

Damped Lyman- α absorbers as a probe of stellar feedback

Simeon Bird^{1*}, Mark Vogelsberger², Martin Haehnelt³, Debora Sijacki³, Shy Genel⁴, Paul Torrey⁴, Volker Springel^{5,6}, Lars Hernquist⁴

¹*Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ, 08540, USA*

²*MKI and Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

³*Kavli Institute for Cosmology and Institute for Astronomy, Madingley Road, Cambridge, CB3 0HA, UK*

⁴*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

⁵*Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnengasse 35, 69118 Heidelberg, Germany*

⁶*Zentrum für Astronomie der Universität Heidelberg, ARI, Mönchhofstr. 12-14, 69120 Heidelberg, Germany*

ABSTRACT

We examine the abundance, clustering and metallicity of Damped Lyman- α Absorbers (DLAs) in a suite of hydrodynamic cosmological simulations using the moving mesh code AREPO. We incorporate models of supernova and AGN feedback, as well as molecular hydrogen formation. We compare our simulations to the column density distribution function at $z = 3$, the total DLA abundance at $z = 2 - 4$, the measured DLA bias at $z = 2.3$ and the DLA metallicity distribution at $z = 2 - 4$. Our preferred models produce populations of DLAs in good agreement with most of these observations. The exception is the DLA abundance at $z < 3$, which we show requires stronger feedback in $10^{11-12} h^{-1} M_{\odot}$ mass halos. While the DLA population probes a wide range of halo masses, we find the cross-section is dominated by halos of mass $10^{10} - 10^{11} h^{-1} M_{\odot}$ and virial velocities $50 - 100 \text{ km s}^{-1}$. The simulated DLA population has a linear theory bias of 1.7, whereas the observations require 2.17 ± 0.2 . We show that non-linear growth increases the bias in our simulations to 2.3 at $k = 1 h \text{ Mpc}^{-1}$, the smallest scale observed. The scale-dependence of the bias is, however, very different in the simulations compared against the observations. We show that, of the observations we consider, the DLA abundance and column density function provide the strongest constraints on the feedback model.

Key words: cosmology: theory – intergalactic medium – galaxies: formation

1 INTRODUCTION

Damped Lyman- α Systems (DLAs) are strong Lyman- α absorption features in quasar spectra. At the high densities necessary to produce neutral hydrogen column densities above the threshold for DLAs, $N_{\text{HI}} > 10^{20.3} \text{ cm}^{-2}$ (Wolfe et al. 1986), hydrogen is self-shielded from the ionising effect of the diffuse radiation background and so is neutral (Katz et al. 1996b). Because lower redshifts can only be observed from space, the bulk of observed DLAs lie at $z = 2 - 4$, where they contain a large fraction of the neutral hydrogen in the Universe (Gardner et al. 1997). DLAs are a direct probe of the distribution of neutral gas at a mean density around one tenth of the star formation threshold (Cen 2012). Thus, they provide a powerful independent check on models of galaxy formation, with the unique advantage that they

do not directly depend on the star formation model. The exact nature of the systems hosting DLAs has historically been uncertain, with kinematic data combined with simple semi-analytic models appearing to favour large rotating discs (Prochaska & Wolfe 1997; Jedamzik & Prochaska 1998; Maller et al. 2001), while early simulations produced small proto-galactic clumps (Haehnelt, Steinmetz & Rauch 1998; Okoshi & Nagashima 2005). However, Barnes & Haehnelt (2009) showed that the kinematic data could be reproduced if DLAs are found in halos with a virial velocity of $50 - 200 \text{ km s}^{-1}$.

We have performed a series of cosmological hydrodynamic simulations with the moving mesh code AREPO (Springel 2010). AREPO performs well in idealised tests (Sijacki et al. 2012) and yields galaxies with disk-like morphologies (Torrey et al. 2012). Compared to its predecessor GADGET, it follows the accretion of cold gas onto the central regions of halos significantly more accurately (Nelson et al.

