results for galaxies with baryonic mass $M_b > 10^{10} M_\odot$ and $M_b > 10^{11} M_\odot$, respectively. Filled symbols show the ASSP observational results, with approximate error bars estimated by a Monte Carlo procedure in which galaxy redshifts are perturbed by $\sim 50\,h^{-1}$ Mpc (comoving) to produce a "random" sample with the same transverse separations.

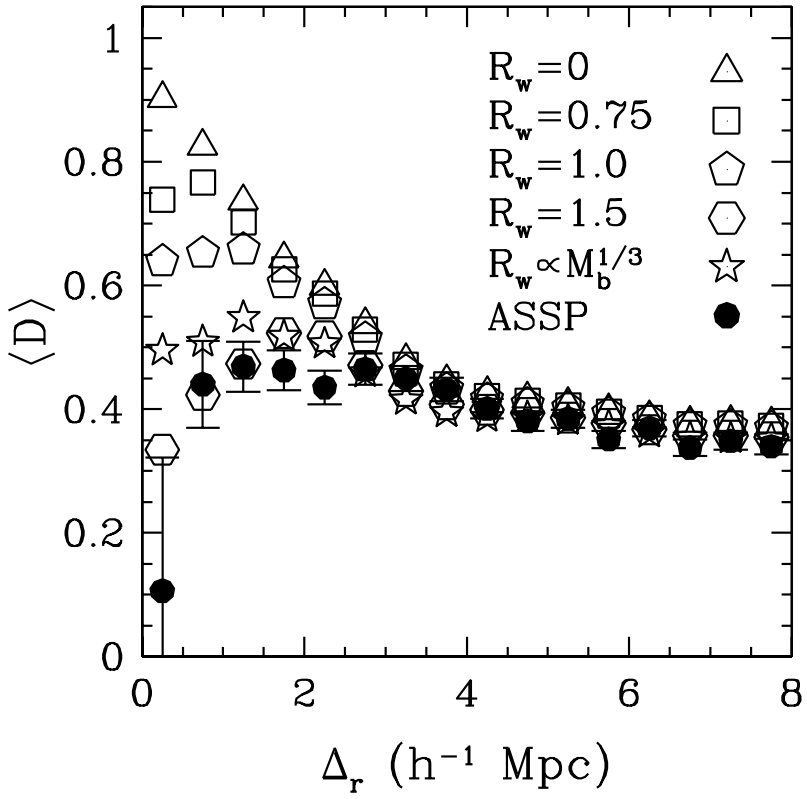

Fig. 18.— Effects of simplified "winds" on the mean flux decrement. Open symbols show our results for the $\Sigma = 3\,\mathrm{arcmin}^{-2}\,\Delta z^{-1}$ sample of the L22n128 simulation, with spectral smoothing length $R_s = 10\,\mathrm{km\ s}^{-1}$ and wind radii of 0.75 (squares), 1.0 (pentagons), 1.5 h^{-1} Mpc (hexagons) and proportional to $M_b^{1/3}$ (stars). Open triangles show our results for no winds, and filled symbols show the ASSP observational results.

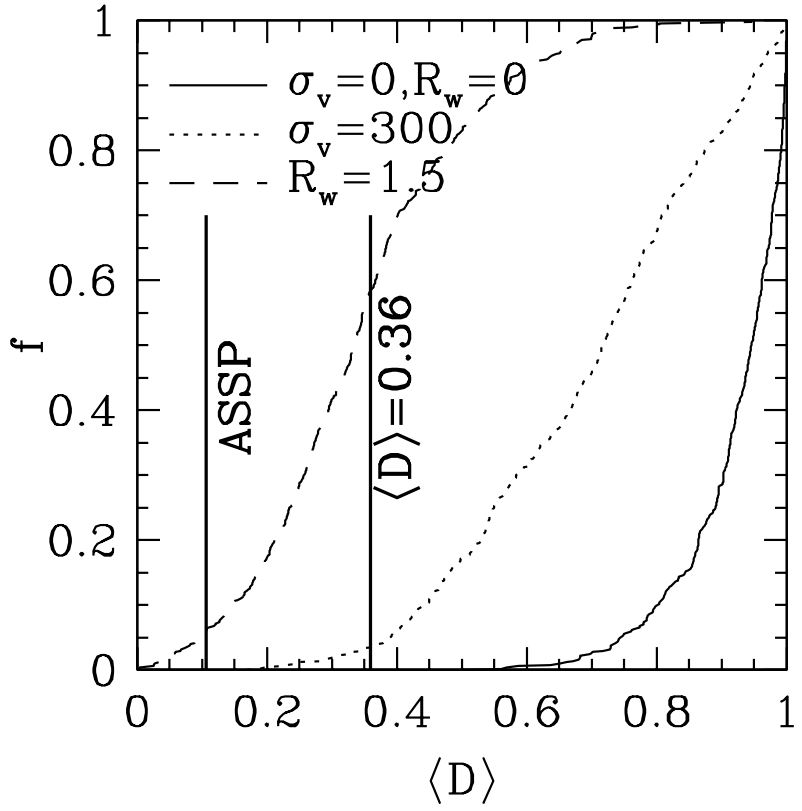

Fig. 19.— Statistical significance of the discrepancy with the innermost ASSP data point, which comes from three galaxy-QSO pairs. Vertical lines mark the ASSP measured value for $\Delta_r < 0.5\,h^{-1}$ Mpc, $\langle D \rangle = 0.11$, and the unconditional mean flux decrement, $\langle D \rangle = 0.36$. For each model, we randomly select 500 3-tuples of close galaxy-spectrum pairs and compute the mean decrement for pixels with $\Delta_r < 0.5\,h^{-1}$ Mpc. Curves show the cumulative distribution of these mean decrement values, for our standard model with no redshift errors (solid, corresponding to the triangles in Fig. 17), the standard model with 300 km s^{-1} rms redshift errors (dotted, corresponding to pentagons in Fig. 17), and the $R_w = 1.5\,h^{-1}$ Mpc wind model with no redshift errors (dashed, corresponding to the hexagons in Fig. 18).