THE NATURE OF DAMPED Lyα SYSTEMS AND THEIR HOSTS IN THE STANDARD COLD DARK MATTER UNIVERSE

RENYUE CEN
Department of Astrophysical Sciences, Princeton University Observatory, Princeton, NJ 08544, USA; cen@astro.princeton.edu
Received 2011 February 28; accepted 2012 January 23; published 2012 March 14

ABSTRACT
Using adaptive mesh refinement cosmological hydrodynamic simulations with a physically motivated supernova feedback prescription, we show that the standard cold dark matter model can account for extant observed properties of damped Lyα systems (DLAs). With detailed examination of DLAs identified for each redshift snapshot through ray tracing through the simulation volumes containing thousands of galaxies, we find the following: (1) While DLA hosts roughly trace the overall population of galaxies at all redshifts, they are always gas-rich and have tendencies of being slightly smaller and bluer. (2) The history of DLA evolution is cosmological in nature and reflects primarily the evolution of the underlying cosmic density, galaxy size, and galaxy interactions. With higher density and more interactions at high redshift the size of DLAs is a larger fraction of their virial radius. (3) The variety of DLAs at high redshift is richer with a large contribution coming from galactic aqueducts, created through close galaxy interactions. The portion of gaseous disks of galaxies where most stars reside makes a relatively small contribution to DLA incidence at $z = 3–4$. (4) The majority of DLAs arise in halos of mass $M_h = 10^{10}–10^{12} \, M_\odot$ at $z = 1.6–4$, as these galaxies dominate the overall population of galaxies then. At $z = 3–4$, 20%–30% of DLA hosts are Lyman break galaxies (LBGs), 10%–20% are due to galaxies more massive than LBGs, and 50%–70% are from smaller galaxies. (5) Galactic winds play an indispensible role in shaping the kinematic properties of DLAs. Specifically, the high velocity width DLAs are a mixture of those arising in high-mass, high velocity dispersion halos and those arising in smaller mass systems where cold gas clouds are entrained to high velocities by galactic winds. (6) In agreement with observations, we see a weak but noticeable evolution in DLA metallicity. The metallicity distribution centers at $[Z/H] = -1.5$ to $-1$ and spans more than three decades at $z = 3–4$, with the peak moving to $[Z/H] = -0.75$ at $z = 1.6$ and $[Z/H] = -0.5$ by $z = 0$. (7) The star formation rate of DLA hosts is concentrated in the range $0.3–30 \, M_\odot \, yr^{-1}$ at $z = 3–4$, gradually shifting lower to peak at $\sim 0.5–1 \, M_\odot \, yr^{-1}$ by $z = 0$. (8) We predict that only 20%–30% of DLAs are within 100 kpc of $L^*$ galaxies at $z = 3–4$, increasing to 30%–50% at $z = 1.6$.

Key words: galaxies: evolution – galaxies: interactions – intergalactic medium – ISM: kinematics and dynamics – methods: numerical

Online-only material: color figures

1. INTRODUCTION

Damped Lyα systems (DLAs) are fundamentally important, because they contain most of the neutral gas in the universe at all times since cosmological reionization (e.g., Storrie-Lombardi & Wolfe 2000; Péroux et al. 2003; Prochaska & Wolfe 2009). Molecular clouds, within which star formation takes place, condense out of cold neutral atomic gas contained in DLAs, evidenced by the fact that the neutral hydrogen (surface) density in DLAs and molecular hydrogen (surface) density in molecular clouds form a continuous sequence (e.g., Kennicutt 1998; Zwaan & Prochaska 2006). Therefore, DLAs are key to understanding the fuel for and ultimately galaxy formation. A substantial amount of theoretical work has been devoted to studying the nature of DLAs (e.g., Gardner et al. 1997a, 1997b, 2001; Haehnelt et al. 1998; Maller et al. 2001; Cen et al. 2003; Nagamine et al. 2004a, 2004b, 2007; Razoumov et al. 2006, 2008; Pontzen et al. 2008; Tescari et al. 2009; Hong et al. 2010; Fumagalli et al. 2011) since the pioneering investigation of Katz (1996) in the context of the cold dark matter (CDM) cosmogony. A very interesting contrast is drawn between the observationally based inference of simple large disk galaxies giving rise to DLAs (Wolfe et al. 1986; Prochaska & Wolfe 1997) and the more naturally expected hierarchical buildup of structures in the CDM cosmogony where galactic subunits may produce some of the observed kinematics of DLAs (Haehnelt et al. 1998). Clearly, the implications on the evolution of galaxies in the two scenarios are very different. The former would require that large galactic disks of radius $\sim 30$ kpc are already in place at redshift $z \sim 3$ (e.g., Schaye 2001), whereas the latter would have important implications on the gas content of galactic subunits and their star formation history.

We have carried out a set of Eulerian adaptive mesh refinement (AMR) simulations with a resolution of 0.65 kpc proper and a sample size of several thousand galaxies with halo mass $\geq 10^{10} \, M_\odot$ to statistically address the physical nature of DLAs in the current standard cosmological constant-dominated CDM model (LCDM) (Komatsu et al. 2011). Mechanical feedback from star formation driven by supernova explosions and stellar winds is modeled by a one-parameter prescription that is physically and energetically sound. Part of the motivation was to complement and cross-check studies to date that are largely based on smoothed particle hydrodynamics (SPH) simulations (e.g., Gardner et al. 1997a, 1997b, 2001; Haehnelt et al. 1998; Nagamine et al. 2004a, 2004b, 2007; Pontzen et al. 2008; Tescari et al. 2009; Hong et al. 2010). With the simulation set and detailed analysis performed here this study represents a significant extension of previous works to simultaneously subject the LCDM model to a wider and more complete range of comparisons with observations. We examine in detail the following properties of DLAs in a self-consistent fashion within the same model: DLA column density distribution evolution, line
density evolution, metallicity distribution evolution, size distribution evolution, velocity width distribution evolution, kinematic structural parameter evolution, neutral mass content evolution, and others. A gallery of DLAs is presented to obtain a visual understanding of the physical richness of DLA systems, especially the effects of galactic winds and large-scale gaseous structures. In comparison to the recent work of Hong et al. (2010), we track the metallicity distribution and evolution explicitly and show that the computed metallicity distribution of DLAs is, in good agreement with observations, very wide, which itself calls for a self-consistent treatment of metal transport. In agreement with the conclusions of Hong et al. (2010), although not in the detailed process, we show that galactic winds are directly responsible for a fraction of wide DLAs at high redshift, by entraining cold clouds to large velocities and causing large kinematic velocity widths for some fraction of DLAs. We find that the simulated Si II $\lambda$1808 line velocity width, kinematic shape measures, and DLA metallicity distributions are all in excellent agreement with observations. Taken all together, we conclude that the standard LCDM model gives a satisfactory account of all properties of DLAs. Finally, we examine the properties of DLA hosts, including their halo mass, star formation rate (SFR), H I content, gas-to-stellar mass ratio, and colors, and show that DLAs arise in a variety of galaxies and roughly trace the entire population of galaxies at any redshift. This may reconcile many apparently conflicting observational evidences.

2. SIMULATIONS

2.1. Hydrocode and Simulation Parameters

We perform cosmological simulations with the AMR Eul- erian hydro code, Enzo (Bryan 1999; Bryan & Norman 2000; O’Shea et al. 2005; Joung et al. 2009). First, we ran a low-resolution simulation with a periodic box of $120 h^{-1}$ Mpc on a side. We identified two regions separately, one centered on a cluster of total mass of $\sim 3 \times 10^{14} M_{\odot}$ and the other centered on a void region at $z = 0$. We then resimulate each of the two regions separately with high resolution, but embedded in the outer $120 h^{-1}$ Mpc box to properly take into account large-scale tidal field and appropriate boundary conditions at the surface of the refined region. We name the simulation centered on the cluster the “C” run and the one centered on the void the “V” run. The refined region for the “C” run has a size of $21 \times 24 \times 20 h^{-3}$ Mpc$^3$ and that for the “V” run is $31 \times 31 \times 35 h^{-3}$ Mpc$^3$. At their respective volumes, they represent 1.8$\sigma$ and 0.7$\sigma$ fluctuations. The initial condition in the refined region has a mean interparticle separation of $117 h^{-1}$ kpc comoving and a dark matter particle mass of $1.07 \times 10^8 h^{-1} M_{\odot}$. The refined region is surrounded by two layers (each of $\sim 1 h^{-1}$ Mpc) of buffer zones with particle masses successively larger by a factor of eight for each layer, which then connects with the outer root grid that has a dark matter particle mass 8 times that in the refined region. Because we still cannot run a very large volume simulation with adequate resolution and physics, we choose these two runs to represent two opposite environments that should bracket the average. At redshift $z > 1.6$, as we will show, the average properties of most quantities concerning DLAs in the “C” and “V” runs are not very different, although the abundances of DLAs in the two runs are already very different. It is only at lower redshift where we see significant divergence of some quantities of DLAs between the two runs, due to different dynamic evolutions of regions contained in the two runs.

We choose the mesh refinement criterion such that the resolution is always better than $460 h^{-1}$ pc physical, corresponding to a maximum mesh refinement level of 11 at $z = 0$. We also run an additional simulation for the “C” run with a factor of two lower resolution to assess the convergence of the results, which we name the “C/2” run, and as we will show in the Appendix, the convergence is excellent for all quantities examined here. The simulations include a metagalactic UV background (Haardt & Madau 1996) and a model for shielding of UV radiation by neutral hydrogen (Cen et al. 2005). They also include metallicity-dependent radiative cooling (Cen et al. 1995). Star particles are created in cells that satisfy a set of criteria for star formation proposed by Cen & Ostriker (1992), very similar to the Schmidt–Kennicutt law and similar to that used by other investigators (Katz 1992; Katz et al. 1996; Steinmetz 1996; Gnedin & Ostriker 1997; Kravtsov 2003; Arielie et al. 2008). A star particle of mass $m_\star = c_\star m_{gas}\Delta t/\tau_\star$ is created (the same amount is removed from the gas mass in the cell), if the gas in a cell at any time meets the following three conditions simultaneously: (1) contracting flow, (2) cooling time less than dynamic time, and (3) Jeans unstable, where $\Delta t$ is the time step, $\tau_\star = \max(t_{dyn}, 10^7 \text{yr})$, $t_{dyn} = \sqrt{3\pi/2} (2\Sigma_0 \rho_{gas})$ is the dynamical time of the cell, $m_{gas}$ is the baryonic gas mass in the cell, and $c_\star \sim 0.03$ is star formation efficiency (e.g., Krumholz & Tan 2007). Each star particle is tagged with its initial mass,
creation time, and metallicity; star particles typically have masses of $\sim 10^6 M_\odot$.

Supernova feedback from star formation is modeled following Cen et al. (2005). Feedback energy and ejected metal-enriched mass are distributed into 27 local gas cells centered at the star particle in question, weighted by the specific volume of each cell, which is to mimic the physical process of supernova blastwave propagation that tends to channel energy, momentum, and mass into the least dense regions (with the least resistance and cooling). We allow the whole feedback process to be hydrodynamically coupled to surroundings and subject to relevant physical processes, such as cooling and heating, as in nature. As we will show later, the extremely inhomogeneous metal enrichment process demands that both metals and energy (and momentum) are correctly modeled so that they are transported into the right directions in a physically sound (albeit still approximate at the current resolution) way. The primary advantages of this supernova energy-based feedback mechanism are threefold. First, nature does drive winds in this way, and energy input is realistic. Second, it has only one free parameter $e_{\text{SN}}$, namely, the fraction of the rest mass energy of stars formed that is deposited as thermal energy on the cell scale at the location of supernovae. Third, the processes are treated physically, obeying their respective conservation laws (where they apply), allowing transport of metals, mass, energy, and momentum to be treated self-consistently and taking into account relevant heating/cooling processes at all times. We use $e_{\text{SN}} = 1 \times 10^{-5}$ in these simulations. The total amount of explosion kinetic energy from Type II supernovae with a Chabrier initial mass function (IMF) translates to $e_{\text{SN}} = 6.6 \times 10^{-6}$. Observations of local starburst galaxies indicate that nearly all of the star formation produced kinetic energy (due to Type II supernovae) is used to power galactic superwinds (e.g., Heckman 2001). Given the uncertainties on the evolution of IMF with redshift (i.e., more top-heavy at higher redshift) and the fact that newly discovered prompt Type I supernovae contribute a comparable amount of energy compared to Type II supernovae, our adopted value for $e_{\text{SN}}$ is consistent with observations and within physical plausibility.

We use the following cosmological parameters that are consistent with the WMAP7-normalized (Komatsu et al. 2011) LCDM model: $\Omega_M = 0.28$, $\Omega_b = 0.046$, $\Omega_L = 0.72$, $\sigma_8 = 0.82$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1} = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and $n = 0.96$.

A convergence test of results is presented separately in the Appendix in order not to disrupt the flow of the presentation in the results section. The tests show that our results are quite converged and should be robust at the accuracies concerned here, suggesting that our resolution has reached an adequate level for the present study. The reader may go to the Appendix any time to gauge the convergence of relevant computed quantities. Given the fact that most of the contributions to DLA incidence come from galaxies of halo mass $\sim 10^{13} M_\odot$ that are well above our resolution, the results of our convergence tests are self-consistent.