* E-mail: spb@ias.edu

2013) and leads to a smoother distribution of Lyman Limit Systems (LLS) (Bird et al. 2013). We incorporate prescriptions for supernova and AGN feedback, based on the implementation presented in Vogelsberger et al. (2013) (V13) and Torrey et al. (2014). Our simulations have been post-processed to produce a simulated DLA population, which we compare to observations of the column density function, DLA metallicity, redshift evolution and bias. These observations strongly constrain the stellar feedback model, and confirm the need for strong stellar feedback even at high redshifts.

A wide range of quasar surveys have over time increased the available sample of high-redshift DLAs (e.g. Wolfe et al. 1995; Storrie-Lombardi & Wolfe 2000; Péroux et al. 2005; Prochaska, Herbert-Fort & Wolfe 2005; Noterdaeme et al. 2012), and constrained the distribution of systems at lower column densities (Péroux et al. 2001; O’Meara et al. 2007; Prochaska & Wolfe 2009; Rudie et al. 2013; Kim et al. 2013; Zafar et al. 2013). Previous simulation work in this subject is also extensive. Pontzen et al. (2008) reproduced the observed DLA metallicity and abundance in SPH simulations with a model incorporating supernova feedback and simple radiative transfer. The effect of galactic winds on DLA abundance and metallicity in SPH simulations was also examined in Nagamine, Springel & Hernquist (2004a); Nagamine, Springel & Hernquist (2004b); Nagamine et al. (2007); Tescari et al. (2009). Similar studies have been done with grid-based codes (Fumagalli et al. 2011; Cen 2012). Faucher-Giguère et al. (2010); Altay et al. (2011); Yajima, Choi & Nagamine (2012); Rahmati et al. (2013a) studied the effects of radiative transfer on DLA self-shielding. Molecular hydrogen formation was considered in Erkal, Gnedin & Kravtsov (2012) and Altay et al. (2011). Davé et al. (2013) used a similar feedback scheme to this work, and examined the galactic HI mass function. Altay et al. (2013) considered the column density distribution within the models examined by the OWLS project, while Rahmati & Schaye (2014) looked at the halo hosts of both DLAs and Lyman Limit Systems (LLSs). Razoumov (2009) examined the kinematic distribution of DLA metal lines in AMR simulations of isolated halos. Cen (2012) presented simulations of one over-dense and one under-dense region, which bracketed many observable properties of DLAs, including the kinematic distribution of metal lines, the metallicity and the column density function.

In this paper, we compare DLAs from a hydrodynamic simulation to recently released data for the column density distribution and DLA abundance (Noterdaeme et al. 2012), DLA metallicity (Rafelski et al. 2012) and the bias of DLAs (Font-Ribera et al. 2012). We use a suite of models based around that of V13, both providing a valuable consistency check on the model and demonstrating the constraints from each observation. This paper is structured as follows. In Section 2, we discuss our methods. Section 3 looks at the properties of our simulated DLAs, and Section 4 compares with observations. We conclude with a summary of our findings in Section 5.

Name	v_w^m (km s ⁻¹)	AGN	κ_w	Notes
DEF	0	Yes	3.7	As V13
WMNOAGN	0	No	3.7	Warm winds
HVEL	600	Yes	3.7	
HVNOAGN	600	No	3.7	
2xUV	0	Yes	3.7	2× UVB amplitude
FAST	0	Yes	5.5	
NOSN	-	No	-	No feedback

Table 1. Simulation parameters varied. v_w^m is the minimum wind velocity and κ_w is the coefficient between local velocity dispersion and wind velocity.

2 METHODS

In this Section we give only a brief overview of our simulation methods and subgrid feedback prescriptions, focussing instead on our post-processing. Details of the moving mesh implementation in AREPO may be found in Springel (2010), while the TreePM approach used to compute gravitational interactions is described in Springel (2005). The framework for supernova and AGN feedback, star formation and metal enrichment within which our simulations are run is presented fully in V13. All our analysis scripts are publicly available at https://github.com/sbird/DLA_script, implemented in Python and C++.

2.1 Parameters of the Simulations

2.1.1 Supernova Feedback

Supernova feedback in our simulations, following V13 and Springel & Hernquist (2003), is modelled by the injection of kinetic energy from star-forming regions into the surrounding gas. This is implemented by stochastically spawning wind particles from gas elements with $\rho \geq 0.13 \text{ cm}^{-3}$, the star formation threshold density. Wind particles interact gravitationally, but are hydrodynamically decoupled. Once they reach a lower density region or a travel time threshold, their energy, momentum, mass and metal content are added to the gas cells at their current locations.