2.2. Simulated Galaxy Catalogs

We identify galaxies in our high-resolution simulations using the HOP algorithm (Eisenstein & Hu 1999), operated on the stellar particles, which is tested to be robust and insensitive to specific choices of concerned parameters within reasonable ranges. Satellites within a galaxy are clearly identified separately. The luminosity of each stellar particle at each of the five Sloan Digital Sky Survey (SDSS) bands is computed using the GISSEL stellar synthesis code (Bruzual & Charlot 2003), by supplying the formation time, metallicity, and stellar mass. Collecting luminosity and other quantities of member stellar particles, gas cells, and dark matter particles yields the following physical parameters for each galaxy: position, velocity, total mass, stellar mass, gas mass, mean formation time, mean stellar metallicity, mean gas metallicity, SFR, luminosities in five SDSS bands (and various colors), and others.

2.3. Simulated Damped Ly$\alpha$ System Samples

While our simulations also solve relevant gas chemistry chains for molecular hydrogen formation (Abel et al. 1997), molecular formation on dust grains (Joung et al. 2009), and metal cooling extended down to 10 K (Dalgarno & McCray 1972), at the resolution of the simulations, molecular clouds are not properly modeled. To correct for that, we use the Hidaka & Sofue (2002) observation that at $n_e = 5 \text{ H}_i \text{ cm}^{-3}$ the $H_2$ fraction is about 50% and then implement the following prescription to remove neutral gas in extrapolated high-density regions and put it in $H_2$ phase. In detail, we assume that the density profile is isothermal below our resolution which would translate the fraction of mass in $H_2$ is $\min(1, 0.5(n_e/n_{\text{H}_i})^{-1/2})$. Thus, we post-process the neutral density in the simulation by the following transformation: $n_{\text{H}_i}(\text{after}) = n_{\text{H}_i}(\text{before})(1 - \min(1, 0.5(n_e/n_{\text{H}_i}(\text{before}))^{-1/2}))$, where $n_{\text{H}_i}(\text{before})$ is the $H_1$ density directly from the simulation and $n_{\text{H}_i}(\text{after})$ is that after this processing step. A very precise choice of the parameter in the above equation is unimportant; changing 0.5 to 1.0 makes marginally noticeable differences in the results. The primary effect of doing this is to remove very high $H_1$ column DLas, which causes the $H_1$ column density distribution function to steepen at $N_{\text{H}_i} > 22.5$, in agreement with observations. In addition, because of that, the total amount of neutral gas in DLas also becomes convergent and stable.

After the above post-process step, we shoot rays through the entire refined region of each simulation along all three orthogonal directions using a cell size of 0.915 $h^{-1}$ kpc comoving. In practice, this is done piecewise, one small volume of the simulation box at a time, due to limited computer memory. The spectral bin size is 3 km $s^{-1}$. All physical effects are taken into account, including temperature broadening and peculiar velocities. Both intrinsic Lorentzian line profiles and Doppler broadening are taken into account for both Ly$\alpha$ lines, although, in practice, for DLas, Doppler broadening is important for the Si$\text{\upalpha}$ $\lambda 1808$ line. All relevant atomic data are taken from Morton (2003). A DLA is defined, as usual, as a system with $H_1$ column larger than $10^{20.3} \text{ cm}^{-2}$. We assume that the fractional abundance of Si$\text{\upalpha}$ is equal to fractional abundance of $H_1$. Since, as we will see later, the $H_1$ regions of DLas are “peaky” with well-defined line-of-sight (LOS) boundaries and since DLas are very optically opaque to ionizing photons, any refined treatment of radiative self-shielding, etc., is unlikely to have any significant effect. Note that we have already included a crude self-shielding method during the simulation, which should work well for optically opaque regions. As an aside, one numerical point to note is that, because of the very large dynamic range of both line cross sections as a function of frequency shift from the line center and the delta function–like cross section shapes in the line core regions, the convolution operations involved in the detailed calculations of optical depths require at least 64-bit precision for floating point numbers.
For each DLA, we compute the H\textsc{i} column weighted metallicity, register its position relative to the center of the primary galaxy (i.e., the impact parameter), and for DLAs that are physically connected by at least one cell side in projection we merge them and in the end compute projected area A of each connected region to define its size $r_{\text{DLA}} = (A/\pi)^{1/2}$. For each galaxy we also register the maximum velocity width $v_{90,\text{max}}$ among its associated DLAs.

While we are able to identify many DLAs through ray tracing at each redshift examined with very small statistical errors, it does not speak to cosmic variance, and as we shall show later, cosmic variance is indeed quite large for quantities that directly or indirectly pertain to the number density of DLAs. Other quantities, such as size, metallicity, kinematic properties, etc., however, depend weakly on environments, and their variances are small. A DLA “belongs” to the largest galaxy in the region, within whose virial radius the DLA lies. For example, a DLA that is physically more closely located to a satellite galaxy that in turn is within the virial radius of a larger galaxy is said to belong to that larger galaxy.

### 2.4. Kinematic Measures for Si \textsc{ii} Line

We do not add instrumental noise to the simulated spectra, but we adopt the same observational procedure to compute the kinematic measures for the Si \textsc{ii} absorption lines. For all relevant measures for the Si \textsc{ii} line, we follow identically the procedures and definitions in Prochaska & Wolfe (1997). We generate synthetic spectra for both Ly\textsc{a} and Si \textsc{ii} lines with 3 km s$^{-1}$ pixels and then smooth them with a 9 pixel boxcar averaging procedure. We define the velocity width of an Si \textsc{ii} absorption line associated with a DLA to be the velocity interval of 90\% of the total optical depth, $v_{90}$. For the three kinematic shape measures for the Si \textsc{ii} line we use all intensity troughs (optical depth peaks) without the 0\% trough and $I_{\text{pk}}$, and then smooth them with a 9 pixel boxcar averaging procedure. We define the velocity width of an Si \textsc{ii} absorption line associated with a DLA to be the velocity interval of 90\% of the total optical depth, $v_{90}$. For the three kinematic shape measures for the Si \textsc{ii} line we use all intensity troughs (optical depth peaks) without the 0\% trough and $I_{\text{pk}}$, and then smooth them with a 9 pixel boxcar averaging procedure. We define the velocity width of an Si \textsc{ii} absorption line associated with a DLA to be the velocity interval of 90\% of the total optical depth, $v_{90}$. For the three kinematic shape measures for the Si \textsc{ii} line we use all intensity troughs (optical depth peaks) without the 0\% trough and $I_{\text{pk}}$, and then smooth them with a 9 pixel boxcar averaging procedure. We define the velocity width of an Si \textsc{ii} absorption line associated with a DLA to be the velocity interval of 90\% of the total optical depth, $v_{90}$.

$$f_{\text{nm}} = \frac{|v_{\text{mean}} - v_{\text{median}}|}{(v_{90}/2)}, \quad f_{\text{edg}} = \frac{|v_{\text{pk}} - v_{\text{median}}|}{(v_{90}/2)},$$

$$f_{2\text{pk}} = \pm \frac{|v_{2\text{pk}} - v_{\text{median}}|}{(v_{90}/2)},$$

where $v_{\text{mean}}$, $v_{\text{median}}$, $v_{\text{pk}}$, and $v_{2\text{pk}}$ are the velocity locations of the mean, median, strongest peak, and second strongest peak, respectively. In Equation (1) the plus sign holds if the second peak is between the velocity of the first (strongest) peak, $v_{\text{pk}}$, and $v_{\text{mean}}$; otherwise, the negative sign holds.

### 3. RESULTS

#### 3.1. Galaxy Luminosity Function

Before presenting detailed analyses for various aspects of DLAs, in this subsection we show that the luminosity functions of simulated galaxies at $z = 2–3$ are in reasonable agreement with observations in the sense that the C and V runs bracket the observed data points. Figure 1 shows that the rest-frame UV luminosity functions from C (red dots) and V runs (green squares) are in approximate agreement with UV observations in the redshift range $z \sim 2–3$ for the range of $M_{AB}(1700) = -22$ to $-19.5$, corresponding to SFR $\geq 4-40 M_{\odot}$ yr$^{-1}$. At the bright end $M_{AB}(1700) \leq -22$, the UV luminosities of galaxies are substantially affected by dust absorption and a large portion of their SFR emerges in infrared as ultraluminous infrared galaxies (ULIRGs) and luminous infrared galaxies (LIRGs). We see that, taking this into account, by including the abundance of the observed ULIRGs and LIRGs, the simulations are in better agreement with observations at $M_{AB}(1700) \leq -19.5$ in the sense that the C and V runs bracket the observations in amplitude and have the approximately correct slope of the luminosity function at $z \sim 2–3$. There is indication that the faint-end slope of the simulated galaxies at $M_{AB}(1700) \geq -19.5$ is significantly flatter than observed. This is due to limited mass resolution in the simulations. Comparably approximate agreement between simulations and observations at $z = 0$ was shown in Cen (2011, Figure 3). In combination, this provides a useful validation of the simulations with respect to properties of the overall population of simulated galaxies, allowing us to examine, with significant confidence, DLA properties in the CDM model. Nevertheless, it is useful to keep in mind that our simulations underestimate the contributions to DLAs from galaxies less massive than $\sim 10^{10} M_{\odot}$. In Nagamine et al. (2007) their simulation Q5 has about 20 times higher mass resolution than ours with a feedback strength, corresponding to a wind velocity of 484 km s$^{-1}$ that is in the right range of the observed LBG winds (e.g., Pettini et al. 2001) and is consistent with the winds that we see in our simulations. Thus, we use their results to gain an estimate of the contribution of smaller galaxies of halo mass $\lesssim 10^{10} M_{\odot}$ that are not well resolved in our simulations. An examination of Figure 5 of Nagamine et al. (2007) reveals that the contribution to DLA cross section from galaxies of halo mass $\lesssim 10^{10} M_{\odot}$ is less then 10\%--15\%. In all the statistical measures presented we do not correct for this.

#### 3.2. Rich Varieties of DLAs

We first present a gallery of 12 DLAs at $z = 3.1$ (Figures 2–13) to show the richness of their physical...
properties. For each DLA six maps and five quantitative panels are displayed. In each of the six maps the LOS intercepting the DLA is shown as a white horizontal line and the exact location of the primary component of the DLA is at the intersection with another, white vertical line. In cases with multiple components along the LOS, the primary component coincides with the highest neutral density. Four of the maps—top left (temperature in Kelvin), middle left (atomic hydrogen density in cm$^{-3}$), bottom left (metallicity in solar units), and top middle (pressure in Kelvin cm$^{-3}$)—have a physical thickness of 1 kpc. Also indicated in top middle (pressure) is the peculiar velocity field with a scale of 5 kpc corresponding to 500 km s$^{-1}$. The remaining two maps—middle right (baryonic overdensity) and bottom middle (stellar surface density in $M_\odot$ kpc$^{-2}$)—are projected over the entire galaxy of depth of order of the virial diameter of the primary galaxy. While these two projected maps give an overall indication of relative projected location of the DLA with respect to the galaxy, the exact depth of the DLA inside the paper is, however, not shown. When we quote distance from the galaxy, we mean the projected distance on the paper plane.

The five panels on the right column show various physical quantities along the LOS (i.e., along the white horizontal line shown in the maps on the left two columns). From top to bottom they are atomic hydrogen density (in cm$^{-3}$; red solid curve with a narrower shape) along with total hydrogen density (dotted green curve), gas metallicity (in solar units), LOS proper peculiar velocity, Ly$\alpha$ flux, and finally Si$\ ii$ 1808 flux. The top three panels are plotted against physical distance, the bottom two versus LOS velocity.

"Figure 2" shows a DLA produced by the LOS intersecting the tip of a long chimney at a distance of $\sim$30 kpc from the galaxy. The velocity structure suggests that it is still moving away (upward) from the galaxy at a velocity of $\sim$500 km s$^{-1}$, (A color version of this figure is available in the online journal.)
caused by galactic winds. The metallicity at the interception is \([Z/H] \sim -1.5\), but there are large gradients and variations of metallicity (we found that some other very nearby DLA systems intersecting different parts of the chimney have metallicity \([Z/H] < -3\), not shown), suggesting a very inhomogeneous enrichment process by galactic winds. While the primary galaxy has a total mass of \(4 \times 10^{12} \, M_\odot\), i.e., a one-dimensional velocity dispersion of 500 km s\(^{-1}\), the kinetic width of this line is only 129 km s\(^{-1}\) with \(N_{\text{HI}} = 21.7\). Although the Ly\(\alpha\) flux appears as a single component, as will be the case in all subsequent examples, the Si\(\text{II}\) absorption has several separate features, reflecting the two-peak structure of the absorbing column and complex velocity structure within. We note that the nearby satellite galaxies may have triggered the starburst and the galactic winds. The responsible gas for this DLA is probably cooling and confined by external pressure due to thermal instability, as seen in the pressure panel.

In Figure 3, a DLA of width 219 km s\(^{-1}\) is created jointly by two major components along the sightline, one at \(x \sim 70\) kpc of metallicity of \([Z/H] \sim [-1.0, -0.5]\) and size of \(\sim 20\) kpc at an impact parameter of \(\sim 30\) kpc and the other at \(x \sim 110\) kpc of metallicity of \([Z/H] \sim 0.0\) and size of \(\sim 30\) kpc. What is striking are the many long gaseous structures in this galaxy. As we will see frequently, there are often long gaseous structures connected with galaxies that are almost always coincidental with visible galaxy interactions of multiple galaxies or galaxies and satellites in close proximity. We shall call these features “galactic aqueducts” hereafter. Some of these galactic aqueducts are cold streams (Kereš et al. 2005; Dekel & Birnboim 2006). However, galactic aqueducts found in our simulations are very rich in variety and disparate in metallicity (spanning three decades or more in metallicity). In other words, they are not necessarily primordial cold streams. In the case of this DLA, the filaments are made of pre-enriched gas, having cooled (as the pressure panel shows) and now rotating about the galaxy (roughly counterclockwise). In contrast, note that the DLA in Figure 2 was still moving away from the galaxy. Like in Figure 2, the rich galactic aqueducts appear to be associated with significant satellite structures in close proximity. The Si\(\text{II}\) absorption has several separate features, reflecting the two separate physical components, as well as substructures within each component.

Figure 4 shows a DLA that is associated with a low-metallicity ([Z/H] \sim [-2.0, -1.5]) filament that is feeding a small satellite, which in turn is interacting and feeding the primary galaxy at a projected distance of \(\sim 20\) kpc. This is yet another example of interacting galaxies producing rich gas-feeding filaments, as already seen in Figures 2 and 3. The relatively large width of 303 km s\(^{-1}\) is produced by steep
velocity gradient in the region from $x \sim 45$ to 50 kpc. One could see that galactic winds are blowing to the upper left corner by the primary galaxy, whose starburst is triggered by the interaction.

Figure 5 shows a DLA that is made up of several filaments at distances of 30–40 kpc from the galaxy. The metallicity of all the components is near solar, indicating that this is probably pre-enriched gas cooling due to thermal instability. The velocity structures show that they are falling back toward the galaxy, in a fashion perhaps similar to galactic fountains (Shapiro & Field 1976), on somewhat extended scales. Once again, galaxy–galaxy interactions may be responsible for the rich gas filaments, as seen in Figures 2–4. There is evidence that winds are blowing upward from the galaxy.

Figure 6 shows another example of a DLA arising from a galaxy with a very rich filament system due to galaxy interactions. The primary galaxy is the same as the one shown in Figure 3, and we are now looking at its south side. These filaments that are responsible for the neutral column of the DLA are enriched to a level of $[Z/H] \sim -1.0$ and have cooled to low temperature. The large width of 420 km s$^{-1}$ is due to the multiple components spanning a spatial range of $\sim 40$ kpc each of physical depth of several kpc and individual velocity width $\leq 100$ km s$^{-1}$. Interestingly, for this DLA system, while most of the gas filaments are now falling back toward the galaxy (not necessarily radially), galactic winds are still blowing toward the upper right corner. Comparison of Figure 6 and Figure 3 indicates that the metallicity in the upper right quadrant of the galaxy is somewhat more metal enriched ($[Z/H] \sim 0$) than other regions ($[Z/H] \sim -1$ or lower), consistent with the directions of ongoing galactic winds. This is strongly suggestive that the metallicity enrichment process not only is episodic, multi-generational, and anisotropic but also in general possesses no parity.

Figure 7 shows a DLA intercepting two filaments at a small inclined angle, giving rise to a broad physical extension of $\sim 25$ kpc. All the visible filaments run roughly top left to bottom right, whereas the metal-enriched regions spread out like a butterfly in the direction roughly perpendicular to the filaments. This is the classic picture that galactic winds tend to blow in directions that are perpendicular to the filaments that are feeding the galaxy. The inner regions of the filaments are less enriched ($[Z/H] \sim -2$) than the outer regions of the filaments ($[Z/H] \sim -0.5, 0$), strongly indicative of galactic winds tending to circumvent the denser filaments. The large width of 510 km s$^{-1}$ is caused by oppositely moving (i.e., converging) flows at $x \sim 15–40$ kpc, probably caused by the
bipolar winds interacting with the complex filament structures. This galaxy has a total mass of $2 \times 10^{11} M_\odot$, and we note that most of its surrounding regions are relatively cold, whereas in Figures 2–6 we consistently see a hot atmosphere permeating the circumgalactic regions. The galaxies in Figures 2–6 all have total mass $\geq 10^{12} M_\odot$, consistent with the mass demarcation of cold and hot accretion modes (Kereš et al. 2005; Dekel & Birnboim 2006). Nevertheless, the existence of cold galactic aqueducts seen in Figures 2–6 is consistent with the suggested cold mode of accretion of massive galaxies at high redshift (Dekel et al. 2009).

Figure 8 shows a “normal” DLA where a relatively quiet galactic disk is pierced edge-on. Its large width of 522 km s$^{-1}$ is simply due to the large halo in which the galaxy is residing of total mass $3 \times 10^{12} M_\odot$ (at $z = 3.1$). The surrounding environment is relatively “pristine” with no widespread metal enrichment at a level of $[Z/H] \geq -1$. However, the temperature panel indicates that there is a hot halo permeating the entire region and embedding and pressure confining (see the pressure panel) the cold neutral clouds. This hot gaseous halo is produced by gravitational shocks rather than galactic wind shocks. There are several filaments attached to the galaxy. This is a good example of cold streams feeding a massive galaxy by penetrating a hot atmosphere.

Figure 9 shows a DLA with a large velocity width arising from a relatively small galaxy of total mass $3 \times 10^{10} M_\odot$. The galactic winds are blowing in the northeast direction, which entrain cold neutral clouds with them. The LOS of the DLA intercepts a high-velocity component at $x = 35–55$ kpc. The combination of this high-velocity component with the low-velocity component at $x \sim 0$ produces the relatively large width of 306 km s$^{-1}$. Note that an isotropic Maxwellian velocity distribution of dispersion equal to that of the halo velocity dispersion would only yield a width of $v_0 = 2.33 v_{\text{vir}} = 176$ km s$^{-1}$. Clearly, galactic winds are directly responsible for the large width of this DLA, by entraining cold gas clouds to a high velocity.

Figure 10 shows another DLA produced by a small galaxy of total mass $2 \times 10^{10} M_\odot$ with a large velocity width. The galaxy system has multiple interaction galaxies at close distances. The galactic winds are blowing primarily, in a bipolar fashion, in northeast and southwest directions, roughly perpendicular to the galactic disk that entrains the cold neutral cloud at $x \sim 40$ kpc to a broad velocity of $v_x = 0–400$ km s$^{-1}$ relative to the galaxy itself. In combination with another complex structure at
with the disk of the galaxy, which would have produced a large width of 501 km s\(^{-1}\). Together, a large width of 501 km s\(^{-1}\) is produced. Given the small mass of the galaxy, the surrounding regions are not embedded in a hot gravitationally shock-heated atmosphere. There are some solid angles with low gas column that have been heated by galactic winds, probably triggered by the binary galaxy interaction, as can be seen by comparing the temperature map and the density map. Note that 2.33\(v_{\text{vir}}\) = 204 km s\(^{-1}\) for this galaxy.

Figure 13 shows the last example of a wide DLA from a small galaxy. Two galactic aqueducts make up this DLA, one at \(x \sim 10\) kpc and the other at \(x \sim 35\) kpc. The galactic system is a primary-satellite binary that is interacting, which has caused both to experience starbursts. The primary galaxy at \((x, y) \sim (23, 36)\) kpc is blowing bipolar galactic winds mainly in the north–south direction, whereas the satellite at \((x, y) \sim (35, 35)\) kpc is blowing bipolar galactic winds in the east–west direction. Together they produce a very complex, multi-stream velocity structure. The total velocity width of this DLA is 500 km s\(^{-1}\), although each of the two components individually has a velocity width of \(<200\) km s\(^{-1}\). Note that 2.33\(v_{\text{vir}}\) = 200 km s\(^{-1}\) for the primary galaxy.

In summary, we see that DLAs arise in a wide variety of cold gas clouds, from galactic disks to cold streams to cooling gas from galactic winds to cold clouds entrained by hot galactic winds.
winds at a wide range of distances from galaxies, with a wide range of metallicity and in galaxies of total mass from $10^{10}$ to $10^{-12.5} M_\odot$ at $z \sim 3$. Inspection of the gallery has already hinted that many large velocity width DLAs may be produced directly or indirectly by galactic winds, that is, directly by entraining cold gas clouds and compressing cold gas clouds with high pressure and indirectly by enhancing cooling and thermal instability with added metals and shock compression. In addition, the composite nature of many large width DLA systems should also help remove the perceived failure of the standard LCDM model with respect to producing large width systems directly or indirectly by galactic winds, that is, directly by thermal instability with added metals and shock compression.

Similar to the situation for the velocity distribution function (Figure 16), the strong environmental dependence of the column density distribution renders it impractical to make vigorous comparisons between the simulations and observations. Given that the amplitude of observed column density distribution lies between that of the “C” and that of the “V” runs, and the shapes of both simulated functions are in reasonable agreement with observations, we tentatively conclude that the standard LCDM model can reasonably reproduce the observed column density distribution. This conclusion is self-consistent with the galaxy luminosity functions shown in Figure 1, where the “C” and “V” runs also bracket the observation. Note that the shape at the highest column end depends on the treatment of high-density regions, for which we have used an empirical relation. Ultimately, when parsec resolution is reached, we can make more definitive tests. What is also interesting is that the variations between different environments are larger than the redshift evolution of the column density distribution in each run. It is further noted that, seen in the lower left panel of Figure 14, the evolution of the column density distribution in the “C” and “V” runs is different. In the “C” run, we see weak evolution from $z = 4$ to $z = 1.6$ and then a relatively large drop in amplitude at $z = 0$. In the “V” run, on the other hand, we see practically little evolution.

3.3. Column Density Distribution, Line Density, and $\Omega_8$ (DLA) Evolution

We begin our quantitative discussion with the most fundamental observable: the column density distribution of DLAs and its evolution. Figure 14 shows the column density distribution at several redshifts from $z = 0$ to $z = 4$. Where comparisons can be reliably made with observations, at $z = 2.5$, $z = 3.1$, and $z = 4$, we see that the overdense run “C” and underdense run “V” appropriately bracket the observational data in amplitude.
from \( z = 3.1 \) to \( z = 0 \). This reflects the dynamical stage of a simulated sample, where the “C” run is more dynamically advanced than the “V” run at the same redshift, consistent with the behavior seen in Figure 26 below.

The left panel of Figure 15 shows the redshift evolution of DLA line density, defined to be the number of DLAs per unit absorption length. The right panel of Figure 15 shows the redshift evolution of neutral gas density in DLAs. Inherited from the situation shown in Figure 14, there is a large variation of both plotted quantities between the two (“C” and “V”) runs. What is reassuring is that the observed data lie sensibly between results from these two bracketing environments. If one assumes that the cosmic mean of each of the two plotted quantities should lie between the “C” and “V” runs, reading the range spanned by the two runs suggests that the LCDM model agreements with observations to within a factor of \( 2 \) with respect to both quantities, although what the overall temporal shape will look like is difficult to guess.

To firmly quantify these important observables and to more precisely assess the agreement/disagreement between the predictions of the LCDM model and observations, a larger set of simulations sampling, more densely, different environments in a statistically correct fashion is necessary, so is a more accurate treatment of the transition from atomic to molecular hydrogen in very high density regions (which affects the shape at the high column density end). We reserve this for future work.

### 3.4 Kinematic Velocity Width Distribution Functions

Since the velocity structure in Ly\( \alpha \) flux of a DLA is “damped” and does not provide the kinematic information of the underlying physical cloud, following Prochaska & Wolfe (1997), the velocity width, \( v_{90} \), is defined to be the velocity interval of 90% of the optical depth of the Si \( \text{II} \) \( \lambda \)1808 absorption line associated with the DLA. Figure 16 shows the velocity width distribution at three redshifts \( (z = 1.6, 3.1, 4.0) \), covering most of the observed redshift range. We see a factor of \( \sim 10 \) variation from the “C” to the “V” run, indicating the need to have a larger statistical set of simulations covering, more densely, different environments, before a more precise comparison can be made with observations. Insofar as the observed velocity width distribution function lies between the two bracketing runs, “C” and “V,” and the shape of the functions is in excellent agreement with observations, including the high velocity tail \( (v_{90} \gtrsim 300 \text{ km s}^{-1}) \), this should be considered a success for the LCDM model—there is no lack of large-width DLAs with \( v_{90} \gtrsim 300 \text{ km s}^{-1} \) in the LCDM simulation. This conclusion is consistent with that of Hong et al.
(2010), who studied this issue with a different code and a different feedback implementation. There is a significant difference between our results and theirs in that we find that galactic winds are directly responsible for many of the large-width DLAs, by entraining neutral dense clouds to large velocities. In addition, they conclude that a large halo mass ($\geq 10^{11} \, M_\odot$) is a necessary condition for producing large velocity widths, while we find that a non-negligible fraction of large velocity width DLAs arise in halos less massive than $10^{11} \, M_\odot$. It is useful to note, however, that a large fraction of less massive halos that give rise to large-width DLAs are satellite galaxies in larger halos.

Note that small galaxies of halo mass $9 < 10^{10} \, M_\odot$ not resolved in our current simulations should contribute the low velocity end ($v_9 < 30 \, \text{km s}^{-1}$) of the distribution in Equation 16.

To help understand the large velocity width DLAs, we plot in Equation 17 the maximum $v_9$ of all DLAs, $v_9 \, \text{max}$, associated with each galaxy against the halo mass of the galaxy, $M_{\text{halo}}$. The black line $v_9 = 2.33 \, v_{\text{vir}}$ is what $v_9$ would be if the velocity distribution were an isotropic and Maxwellian distribution with its dispersion equal to $v_{\text{vir}}$, and the Si II gas density is constant across the DLA. We see that the Maxwellian velocity distribution (the black line) approximately provides a lower bound to $v_9$, although there is, unsurprisingly, some fraction of systems that lie below (see Figure 2 for an example). What is very interesting is that at $z = 3.1$ there are a large number of galaxies whose $v_9 \, \text{max}$ are substantially larger than what $v_{\text{vir}}$ could produce, i.e., “super-gravitational motion” in the terminology of Hong et al. (2010). This super-gravitational motion is produced by galactic winds for some halos, as we have seen clearly in Figures 9–13 in Section 3.1. For others, it simply reflects the fact that some of the small galaxies producing large-width DLAs are satellite galaxies of larger halos; therefore, this “super-gravitational motion” for some small halos is merely gravitational motion in the potential well of the primary galaxies. We also note that at $z = 1.6$ for both the “C” and “V” runs (and especially the “C” run), the correlation between $v_9$ and $v_{\text{vir}}$ becomes substantially better with much reduced scatter, and the excess of DLAs with large $v_9 / v_{\text{vir}}$ is much removed. This is circumstantial but strong evidence that galactic winds are responsible for at least some of the large $v_9 / v_{\text{vir}}$ DLAs, because of higher star formation activities and hence galactic winds at $z = 3.1$ than at $z = 1.6$. Figure 19 below for $z = 0$ will further strengthen this point.

We see that the redshift evolution at a fixed environment is relatively mild in the redshift range $z = 4.0$ to $z = 1.6$. We speculate that the weak evolution of the velocity width
distribution from $z = 4.0$ to $z = 1.6$ may be coincidental and attributable to counteracting processes: growth of halo mass and hence virial velocity with time on the one side and diminution of super-gravitational motion produced by galactic winds (due to reduced star formation activities with time at $z < 2$) and decreasing density with time on the other. This prediction of a weak evolution of velocity width distribution with redshift is verifiable with future larger DLA samples and is a powerful test for the model.

Figure 17 does not, however, fairly characterize the relative contribution of halos of different masses to velocity width distribution function, because it does not specify the number of DLAs at a given halo mass. In Figure 18 we show the halo mass probability distribution function for DLAs above three velocity width cuts, $v_{90} \geq 150, 300, 600$ km s$^{-1}$, respectively. We see a clear trend that larger halos make larger contributions to larger width DLAs, as one would have expected. For example, about one-half of all DLAs with $v_{90} \geq 600$ km s$^{-1}$ arise in halos of mass greater than $10^{12} M_\odot$ at $z = 3.1$, whereas that division line drops to $2 \times 10^{11} M_\odot$ for $v_{90} \geq 150$ km s$^{-1}$. It should be noted that the ratio of the virial velocity of a halo of mass $2 \times 10^{11} M_\odot$ to that of $10^{12} M_\odot$ is $0.58$, significantly greater than $0.25 = 150/600$, indicating an overweight of DLA cross section by large galaxies. For moderate to large velocity width of $v_{90} \geq 150$ km s$^{-1}$, halos of mass $1 \times 10^{11} M_\odot$ dominate the contribution to DLA incidence, largely in agreement with Hong et al. (2010). Slightly at odds with Hong et al. (2010), however, we find a significant fraction of these relatively wide systems arising in galaxies of mass less than $1 \times 10^{11} M_\odot$: (24%, 18%, 12%) of DLAs with velocity width larger than $(150, 300, 600)$ km s$^{-1}$ are due to galaxies with total mass less than $1 \times 10^{11} M_\odot$.

We note that our definition of associating DLAs with galaxies biases associating them with larger galaxies; Figure 4 gives an example, where the DLA is defined to arise from the larger galaxy of total mass $8 \times 10^{11} M_\odot$, even though it is more closely related to a much smaller satellite galaxy that is orbiting around the larger galaxy. Our results are perhaps unsurprising in the sense that one would have expected that galactic winds, when they are blowing, should be stronger, or at least not weaker, in dwarf starburst galaxies than larger galaxies thanks to shallow gravitational potential wells in the former, when cold gas is still abundant at high redshift. Both Figure 18 and gallery pictures in Section 3.1 confirm this point. Galactic winds, however, could be weaker in dwarf galaxies if star formation is in proportionally less vigorous. This may be the case at lower redshift, as shown in Figure 19. What is interesting, and further evidence, is that...
at \( z = 0 \) the dwarf galaxies in the “V” run have more super-gravitational motion than in the “C” run, simply because the former are gas richer and have higher SFR than the latter. Thus, galactic winds are a bivariate function of total galaxy mass and SFR, in a fashion that is consistent with observations (e.g., Martin 2005).

3.5. Metallicity Distribution and Evolution

The current set of simulations is vastly superior to those used in our earlier work addressing the observed relatively weak but non-negligible evolution of DLA metallicity (Cen et al. 2003), and here we return to this critical issue. Figure 20 shows the DLA metallicity distributions at four redshifts, \( z = 0, 1.6, 3.1, 4.0 \). For the three redshifts, \( z = 1.6, 3.1, 4.0 \), where comparisons can be made, we find that the agreement between simulations and observations at \( z = 1.6 \) to \( z = 0 \) is excellent, as K-S tests show. This is a non-trivial success, given that our feedback prescription has essentially one free parameter, that is, the supernova energy that is driving galactic winds transporting energy, metals, and mass throughout interstellar (ISM), circumgalactic (CGM), and intergalactic space (IGM). Furthermore, the absolute amount of metals is totally fixed by requiring that 25% of stellar mass with metallicity equal to \( 10 Z_\odot \) returns to the ISM, CGM, and IGM, consistent with the Salpeter IMF adopted. The agreement indicates that our choices of both the supernova ejecta mass and its metallicity and the explosion energy, which are inspired by theories of stellar interior and direct observations, provide a reasonable approximation of truth.

We see that the peak of the DLA metallicity distribution evolves from \( [Z/H] = -1.5 \) at \( z = 3-4 \), to \( [Z/H] = -0.75 \) at \( z = 1.6 \), and to \( [Z/H] = -0.5 \) at \( z = 0 \). Thus, both simulations and nature indicate that there is a weak but real evolution in DLA metallicity. What is also important to note is that, in agreement with observations, simulations indicate that the distribution of metallicity is very wide, spanning three or more decades at \( z \geq 1.6 \). This wide range reflects the rich variety of neutral gas that composes the DLA population, from relatively pristine gas clouds falling onto or feeding galaxies, to metal-enriched cold clouds that are falling back to (galactic fountain) or still moving away from (due to entrainment of galactic winds) galaxies, to cold neutral gas clouds in galactic disks. There is a metallicity floor at \( [Z/H] = -3 \) at \( z = 1.6-4 \), and that floor moves up to \( [Z/H] = -1.5 \) by \( z = 0 \), consistent with observations (Prochaska et al. 2003). The distribution at \( z = 0 \) is significantly narrower, partly reflecting the overall enrichment of the IGM and partly due to much reduced variety of DLAs.
with galactic disks becoming a more dominant contributor to DLAs (see discussion below).

Ellison et al. (2010) find that proximate DLAs (PDLAs), those within a velocity distance from the QSO $\Delta v < 3000$ km s$^{-1}$, have metallicity higher than the more widely studied, intervening DLAs. Thus, the total sample of PDLAs plus conventional (intervening) DLAs may somewhat shift the metallicity distribution to the right, perhaps bringing it to a still better agreement with our simulations.

Observations have found a strong positive correlation between galaxy stellar mass and metallicity (e.g., Erb et al. 2006). We divide the simulated DLA sample at $z = 3.1$ into two subsets, one with metallicity less than $[Z/H] = -1$ and the other with metallicity more than $[Z/H] = -1$. We then compute the velocity width functions separately for each subset, which are shown as solid dots (lower metallicity) and solid squares (higher metallicity) in the left panel of Figure 21. What we see is that there is a small excess of large velocity width DLAs for the higher metallicity subset compared to the lower metallicity one. This is, of course, in the sense that is consistent with the observed metallicity–mass relation. However, the current observational data sample is consistent with simulations, and the difference between the two simulated subsets and between the two observed subsets is statistically insignificant. A larger sample (by a factor of four) may allow for a statistically significant test. Do we expect a larger difference in the disk model (Wolfe et al. 1986; Prochaska & Wolfe 1997)? We do not have a straight answer to this question, without a very involved modeling. However, we suggest that the picture we have presented, where DLAs arise from a variety of galactic systems, in a variety of locations of widely varying metallicity (see the gallery in Section 3.1), would be consistent with the small difference found, because the velocity widths of large-width DLAs do not strongly correlate with galaxy stellar mass (see Figure 17). In other words, the observed correlation between metallicity and galaxy stellar mass is largely washed out by DLAs that do not arise in disks and whose metallicities do not strongly correlate with galaxy stellar mass. If one combines the information provided by Figures 18 and 21, one may reach a similar conclusion.

### 3.6. Size Distribution

Binary quasars, physical or lensed, provide a unique tool to probe the size of DLAs. Here, we present our predictions of size distributions of DLAs in the LCDM model. As we have described in Section 2.3, any cells (of size $0.915 h^{-1}$ kpc comoving) that are connected by one side in projection are merged into a “single isolated” DLA. The area of each “isolated” DLA, $A$, is then used to define the size (radius) of the DLA by

![Figure 12](image-url)
$r_{DLA} = (A/\pi)^{1/2}$. The total area of all isolated DLAs associated with a galaxy along three orthogonal directions ($x$, $y$, $z$), $A_x$, $A_y$, and $A_z$, are summed to obtain $A_{tot} = A_x + A_y + A_z$, and the total DLA size (radius) of the galaxy is defined to be $r_{tot} = (A_{tot}/\pi)^{1/2}$. Note that, if DLAs arises from a thin disk, the $A_{tot}$ computed this way will be the exact size of the disk face-on, regardless of its orientation. On the other hand, if each DLA cloud is a sphere, this method overestimates the size (area) by a factor of $\sqrt{3}$.

Figure 22 shows the size (radius) distribution at redshift $z = 0, 1.6, 3.1, 4.0$ for individual DLA size and total DLA size of each galaxy. Figure 23 presents the size information in a different way, where we show the probability that, for a random pair of sightlines separated by (30, 20, 10, 5, 3) kpc at the redshift in question, one sightline intercepts a DLA and the other intercepts a column density lower than shown in the $x$-axis. We see that, in the sense that is consistent with distance distributions shown in Figure 26, on average, individual DLA size, as well as the total DLA size of galaxies, is larger at high redshift than at lower redshift. The individual DLA size distribution sharply peaks at $r_{DLA} \sim 15$ kpc at $z = 4$, then moves left to a broader peak at $r_{DLA} \sim 10$ kpc at $z = 3.1$, then sharpens somewhat to peak at $r_{DLA} \sim 5$ kpc at $z = 1.6$, and finally moves rightward slightly to peak at $r_{DLA} \sim 7$ kpc at $z = 0$ for the “C” run. The numbers for the “V” run are largely comparable to the “C” run: (15, 12, 4, 6) kpc at $z = (4, 3.1, 1.6, 0)$, respectively. What is quite remarkable is that, at $z = 1.6$ where comparison can be made, the predicted size distribution and the available observation agree better than anyone would have expected. This is a testament to the success of the LCDM model and physical treatment of our simulations, especially considering other agreements that we have already found, for example, with respect to metallicity distribution, column density distribution, and kinematics (see Section 3.8).

Intriguingly, Rauch et al. (2008) find a new population of faint Lyα emitters that they advocate may be the host galaxies of DLAs at $z \sim 3$. Their estimated size, based on the observed Lyα line emission profile, shown as the open circle in the $z = 3.1$ panels of Figure 22, would be consistent with our model, although the true nature of this population is relatively uncertain. Given the reasonable agreement in size, we will make some further speculation to gauge if our model can accommodate or explain this population as DLAs. One possible physical mechanism for the Lyα emission of this population, if they are indeed DLA hosts, would be fluorescence due to ionizing photons from the host galaxies. Assuming that the distance of each DLA cloud from the host galaxy is 25 kpc (Figure 26), the size of each DLA cloud is 8 kpc (Figure 22), and the SFR.
Figure 14. Column density distributions, defined to be the number of DLAs per unit column density per unit absorption length, at $z = 2.5$ (lower right), $z = 3.1$ (upper left), and $z = 4.0$ (upper right), separately, and together for $z = 0$, 1.6, 3.1, 4.0 (lower left). In each panel, two sets of simulation results are shown, one for the “C” run (solid dots) and one for the “V” run (open circles). The corresponding observational data for each of the individual redshifts are an updated version with SDSS DR7 from Prochaska et al. (2005), shown as open squares. Also shown in three of the panels as stars are the observational data from Noterdaeme et al. (2009) for the entire SDSS DR7 sample in the redshift range $z = 2.2$–5.2.

(A color version of this figure is available in the online journal.)

Figure 15. Left panel: the redshift evolution of the DLA line density for the “C” run (solid dots) and the “V” run (solid squares). The observational data at $z > 2$ are an updated version with SDSS DR7 from Prochaska et al. (2005), shown as open squares; the observational data at $z < 2$ are from Rao et al. (2006), shown as open circles. Right panel: the redshift evolution of the neutral gas density in DLAs for the “C” run (solid dots) and the “V” run (solid squares). The observational data at $z > 2$ are an updated version with SDSS DR7 from Prochaska et al. (2005) shown as open squares, the data using SDSS DR7 are from Noterdaeme et al. (2009) shown as open triangles, the observational data at $z < 2$ are from Rao et al. (2006) shown as open circles, the observational data point at $z = 0$ is from Zwaan et al. (2005) shown as a diamond, and the observational data point at $z = 1.96$ is from Péroux et al. (2003) shown as a star.

(A color version of this figure is available in the online journal.)

of the host galaxy is $10 \ M_{\odot} \ yr^{-1}$ (see Figure 35 below), then a typical expected SFR inferred from their fluorescence would be $(8^2 \pi)/(4\pi 25^2) \times 10 = 0.26 \ M_{\odot} \ yr^{-1}$. This very rough estimate curiously falls within the range of 0.07–1.5 $M_{\odot} \ yr^{-1}$ estimated by Rauch et al. (2008) for the observed emitters if there are Ly$\alpha$ emitters at $z \sim 3$.

Now we make more direct comparisons between our model and the available observed 18 pairs of QSOs (all lensed
Figure 16. Velocity distribution functions, defined to be the number of DLAs per unit width velocity per unit absorption length, at $z = 1.6, 3.1, 4.0$. Two sets of simulation results are shown, one for the “C” run (solid symbols) and one for the “V” run (open symbols). The corresponding observational data for each of the individual redshifts (Prochaska et al. 2005) are shown as open squares, which span the redshift range of $z = 1.7–4.5$.

(A color version of this figure is available in the online journal.)

Figure 18. DLA incidence weighted halo mass probability distribution function for DLAs above three velocity width cuts, $v_{90} \geq 150, 300, 600$ km s$^{-1}$ at $z = 3.1$ for the “C” run. Note that a DLA associated with a satellite galaxy or any gas cloud within the virial radius is given the halo mass of the primary galaxy.

(A color version of this figure is available in the online journal.)

Figure 17. Maximum $v_{90}$ of all DLAs associated with each galaxy against the halo mass of the galaxy $M_{\text{halo}}$ for $z = 1.6$ and $z = 3.1$ for the “C” and “V” runs. The black line $v_{90} = 2.33v_{\text{vir}}$ is what $v_{90}$ would be if the velocity distribution is an isotropic and Maxwellian distribution with its dispersion equal to $v_{\text{vir}}$, and the Si II gas density is constant across the DLA.

(A color version of this figure is available in the online journal.)
pairs/multiples except one real physical binary) at $z \sim 1.6$, given in Table 1. Columns 1–7 list parameters of each observed pair, and Columns 8–9 give the probability of each pair occurring in the “C” [P(C)] and “V” [P(V)] runs, respectively, at $z = 1.6$. We have made the following simplification for computing the probability: we treat all DLAs with $N_{\text{HI}} \geq 2 \times 10^{20} \text{ cm}^{-2}$ the same regardless of column density values, and for non-DLA absorbers the probability that we present is the probability of having a column equal to or less than the indicated value. We see from Table 1 that the observed QSO pairs are all statistically consistent with our model at the $\sim 1.5\sigma$ level or better, entirely consistent with the agreement shown in Figure 22 between the computed size distribution and observationally inferred size, which is somewhat model dependent.

In Figure 24, we plot the total DLA cross section of each galaxy $(A_{\text{tot}})$ against its halo mass at $z = 3.1$ (left panel) and $z = 0$ (right panel). Also shown as the black line is a linear fit assuming that the effective radius $r_{\text{tot}} = (A_{\text{tot}}/\pi)^{1/2}$ is proportional to the virial radius for simplicity. Note that a least-squares linear fit is slightly flatter than the black curve. It is noted that there are a small fraction of galaxies that do not have significant neutral gas and thus do not show up in the plotted range. What is most striking is that, for gas-rich galaxies that give rise to DLAs at $z = 3.1$, on average, the effective area that gives rise to DLAs occupies about 1/3 of the total area within the virial radius, i.e., $r_{\text{tot}} = 0.6r_{\text{vir}}$. By $z = 0$ the effective radius has reduced to $r_{\text{tot}} = 0.08r_{\text{vir}}$, becoming comparable to the size of the stellar disk. We note that the method we use to compute the total DLA cross section gives a correct value for the case of a thin disk but could overestimate it by a factor of $\sqrt{3}$ in the case of a sphere.

### Table 1

| QSO          | $z_{\text{abs}}$ | Pair | $\theta_{\text{abs}}$ (kpc) | $d$ | $N_1$ | $N_2$ | $P(C)$ | $P(V)$ |
|--------------|------------------|------|-----------------------------|-----|------|------|--------|--------|
| H1413+1143   | 1.440            | B–A  | 0.753                       | 3.13| 60   | 9.0  | 0.84   | 0.80   |
|              |                  | B–D  | 0.967                       | 4.02| 60   | 0.25 | 0.11   | 0.13   |
|              |                  | B–C  | 1.359                       | 5.65| 60   | 0.20 | 0.17   | 0.21   |
|              |                  | A–D  | 1.118                       | 4.64| 9.0  | 0.25 | 0.13   | 0.16   |
|              |                  | A–C  | 0.872                       | 3.62| 9.0  | 0.20 | 0.093  | 0.11   |
|              |                  | D–C  | 0.893                       | 3.71| 0.25| 0.20 | 0.097  | 0.12   |
| H1413+1143   | 1.486            | D–A  | 1.118                       | 4.32| 2.0  | <0.05| 0.097  | 0.11   |
|              |                  | D–B  | 0.967                       | 3.73| 2.0  | <0.1 | 0.084  | 0.097  |
|              |                  | D–C  | 0.893                       | 3.45| 2.0  | <0.05| 0.070  | 0.079  |
| H1413+1143   | 1.662            | B–A  | 0.753                       | 2.17| 6.0  | 1.5  | 0.92   | 0.89   |
|              |                  | B–C  | 1.359                       | 3.91| 6.0  | 0.6  | 0.13   | 0.17   |
|              |                  | B–D  | 0.967                       | 2.78| 6.0  | 0.3  | 0.064  | 0.078  |
|              |                  | A–C  | 0.872                       | 2.51| 1.5  | 0.6  | 0.058  | 0.074  |
|              |                  | A–D  | 1.118                       | 3.22| 1.5  | 0.3  | 0.086  | 0.11   |
|              |                  | C–D  | 0.893                       | 2.57| 0.6  | 0.3  | 0.052  | 0.063  |
| HE 1104–1805 | 1.662            | A–B  | 3.0                         | 4.47| 6.3  | <0.013| 0.10   | 0.11   |
| UM 673       | 1.6253           | A–B  | 2.22                        | 2.71| 6.3  | <0.037| 0.036  | 0.036  |
| LBQS 1429-0053 | A–B  | 5.11 | 43.9                       | 3.0  | 0.1  | 0.051  | 0.050  |

**Notes.** Column 1: QSO name; Column 2: DLA redshift; Column 3: pair name; Column 4: separation of two images in the sky in arcseconds; Column 5: physical separation of two sightlines at $z_{\text{abs}}$; Column 6: $H_1$ column density of the first sightline in units of $10^{20} \text{ cm}^{-2}$; Column 7: $H_1$ column density of the second sightline in units of $10^{20} \text{ cm}^{-2}$; Column 8: probability from the “C” run; Column 9: probability from the “V” run. The observed data are taken from Table 1 of Cooke et al. (2010a), which are all lensed binaries (or multiples). Note that the physical separations at the redshift of DLAs are computed correctly consistent with the agreement shown in Figure 22 between the observed and computed size distributions.

### Table 2

| $z$ | $a$ | $b$ |
|-----|-----|-----|
| 3.1 | 1700 | 0.40 |
| 1.6 | 800  | 0.60 |
| 1.0 | 600  | 0.40 |
| 0.0 | 330  | 0.30 |

Finally, in Figure 25, we plot the LOS DLA cross section of each galaxy as a function of halo mass that can be compared directly with other simulations. We see that, within a factor of two, there is a broad agreement among simulations by different groups at the high-mass end ($M_h \geq 10^{12} M_\odot$). While the DLA cross sections of smaller halos ($M_h \leq 10^{12} M_\odot$) in our simulations are larger than those found in the quoted SPH simulations, the larger box “Cosmo” simulation of size $(17.5 h^{-1} \text{ Mpc})^3$ of Pontzen et al. (2008) overlaps with our simulations of sizes $(22.2 h^{-1} \text{ Mpc})^3$ and $(32 h^{-1} \text{ Mpc})^3$. An examination of all simulations reveals an interesting trend: smaller box simulations of isolated dwarf galaxies tend to produce smaller DLA cross sections and have smaller dispersions at a fixed halo mass than larger box simulations. This is most easily seen in Figure 4 of Pontzen et al. (2008), where there is dramatic drop of dispersion from their “Cosmo” (i.e., larger box) simulation to the smaller box isolated galaxy simulations. We do not include the results of Razoumov et al. (2006) in the comparison because of the following difficulty: their fitting curves (their Equation (8)) in their Figure 8 do not include the so-called intergalactic DLAs (i.e., if a DLA is not hosted by a halo in the same base grid cell of size $16-64 h^{-1} \text{ kpc}$ comoving), whereas in the present analysis all DLAs within the virial radius are included. Thus, their fitting curves are underestimates of true DLA cross sections of galaxies, not inconsistent with the fact that their curves are substantially lower than those of Nagamine et al. (2007), shown in their Figure 8. It is interesting to note that some of their so-called intergalactic DLAs are likely what we term galactic aqueducts.

The DLA sizes for some of the dwarf galaxies in our simulations are overestimated, due to the ambiguity of associating a DLA with a galaxy where no clear boundary can be defined. These include cases where a dwarf galaxy is a satellite galaxy interacting with other satellites and/or the primary and is thus associated with a larger DLA structure. We attribute a large portion of the difference between our simulations and all these previous SPH simulations to the difference in the radiative transfer treatments. Our simulations include an approximate method to provide self-shielding of regions that are optically thick to Lyman continuum radiation, and this is done on the fly during the simulation, not a post-processing step. This method provides column density–dependent self-shielding for regions, starting at $N_{\text{HI}} \geq 10^{17} \text{ cm}^{-2}$. At the DLA column density there is complete self-shielding. This treatment of self-shielding has important dynamic effects, especially for smaller halos where appropriate self-shielding allows for cooling and condensation of gas, which further enhances self-shielding. Table 2 gives the parameters to the fits to the median relation in Figure 25.

Let us fast-forward to Figure 36 and note that at $z = 3.1$ the typical halo mass for DLA incidence is $M_h \sim 10^{11}$, having a virial radius of $r_{\text{vir}} = 25 \text{ kpc}$, which then gives $r_{\text{tot}} = 15 \text{ kpc}$ (see above relation between $r_{\text{tot}}$ and $r_{\text{vir}}$).
Figure 19. Maximum $v_{90}$ of all DLAs associated with each galaxy against the halo mass of the galaxy $M_{\text{halo}}$ at $z = 0$ for the “C” and “V” runs. The black line $v_{90} = 2.33v_{\text{vir}}$ is what $v_{90}$ would be if the velocity distribution were an isotropic and Maxwellian distribution with its dispersion equal to $v_{\text{vir}}$, and the Si II gas density is constant across the DLA.

(A color version of this figure is available in the online journal.)

Figure 20. DLA metallicity distributions at four redshifts, $z = 0, 1.6, 3.1, 4.0$, for both the “C” (red histograms) and “V” (green histograms) runs. The observational data are from Prochaska et al. (2005), shown as black histograms. Because there is non-negligible evolution, the comparisons between simulations at a given redshift are only made with observed DLAs within a narrow redshift window, as shown. Probabilities that simulated and observed samples are drawn from the same underlying distribution are indicated in each panel, separately for the “C” and “V” runs.

(A color version of this figure is available in the online journal.)

From Figure 26 we see that the typical distance of DLAs from galaxy center is $d_{\text{DLA}} \sim 25$ kpc, thus $r_{\text{tot}} \sim d_{\text{DLA}}$. This near equality between the distance and size suggests that a large fraction of the sky seen by an object sitting in the galaxy, up to $\pi r_{\text{tot}}^2/4\pi d_{\text{DLA}}^2 \sim 10\%$, may be covered by DLAs! If that object is a QSO, the situation may be different, because the QSO, with its intense radiation (and possibly other effects, such as winds), may have significantly modified its surroundings, including the nearby DLAs. Note that a QSO proximity radius is about 5–10 Mpc. The evidence we see in the gallery...
examples is that the surge in DLA incidence of a galaxy is almost always associated with starbursts and in gas-rich (probably) spiral-like galaxies, whereas QSOs tend to reside in gas-poor elliptical galaxies. These two facts together suggest that it is not straightforward to estimate the fraction of QSOs that have this type of proximity to DLAs. The situation for gamma-ray bursts (GRBs) may be relatively more straightforward to analyze, because GRBs are associated with star formation, hence coinciding with the surge in DLA cross section of interacting galaxies, and because radiation from GRBs does not significantly affect their surroundings at a distance of ~25 kpc. Therefore, we suggest that at high redshift \( z \geq 3 \), a significant fraction (up to 10%) of GRBs, which depends on detailed geometry/orientation of DLA clouds, have associated DLAs with a typical velocity difference from its systemic redshift of \( \Delta v \lesssim 100 \text{ km s}^{-1} \) (see Figure 40 below for \( \Delta v \)). One complication factor is that some (or possibly most of) GRB DLAs may be due to dense gas in the close vicinity of GRBs nearby DLAs. Since GRBs are thought to occur in star-forming regions, i.e., embedded in molecular clouds, one might expect to see a significant amount of molecules associated with nearby DLAs, and they should be dusty and relatively more metal enriched. However, if progenitors of GRBs are biased metal-poor, such as preferred in the collapsar model (MacFadyen & Woosley 1999; Woosley & Heger 2006), then it may become more complicated, although the latest observations indicate otherwise (e.g., Levesque et al. 2010). Thus, we
suggest that a subset of GRB DLAs with little molecular hydrogen column density and being relatively metal-poor and dust-poor, such as those in Hjorth et al. (2003), Tumlinson et al. (2007), and Ledoux et al. (2009), may be identified with circumgalactic DLAs proposed here.

This fraction of GRBs with DLAs is, however, likely to decrease rapidly toward lower redshift because of (1) a smaller \( r_{\text{tot}} / d_{\text{DLA}} \) ratio and (2) a smaller fraction of DLAs outside the stellar disk. The unknown in the above estimate is the geometry of DLAs. We will reserve a detailed analysis of this issue for a future study.

### 3.7. Locations of DLAs and Correlation between DLAs and Galaxies

Several lines of evidence so far suggest that DLAs do not arise predominantly in gaseous disks of spiral galaxies at high redshift, in agreement with Maller et al. (2001) and Hong et al. (2010). Where are they exactly?

In Figure 26 we show the distribution of physical distance of DLAs from the galactic center (i.e., impact parameter in proper physical units) at four redshifts, \( z = 0, 1.6, 3.1, \) and 4.0. Since we have shown in Figure 18 that the DLA incidence...
contribution peaks at $\sim 10^{11.5} M_\odot$, let us make a simple estimate of their size at $z \sim 3$. As a reference, let us take the radius of the Milky Way (MW) stellar disk to be 15 kpc. Taking MW to $z = 3$ self-similarly would give a radius of 3.8 kpc, and for a $10^{11.5} M_\odot$ galaxy the stellar disk radius would be 2.5 kpc at $z = 3$, corresponding to 0.40 in log distance shown in Figure 26. Observed large galaxies (in the total mass range $\sim 10^{11.5} - 10^{12} M_\odot$) at $z \sim 3$ have sizes of $\sim 1$–10 kpc (Lowenthal et al. 1997; Ferguson et al. 2004; Trujillo et al. 2006; Toft et al. 2007; Zirm et al. 2007; Buitrago et al. 2008), roughly consistent with the simple scaling. The distance distribution peaks at $d_{\text{DLA}} \sim 20$–30 kpc at $z = 3$–4, which is much larger than a few kpc of the observed (or expected based on $z = 0$ galaxies) stellar disk size at $z \sim 3$. It is noted that the virial radius of an MW-size galaxy is about $\sim 50$ kpc at $z = 3$. Thus, we conclude that at $z = 3$–4 most of the DLAs do not arise from large galactic stellar disks. They come from regions that are $\sim 5$–8 times larger than the stellar disks and occur at about half the virial radius at $z = 3$. The ubiquitous extended structures—galactic aqueducts—are at the right distances of $d_{\text{DLA}} \sim 20$–30 kpc, seen in the gallery examples in Section 3.1. In Figures 27 and 28, we show 12 randomly chosen examples of galaxies that contain DLAs at $z = 3.1$ and $z = 0$, respectively. We see a striking difference between these two redshifts: every single galaxy at $z = 3.1$ shows an extended H I structure, thanks to strong galaxy–galaxy interactions and high gas density at $z = 3.1$, whereas at $z = 0$ we only see 1 out of 12 that shows a similar, filamentary gaseous structure, apparently also caused by a strong galaxy–galaxy interaction, but it is notable that the filamentary structure is much “slimmer” than all the ones at $z = 3.1$. Taken together, it is clear that higher densities and strong interactions are both responsible for extended, large cross section gaseous structures producing DLAs. Visual inspection of Figures 27 and 28 also confirms the quantitative findings in Figure 26 that, while galactic disks make a large contribution to the overall DLA incidence at $z = 0$, the DLAs are mostly due to more extended gas structures on scales larger than the stellar disks at high redshift.

While the universal association of galactic aqueducts with galaxy interactions suggests that the host galaxies are experiencing starbursts, as seen in the gallery examples, the clouds that give rise to DLAs do not have significant ongoing in situ star formation. Clearly, most DLAs do not arise in disks, and most DLAs that have low metallicities, as we have shown, are self-consistent. In other words, aside from those DLAs that arise from galactic disks and are metal-rich, the vast majority of more metal-poor DLAs are not actively forming stars. We note that the simulations adopt a star formation prescription that is
similar to the Kennicutt–Schmidt law and hence star formation in low-density regions is permitted. Therefore, the gas in the galactic aqueducts forms stars, albeit slowly, determined by the local dynamical time. Figure 29 shows the ratio of gas metallicity for DLAs for different subsets of DLAs with different column density ranges to the mean metallicity of ongoing star-forming gas. It is clear that the high column density ($N_{\text{HI}} \geq 10^{22}$ cm$^{-2}$) DLAs dominate forming stars; most of the DLAs have little star formation. While this particular result in principle may depend on the star formation scheme used, we do not impose a density threshold for star formation in the simulations. When low-density DLA gas is incorporated into disks of galaxies to become high-density gas, it is likely that it either becomes more metal-rich due to increased star formation or is destroyed by significant star formation feedback and removes itself from the DLA category. We suggest that our model gives a natural explanation to the apparent puzzle of the lack of active star formation of gas-rich DLAs at high redshift (Wolfe & Chen 2006). On the other hand, the inferred cooling rates of DLAs may be provided, in part, by radiative heating from the host galaxy (see Figure 35 below) and in part by compression heating, as we frequently see higher external pressure in Section 3.1; we will leave a more quantitative investigation of this issue to a future study.

Returning to Figure 26, at $z = 1.6$ there is a very interesting divergence between the two distributions for the “C” and “V” runs, where the distribution for the “C” run peaks at $d_{\text{DLA}} \sim 40$ kpc and for the “V” run at $d_{\text{DLA}} \sim 10$ kpc. This is consistent with the expectation that the overdense region in the “C” run and the underdense region in the “V” run start to “feel” the difference in their respective local large-scale density environment and evolve differently dynamically. That is, in the “C” run gravitational shock heating due to large-scale structure formation begins to significantly affect the cold gas in galaxies, whereas in the “V” run the galaxies have not changed significantly since $z = 3–4$ except that they are now somewhat smaller due to lower gas density at lower redshift. By $z = 0$ the two distributions once again become nearly identical; this is rather intriguing and may reflect the following physical picture: while galaxies in the “V” run have by now dynamically “caught up” with the field galaxies in the “C” run, giving rise to the similar Gaussian-like distribution centered at $d_{\text{DLA}} = 10$ kpc, the original gas-rich galaxies in the “C” run have fallen into the cluster, lost gas, and “disappeared” from the DLA population. While there is almost no DLA that is farther away than 50 kpc at $z = 3–4$, there is a second bump at $d_{\text{DLA}} = 100–300$ kpc in the distribution for the “C” run at $z = 0$. This bump is probably due to gas-rich satellite galaxies orbiting larger galaxies or small groups of total mass $10^{11}–10^{13} M_\odot$. Beyond $d_{\text{DLA}} = 300$ kpc, there is no DLA in the “C” run, which is due to gas starvation of galaxies in still larger groups or clusters at $z = 0$. With direct
Figure 27. Twelve examples of randomly chosen galaxies that have associated DLAs at $z = 3.1$. The odd rows are the logarithm of the stellar luminosity surface density maps in SDSS $z$ band in units of $L_\odot$ kpc$^{-2}$. The even rows are the logarithm of the corresponding neutral gas column density maps in units of cm$^{-2}$. The lengths are in proper kpc, and the depth of each projection is about the virial diameter of each galaxy.

(A color version of this figure is available in the online journal.)
Figure 28. Same as Figure 27 but for $z = 0$.
(A color version of this figure is available in the online journal.)
Figure 29. Distribution of the ratio of gas metallicity for DLAs at different column density ranges to the mean metallicity of ongoing star-forming gas in the “C” run (left set of four panels) at $z = 0, 1.6, 3.1, 4.0$ and in the “V” run (right set of four panels). We expect that gas that has the $x$-axis value close to or greater than 0 may be forming stars.

(A color version of this figure is available in the online journal.)

Figure 30. Cumulative probability distribution functions for finding a galaxy of the indicated luminosities (rest-frame SDSS $z$-band luminosities of $0.01L^*, 0.1L^*$, and $L^*$) within a proper distance ($x$-axis) for the C and V runs at $z = 0, 1.6, 3.1, 4.0$.

(A color version of this figure is available in the online journal.)
Figure 31. Left panel: \( v_{90} \) as a function of log \( N_{HI} \), assuming \( Z = 0.1 Z_\odot \). Right panel: the percentage of DLAs that have multiple components, as a function of \( v_{90} \).

(A color version of this figure is available in the online journal.)

Figure 32. Left set of four panels: \( f_{mm} \) distributions for the “C” run at redshift \( z = 1.6 \) (top left), \( z = 3.1 \) (top right), and \( z = 4.0 \) (bottom left). For \( z = 1.6 \) we compared to observed DLAs in the redshift range \( z = 1.2–2.2 \); for \( z = 3.1 \) we compared to observed DLAs at \( z = 2.9–3.3 \); for \( z = 4.0 \) we compared to observed DLAs at \( z = 3.8–4.2 \). The observed sample is an updated version of Prochaska & Wolfe (1997), shown as the black histogram. Also shown in each K-S test probability is that the two distributions (computed and observed) are drawn from the same underlying distribution. In the bottom right panel, we compare computed \( z = 1.6 \) and \( z = 3.1 \) distributions along with the K-S test probability to show a significant evolution of this shape distribution function with redshift. Right set of four panels: \( f_{mm} \) distributions for the “V” run.

(A color version of this figure is available in the online journal.)

The inspection of simulation data we find that there are virtually no gas-rich galaxies within the virial radius of the primary cluster in the “C” run.

The peak distance of \( d_{DLA} \sim 10 \) kpc at \( z = 0 \) and visual impression from Figure 28 have made it clear that DLAs at \( z = 0 \) do not extend significantly beyond the normal stellar disks. The right panel of Figure 21 shows the velocity distributions of two subsets of DLAs, divided at metallicity at \( [Z/H] = -1 \) at \( z = 0 \). Here, we see a very clear difference between the two distributions: the higher metallicity subset has large velocity widths, i.e., there is a strong positive correlation between metallicity and velocity width at \( z = 0 \). This supports the picture that a large fraction of DLAs arise in gaseous disks of large field galaxies. Most of the DLAs at \( z = 0 \) have a higher metallicity of \( [Z/H] \geq -1.0 \) with the overall distribution peaking at \( [Z/H] = -0.5 \), also providing support for this picture. Therefore, by \( z = 0 \) the situation has reversed: galactic disks of large galaxies make a major contribution to DLAs at \( z = 0 \). The fact that the peak distance has dropped from 30–40 kpc at \( z = 3–4 \) to 10 kpc at \( z = 0 \) is physically due in part to a large decrease (a factor of \( \sim 100 \)) in the mean gas density of the universe from \( z = 3–4 \) to \( z = 0 \).

Finally, Figure 30 shows a prediction from this model in terms of association between DLAs and galaxies. We see that at \( z = 3.1 \) (60%, 70%) of DLAs in the (C,V) runs are not associated with \( L^* \) galaxies within 100 kpc. That fraction drops to (20%, 30%), if one lowers the luminosity to 0.1 \( L^* \) galaxies. The probability of finding a \( \geq 0.01 L^* \) galaxy within \( \sim 35 \) pc of
a DLA is nearly 100% at \( z = 3 \). At \( z = 1.6 \) the probability of finding an \( L^* \) galaxy within 100 kpc increases to 30%–50%.

### 3.8. \( \text{Si} \, \text{ii} \) Line Profile Shape Measures

Given the overall good agreement with observations with respect to the velocity width distribution, we now turn to shape measures of the \( \text{Si} \, \text{ii} \) absorption profile. Before comparing to observational data from Prochaska & Wolfe (1997), we shall first try to understand the relationship among the optical depth of an \( \text{Si} \, \text{ii} \) line, \( \text{H} \, \text{i} \) column density, metallicity, and velocity width. Assuming that the optical depth profile of the \( \text{Si} \, \text{ii} \) line is a simple top-hat (assuming a different profile such as a Gaussian makes no material difference for our purpose) and \( \text{Si} \, \text{ii} \) abundance follows that of neutral hydrogen, it can be shown that

\[
\tau_{\text{Si} \, \text{ii}} = 0.01 \left( \frac{N_{\text{H} \, \text{i}}}{2 \times 10^{21} \text{ cm}^{-2}} \right) \left( \frac{Z}{Z_{\odot}} \right) \left( \frac{v_{90}}{100 \text{ km s}^{-1}} \right)^{-1},
\]

where \( Z \) is the metallicity of DLA in solar units and all atomic data (frequency, oscillator strength) are taken from Morton (2003). The left panel of Figure 31 shows \( v_{90} \) as a function of log \( N_{\text{H} \, \text{i}} \) for \( Z = 0.1 Z_{\odot} \), based on Equation (2). As expected, an increase in velocity width requires a corresponding increase in column density to produce the same optical depth. More important is that, quantitatively, in order to achieve an optical depth of 0.1, with a width \( v_{90} \approx 100 \text{ km s}^{-1} \), requires a \( \text{H} \, \text{i} \) column of \( \sim 2 \times 10^{22} \text{ cm}^{-2} \), if the DLA is composed of one single component with \( [Z/H] = -1 \). Since the abundance of DLAs with \( N_{\text{H} \, \text{i}} \geq 10^{22.5} \text{ cm}^{-2} \) declines rapidly (see Figure 14) but the abundance of the \( \text{Si} \, \text{ii} \) line peaks near \( v_{90} \sim 100 \text{ km s}^{-1} \) (Figure 16), this suggests that a significant fraction of \( \text{Si} \, \text{ii} \) lines must have multiple components. To quantitatively illustrate this, we define a new simple two-component measure as follows. If there are at least two peaks in the optical depth profile that are separated by more than 0.5\( v_{90} \) and the ratio of the peak heights is greater than 1/15, we define the DLA to be a two-component DLA. The ratio, 1/15, comes about such that the lower peak is guaranteed to be included in the accounting of the \( v_{90} \) interval, although changing it to, say, 1/10 makes no dramatic difference in the results. Note that DLAs with more than two components are included as two-component systems. The right panel of Figure 31 shows the percentage of two-component DLAs as a function of \( v_{90} \). In good agreement with the simple expectation, we see that at \( v_{90} = 100 \text{ km s}^{-1} \), about 50% of DLAs have more than one component, and that number increases to \( \sim 90% \) at \( v_{90} = 300 \text{ km s}^{-1} \). This result is also consistent with the anecdotal evidence shown in the gallery examples in Section 3.1, where most of the large-width DLAs contain more than one physical component.

We now turn to the three kinematic shape measures defined in Prochaska & Wolfe (1997), \( f_{\text{mm}} \), \( f_{\text{edg}} \), and \( f_{2\text{pk}} \), representing, respectively, measures of the symmetry, leading edgeness, and two peakness of the profile of \( \text{Si} \, \text{ii} \) lines associated with DLAs (see the bottom right panels of the gallery pictures in Section 3.1). Figures 32–34 show comparisons of simulation results with observations at three redshifts \( z = 1.6, z = 3.1, \) and \( z = 4.0 \). We see that the overall agreement between simulations and observations is excellent, with K-S tests (indicated in the figures) for both runs ("C" and "V") at three compared redshifts \( z = 1.6, 3.1, 4.0 \) all being at acceptable levels. Our results are in good agreement with one of the models with feedback in Hong et al. (2010), except for the case of \( f_{2\text{pk}} \); our simulations find acceptable K-S test values of 26%–29%, 4%–34%, and 23%–34% at \( z = 1.6, 3.1, \) and \( z = 4.0 \), respectively, whereas they find that none of their models have probability higher than 5% at \( z = 3.1 \). We speculate that the difference in the detailed treatments of the metal transport process and the feedback prescription between our simulations and theirs may have partly contributed to this difference; with detailed metal transport we find very inhomogeneous metallicity distributions across space and among DLAs in our simulations (see Figure 20 below), whereas they assume a constant metallicity of \( [Z/H] = -1 \) for all DLAs. It is also noted that

![Figure 33. Left set of four panels: \( f_{\text{edg}} \) distributions for the "C" run at redshift \( z = 1.6 \) (top left), \( z = 3.1 \) (top right), and \( z = 4.0 \) (bottom left). For \( z = 1.6 \) we compared to observed DLAs in the redshift range \( z = 1.2–2.2 \); for \( z = 3.1 \) we compared to observed DLAs at \( z = 2.9–3.3 \); for \( z = 4.0 \) we compared to observed DLAs at \( z = 3.8–4.2 \). The observed sample is an updated version of Prochaska & Wolfe (1997), shown as the black histogram. Also shown in each K-S test probability is that the two distributions (computed and observed) are drawn from the same underlying distribution. In the bottom right panel, we compare computed \( z = 1.6 \) and \( z = 3.1 \) distributions along with the K-S test probability to show a significant evolution of this shape distribution function with redshift. Right set of four panels: \( f_{\text{edg}} \) distributions for the "V" run. (A color version of this figure is available in the online journal.)

```
Figure 34. Left set of four panels: $f_{2pk}$ distributions for the “C” run at redshift $z = 1.6$ (top left), $z = 3.1$ (top right), and $z = 4.0$ (bottom left). For $z = 1.6$ we compared to observed DLAs in the redshift range $z = 1.2–2.2$; for $z = 3.1$ we compared to observed DLAs at $z = 2.9–3.3$; for $z = 4.0$ we compared to observed DLAs at $z = 3.8–4.2$. The observed sample is an updated version of Prochaska & Wolfe (1997), shown as the black histogram. Also shown in each K-S test probability is that the two distributions (computed and observed) are drawn from the same underlying distribution. In the bottom right panel, we compare computed $z = 1.6$ and $z = 3.1$ distributions along with the K-S test probability to show a significant evolution of this shape distribution function with redshift. Right set of four panels: $f_{2pk}$ distributions for the “V” run.

(A color version of this figure is available in the online journal.)

Figure 35. Cumulative distribution of incidence rate as a function of SFR, for both the C and V runs, compared to their respective total galaxy population at four different redshifts, $z = 0, 1.6, 3.1, 4.0$. For the distributions of the total galaxy populations we assume that the DLA cross section of a galaxy is proportional to its total mass to the two-thirds power, i.e., proportional to its virial radius squared.

(A color version of this figure is available in the online journal.)
Figure 36. Cumulative distribution of incidence rate as a function of halo mass for both the C and V runs, compared to their respective total galaxy population at four different redshifts, $z = 0, 1.6, 3.1, 4.0$. For the distributions of the total galaxy populations we assume that the DLA cross section of a galaxy is proportional to its total mass to the two-thirds power, i.e., proportional to its virial radius squared. The open squares shown in the $z = 1.6$ and $z = 3.1$ panels are the observationally inferred halo mass range for LBGs (Adelberger et al. 2005).
(A color version of this figure is available in the online journal.)

Figure 37. DLA incidence distribution as a function of host H\textsubscript{I} mass for both the C and V runs.
(A color version of this figure is available in the online journal.)
galaxies at any particular time has statistically the same mass function as the galaxy population as a whole. There is, however, a noticeable bias for lower mass galaxies for DLA hosts compared as the galaxy population as a whole. There is, however, a noticeable bias for lower mass galaxies for DLA hosts compared to observations of DLAs in many different aspects. From these detailed comparisons we conclude that the model is consistent with all extant observations of DLAs, including kinematic properties, metallicity, column density distribution, line density and neutral hydrogen density, size distribution, and their evolution, wherever comparisons could be made. Proper modeling of galactic winds along with high numerical resolution has contributed to the success. What is clear is that galactic winds play an indispensable role in alleviating previous tension between the LCDM model and observations, especially in producing large velocity width DLAs and disparate and low metallicities of DLA gas. We now examine the properties of DLA host galaxies.

Figure 36 shows the DLA incidence as a function of the halo mass. Our model makes this rather simple prediction: DLAs closely sample the overall population of galaxies at $z = 0–4$. This is not to say that every galaxy has a DLA in it. Rather, this simply says that the portion of galaxies that give rise to DLAs at any particular time has statistically the same mass function as the galaxy population as a whole. There is, however, a noticeable bias for lower mass galaxies for DLA hosts compared to the general galaxy population, with the bias increasing with decreasing redshift: at $z = 4$ the median mass offset between the total population and DLA hosts is about 0.3–0.5 dex, which increases to almost 1 dex by $z = 0$. At $z = 1.6–4$, the majority of DLAs arise in halos of total mass $M_h = 10^{10}–10^{12} \, M_\odot$, as these galaxies dominate the overall population of galaxies. We expect the clustering of DLAs to be only slightly weaker than that of the overall population of galaxies at $z \gtrsim 3$ and weakens with time.

Based on clustering analyses, the total mass range of the observed LBGs is inferred to be $10^{11.2}–10^{11.8} \, M_\odot$ at $z = 2.9$, $10^{11.3}–10^{12.2} \, M_\odot$ at $z = 2.2$, and $10^{11.9}–10^{12.9} \, M_\odot$ at $z = 1.7$ (Adelberger et al. 2005). Comparing these observationally inferred ranges to the cumulative distributions in Figure 36, we find that 20%–30% of DLAs at $z \sim 3$ may be LBGs and the fraction drops to 10%–20% at $z \sim 1.7$. Schaye (2001) pointed out that, if DLAs have a radius of 27 kpc, LBGs could account

3.9. Properties of DLA Host Galaxies

So far we have presented our simulations and compare to observations of DLAs in many different aspects. From these detailed comparisons we conclude that the model is consistent with all extant observations of DLAs, including kinematic properties, metallicity, column density distribution, line density and neutral hydrogen density, size distribution, and their evolution, wherever comparisons could be made. Proper modeling of galactic winds along with high numerical resolution has contributed to the success. What is clear is that
for all the DLAs at \( z \approx 3 \). Working backward, the estimated 20%–30% of DLAs at \( z \approx 3 \) being LBGs would imply a DLA disk radius of 12–15 kpc for LBGs, which is significantly larger than the size of LBG stellar disks, suggesting that the distribution of neutral DLA gas is in a much more extended region than stellar disk, a combination of some extended gaseous disk and galactic aqueducts. While a significant overlap between DLAs and LBGs is expected, the overall clustering of DLAs is expected to be comparable but slightly weaker than LBGs, since the median halo mass for DLA incidence distribution is slightly smaller than the typical LBG halo mass (a more quantitative comparison will be performed in future work). Møller et al. (2002) conclude that the properties of DLA hosts, including half-light radius, radial profile, optical to near-infrared color, morphology, Ly\( \alpha \) emission equivalent width, and Ly\( \alpha \) emission velocity structure, lie within the measured range for the general population of LBGs. Since those DLA hosts that are sufficiently luminous to be detected are large galaxies in the total mass range of LBGs and because DLAs tend to arise in gaseous galactic aqueducts of galaxies that experience starbursts (via galaxy interactions; see Section 3.1), like LBGs, it is naturally expected in our model that they find similarity between their DLA hosts and LBGs. Of the remaining 70%–80% that are not covered by LBGs at \( z = 3–4 \), roughly 10%–20% are due to more massive galaxies, and 50%–70% are from smaller galaxies.

When systems become more dynamically advanced with time and large galaxies are no longer abundant with cold gas, a slightly more significant shift occurs between DLA hosts and the overall population of galaxies at \( z = 0 \). At \( z = 0 \) we see that the contribution to the DLA population is broadly peaked at \( M_h = 10^{12} M_\odot \), whereas most of the sum of the square of virial radius is also broadly distributed but, unlike for DLA host, having a significant contribution from halos of total mass \( M_h = 10^{12}–10^{13} M_\odot \) in the “C” run; for the “V” run most of DLAs are in halos of total mass \( M_h = 10^{10}–10^{12} M_\odot \), whereas most of the sum of the square of virial radius comes from halos of total mass \( M_h = 10^{11}–10^{13} M_\odot \).

Figure 35 shows the DLA incidence distribution as a function of host SFR and the corresponding notional distribution if the total DLA cross section of a galaxy is proportional to its stellar mass to the two-thirds power, i.e., proportional to its virial radius squared.

Figure 39. Cumulative distribution of incidence rate as a function of SDSS rest-frame \( g - r \) color for both the C and V runs, compared to their respective total galaxy population at four different redshifts, \( z = 0, 1.6, 3.1, 4.0 \). For the distributions of the total galaxy populations we assume that the DLA cross section of a galaxy is proportional to its total mass to the two-thirds power, i.e., proportional to its virial radius squared.

(A color version of this figure is available in the online journal.)
the found range above. The trend of moving to lower SFR galaxies at lower redshift is perhaps due to the transition from frequent at high redshift \((z = 3–4)\) to less frequent galaxy-galaxy interactions and lack of cold gas in very large galaxies at \(z = 0\). All evidence—size (Figure 22), distance (Figure 26), metallicity (Figure 20), and now SFR (Figure 35)—points to the picture that by \(z = 0\) most DLAs arise in disks of large galaxies of total mass \(10^{11}–10^{12.5} M_\odot\), that are relatively quiescent. There is a significant (~25%) galaxy population at \(z \approx 0\) that has very low SFR and does not host DLAs; these are the gasless elliptical galaxies, most of which reside in the clusters of galaxies. That population is absent in the V run.

Figure 37 shows the DLA incidence distribution as a function of host \(\text{H}1\) mass. It is interesting that the \(\text{H}1\) mass peaks at \(10^{9.5}–10^{10} M_\odot\) at all redshifts with \(10^9–10^{11} M_\odot\) covering nearly all the contributions. The slight shift to a lower \(\text{H}1\) mass for DLA hosts in the V run compared to DLA hosts in the C run is simply due to the fact that halos in the V runs are smaller than those in the C run (see Figure 36). Why the peak position of \(M_{\text{H}1}\) changes little with time is intriguing and needs to be further explored.

Figure 38 shows the DLA incidence distribution as a function of host \(\text{H}1\) mass-to-stellar mass ratio. We see that at \(z = 3–4\) the majority of galaxies are gas-rich \((M_{\text{gas}}/M_{\text{star}} \geq 0.1)\) and DLA hosts closely follow the general population of galaxies, consistent with what we saw in Figure 36. This statement remains true for the “V” run until \(z = 0\), whereas for the regions that are more dynamically advanced, i.e., \(z = 0\) in the “C” run, a significant segregation is evident in that most of the galaxies are now gas-poor \((M_{\text{gas}}/M_{\text{star}} \leq 0.1)\) but a majority of DLA hosts are still gas-rich \((M_{\text{gas}}/M_{\text{star}} \geq 0.1)\). Moreover, at \(z = 0\) in the “C” run, there is a population of old, red and “dead” galaxies with virtually no cold gas; these galaxies are mostly in clusters of galaxies, which DLAs largely avoid.

Figure 39 shows the DLA incidence distribution as a function of host SDSS \(g - r\) color. While the individual color of the simulated galaxies may not be totally consistent with observations due to its sensitivity to a relatively small variation in the amount of recent star formation that the simulation may not necessarily properly capture, a comparative statement is still valid statistically. We see that DLA hosts trace the general population of galaxies at \(z = 3–4\) in all environments with a slight bias to a bluer color peaking sharply at \(g - r \sim 0\). At lower redshift DLA hosts become significantly bluer than average galaxies by roughly \(\Delta(g - r) \sim 0.2\) and they avoid the reddest galaxies that are presumably elliptical galaxies or galaxies in clusters of galaxies.

Finally, in Figure 40 we show the distribution of the LOS velocity difference between the DLA and its host, for “C” (red histograms) and “V” (green histograms) runs. As shown in blue (for “C” run) and black (for “V” run) are fits to the histograms with the dispersion parameters \(\sigma\) indicated.

\[
\text{PDF}(v_{\text{gal}} - v_{\text{DLA}}) = \frac{1}{2\sigma} \exp\left(-\frac{|v_{\text{gal}} - v_{\text{DLA}}|}{\sigma}\right).
\]

We find that the dispersion parameter \(\sigma\) only increases mildly with decreasing redshift from \(\sigma \sim 75–90\) km s\(^{-1}\) at \(z = 3–4\) to \(\sim 90–110\) km s\(^{-1}\) at \(z = 0\). The relatively small velocity dispersion and a symmetric distribution may make DLAs a useful proxy for the systemic velocity of host galaxies. This property may be utilized to measure galactic wind velocities, using a large statistical sample of DLA hosts with measured interstellar absorption lines, such as the Na I line (e.g., Heckman et al. 2000).
In summary, we show that, at $z = 3–4$, DLA hosts approximately follow the general population of galaxies with respect to halo mass, SFR, H\textsc{i} mass, gas mass–to–stellar mass ratio, and color, with a small but noticeable tendency of being more gas-rich, slightly smaller, and bluer. It is seen that LBGs make up about 20%–30% of the overall DLAs at $z = 3–4$, decreasing to 10%–20% by $z = 1.7$. Of the remaining 70%–80% that are not covered by LBGs at $z = 3–4$, roughly 10%–20% are due to galaxies more massive than LBGs, and 50%–70% are from smaller galaxies. The comparisons at lower redshifts become not as clean-cut as at high redshift, but in a way that is physically understandable. Specifically, at lower redshifts, while the general population of galaxies gradually become less cold and gas-rich, DLA hosts seem to “pick out” those that are gas richer and bluer, but not necessarily of higher SFR.

4. CONCLUSIONS

With high resolution and a physically sound treatment of all relevant physical processes, our state-of-the-art, AMR Eulerian cosmological hydrodynamic simulations reproduce a whole array of observables of DLA systems, including the distributions of the following quantities and their evolution with redshift: column density, S\textsc{iii} λ1808 absorption line velocity width, kinematic shape measures of the S\textsc{ii} line, metallicity, size, line density, and H\textsc{i} mass density. This allows us to examine, in addition, with significant confidence, the properties of DLA host galaxies. We are able to reach the following conclusions:

1. DLA hosts roughly trace the overall population of galaxies at all redshifts, with respect to halo mass, SFR, H\textsc{i} mass, and color, with a noticeable tendency of being smaller and bluer, on average. Most DLA hosts have H\textsc{i} mass peaked at $10^{10} M_\odot$ at all redshifts with $10^{9}–10^{11} M_\odot$ covering nearly all DLA hosts. The majority of DLA hosts are gas-rich ($M_{\text{gas}}/M_{\text{star}} \geq 0.1$), closely following the general population of galaxies at high redshift; by $z = 0$ most of the stars are in systems with ($M_{\text{gas}}/M_{\text{star}} \leq 0.1$), but the majority of DLA hosts are still gas-rich ($M_{\text{gas}}/M_{\text{star}} \geq 0.1$).

2. The history of DLA evolution is cosmological in nature and reflects the underlying cosmic density evolution, galaxy evolution, and galaxy interactions. With higher density and more interactions at high redshift, DLAs are larger in both absolute terms and relative terms with respect to virial radii of halos. At $z = 3–4$ galactocentric distance of DLAs is, on average, $0.6 r_{\text{vir}}$, whereas at $z = 0$ it decreases to $0.08 r_{\text{vir}}$. The typical DLA impact parameter is $d = 20–30$ kpc at $z = 3–4$ and 10 kpc at $z = 0$. The typical size (radius) of individual DLAs is $\sim 10$ kpc at $z \sim 3–4$, dropping to $\sim 5$ kpc at $z \sim 1.6$, in agreement with observations, then increasing to $\sim 7$ kpc at $z = 0$.

3. The variety of DLAs at high redshift is richer with a large contribution coming from galactic aqueducts—filamentary gaseous structures sometimes extending to as far as virial radius—that are created through close galaxy interactions. The portion of gaseous disks of galaxies where most stars reside makes a relatively small contribution to DLA incidence at $z = 3–4$. By $z = 0$ galaxy interactions have become rare, and cosmic mean density has decreased.

**Figure 41.** Comparison of DLA column density distributions at $z = 2.5$ for the “C” run (solid dots) and “C/2” run (open squares). (A color version of this figure is available in the online journal.)

**Figure 42.** Comparison of DLA velocity width distributions at $z = 2.5$ for the “C” run (solid dots) and “C/2” run (open squares). (A color version of this figure is available in the online journal.)

**Figure 43.** Comparison of DLA metallicity distributions at $z = 2.5$ for “C” run (red histogram) and “C/2” run (green histogram). (A color version of this figure is available in the online journal.)
In agreement with observations, we see a weak but noticeable contribution to DLA incidence is gas-rich disks of galaxies at low redshift, which explains the significant increase of metallicity.

7. The SFR of DLA hosts is, however, quite strong, heavily concentrated in the range $0.3–30 M_\odot \text{yr}^{-1}$ at $z = 3–4$, gradually shifting lower to peak at $\sim 0.5–1 M_\odot \text{yr}^{-1}$ by $z = 0$. The finding that the typical size and galactocentric distance of DLAs are comparable gives a moderate, “apparent” SFR seen in Ly$\alpha$ emission due to fluorescence of ionizing photons from the host galaxies. Both the size of Ly$\alpha$ emission and the magnitude of this fluorescence are, curiously, consistent with the population of faint Ly$\alpha$ emitters observed by Rauch et al. (2008), if they are at $z \sim 3$.

8. Finally, we predict that only 20%–30% of DLAs are within 100 kpc of $L^*$ galaxies at $z = 3–4$, increasing to 30%–50% at $z = 1.6$.

I thank Dr. M. K. R. Joung for help on generating initial conditions for the simulations and running a portion of the simulations and Greg Bryan for help with Enzo code. I also thank an anonymous referee for detailed and constructive reports. I thank Dr. Jason X. Prochaska, Dr. Andrew Pontzen, and Dr. P. Noterdaeme for kindly providing the observational data, Dr. Sara L. Ellison for sending a paper draft before publication and for discussion, Dr. Edward Jenkins for discussion on atomic data, and Dr. Jerry Ostriker for many useful comments.

Computing resources were in part provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center. This work is supported in part by grants NNX08AH31G and NAS8-03060. The simulation data are available from the author on request.

APPENDIX

RESOLUTION CONVERGENCE TESTS

In Figure 41, we show a comparison of DLA column density distributions for the “C” and “C/2” runs at $z = 2.5$. Recall that the “C/2” run has a spatial resolution lower than the “C” run dramatically, so what is left to contribute to DLA incidence is gas-rich disks of galaxies in the total mass range of $10^{10}–10^{12} M_\odot$.

4. Galactic winds play an indispensable role in shaping the kinematic properties of DLAs. Specifically, the high velocity width DLAs are a mixture of those arising in high-mass, high velocity dispersion halos and those arising in smaller mass systems where cold gas clouds are entrained to high velocities by galactic winds. Closer examination reveals that most of the large width DLAs are due to multiple physically distinct components of varying LOS velocities within a 100 kpc separation (along the LOS).

5. Quantitatively, the majority of DLAs arise in halos of total mass $M_h = 10^{10}–10^{12} M_\odot$ at $z = 1.6–4$, as these galaxies dominate the overall population of galaxies then. At $z = 3–4$, 20%–30% of DLA hosts are LBGs, 10%–20% are due to galaxies more massive than LBGs, and 50%–70% are from smaller galaxies. The fraction of LBG DLA hosts drops to 10%–20% by $z \sim 1.7$.

6. In agreement with observations, we see a weak but noticeable evolution in DLA metallicity. The metallicity distribution centers at $[Z/H] = -1.5$ to $-1$ and spans more than three decades at $z = 3–4$, with the peak moving to $[Z/H] = -0.75$ at $z = 1.6$ and $[Z/H] = -0.5$ by $z = 0$. The overall metallicity is floored at $[Z/H] \sim -3$ at $z = 1.6–4$ and at $[Z/H] \sim -1.5$ at $z = 0$. When most of the DLA incidence is due to cold gas outside the stellar disks at high redshift, not only is the metallicity relatively low, but also there is little star formation within. This explains self-consistently their low metallicity and lack of star formation activity. If and when significant internal star formation occurs, there are two possible scenarios. If that takes place outside the galactic disk, the DLA will be destroyed and remove itself from the DLA category. If it takes place on the galactic disk, the DLA will either be destroyed permanently or temporarily and then cool back down to become a more metal-enriched one, after star formation has stopped. The former occurs predominantly at high redshift, which is why most DLAs are metal-poor. The latter occurs at low redshift, which explains the significant increase of metallicity.
by a factor of two. While the agreement is not perfect, it is reasonably good and the difference is at 10%–30% across the dynamic range. Figure 42 compares velocity width ($v_{90}$) column density distributions for the “C” and “C’/2” run at $z = 2.5$, where we see that the level of agreement is comparable to that seen in Figure 41. Figure 43 shows a comparison of the metallicity distribution for the “C” and “C’/2” run at $z = 2.5$. It is seen that the metallicity distribution agrees to within about 0.2 dex and it appears that a still higher resolution simulation might give a somewhat lower metallicity, perhaps in still better agreement with observations. Figure 44 shows a comparison of the DLA size distribution for the “C” and “C’/2” runs at $z = 2.5$. Since the spatial resolution more directly affects the size, here we see the worst disagreement between the two runs in terms of the shape of the distribution, but overall still mostly at a level of ~0.05–0.1 dex. However, a direct comparison of probability for QSO binary sightlines shown in Figure 45 suggests that the overall convergence is quite good.

REFERENCES

Abel, T., Anninos, P., Zhang, Y., & Norman, M. L. 1997, New Astron., 2, 181
Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697
Arieli, Y., Rephaeli, Y., & Norman, M. L. 2008, ApJ, 683, L111
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bryan, G. L. 1999, Comput. Sci. Eng., 1, 46
Bryan, G. L., & Norman, M. L. 2000, Institute for Mathematics and Its Applications, 117, 165
Cen, R. 2011, ApJ, 741, 99
Cen, R., Kang, H., Ostriker, J. P., & Ryu, D. 1995, ApJ, 451, 436
Cen, R., & Ostriker, J. P. 2005, ApJ, 635, 86
Cen, R., & Ostriker, J. P. 1992, ApJ, 391, 502
Ellison, S. L., Prochaska, J. X., Hennawi, J., et al. 2010, MNRAS, 406, 1435
Ferguson, H. C., Dickinson, M., Giavalisco, M., et al. 2004, ApJ, 600, L107
Gardner, J. P., Katz, N., Weinberg, D. H., & Hernquist, L. 1997b, ApJ, 486, 42