Table 1 summarises the parameters of our feedback model. DEF is the reference simulation of V13, while our other simulations change the parameters listed. As a check on the effects of feedback, we include a simulation with metal enrichment but lacking supernova or AGN feedback. In our other simulations, the wind energy per unit stellar mass is a function of the wind velocity, v_w^2 , and mass loading, η_w

$$\text{egy}_w = \frac{1}{2} \eta_w v_w^2 \quad (1)$$

$$= 1.09 \text{ egy}_w^0 \quad (2)$$

where $\text{egy}_w^0 = 1.73 \times 10^{49} \text{ erg } M_\odot^{-1}$ is the expected available supernova energy per stellar mass.¹ In order to obtain the large mass loadings necessary to adequately suppress star formation in low-mass halos (Okamoto et al. 2010; Puchwein

¹ Note that the values of these parameters are incorrectly listed in V13; see Vogelsberger et al. (2014a).

et al. 2012), v_w is allowed to depend linearly on the local DM velocity dispersion, $\sigma_{\text{DM}}^{\text{1D}}$. Thus we have

$$v_w = \kappa_w \sigma_{\text{DM}}^{\text{1D}}. \quad (3)$$

Oppenheimer & Davé (2008) showed that this correlates with the virial velocity of the host halo; smaller halos thus produce winds at a lower velocity and so a larger mass loading. We consider two wind speeds; $\kappa_w = 3.7$, following V13, and, to test the effect of faster winds, $\kappa_w = 5.5$. The latter produces a wind velocity mid-way between the default model and the fast winds simulation of V13, in which star formation at $z = 0$ is almost completely suppressed.

To evaluate the extent to which the increased mass loading in small halos is affecting the DLA population, we run a simulation where the winds have a minimum velocity of 600 km s^{-1} . This keeps the wind velocity constant for the halos which dominate the DLA cross-section and increases it when compared to our default model. We also considered a minimum wind velocity of 200 km s^{-1} , and found that it produced an effect similar to changing the UVB amplitude, described below.

Following Marinacci, Pakmor & Springel (2014) we considered depositing 40% of the supernova energy in thermal, rather than kinetic, energy, producing warmer winds. This had negligible effect on the simulated DLA population. It is likely that within the high density gas which produce DLAs the extra thermal energy is quickly dissipated by cooling.

In order to match the mass-metallicity relation, V13 allow the wind metal loading to vary independently of the wind mass loading, so that the wind metallicity is given by $Z_w = \gamma_w Z_{\text{ISM}}$. We adopt $\gamma_w = 0.4$, following V13. We have checked explicitly by running a simulation with $\gamma_w = 0.7$ that none of our results are sensitive to this parameter.

2.1.2 AGN Feedback

AGN feedback is implemented following Springel, Di Matteo & Hernquist (2005) and Sijacki et al. (2007). Massive halos contain accreting black hole particles, which affect their surroundings in three different ways; quasar-mode, radio mode and by suppressing cooling. The radio mode is the only one which has a significant effect on the surrounding galactic environment. Here thermal energy is periodically dumped into the gas around the black hole, allowing sufficient energy to be released in a single burst to overcome the short cooling times of the dense gas. We will show that the effect of AGN feedback on the DLAs is in most circumstances small, so we will not explore the parameter space of the AGN model further in this work.

2.1.3 Cooling and Enrichment

We include radiative cooling implemented using a rate network, including line cooling, free-free emission and inverse Compton cooling off the cosmic microwave background (Katz, Weinberg & Hernquist 1996a). The UV background (UVB) in most of our simulations follows the estimates of Faucher-Giguère et al. (2009), which was chosen to match constraints from the mean temperature and opacity in the Lyman- α forest (Faucher-Giguère et al. 2008b,a). However, Becker & Bolton (2013) find a photo-ionisation rate larger

by a factor of two, due primarily to an improved measurement of the mean IGM temperature (Becker et al. 2011), rather than a change in the measured mean optical depth. We perform one simulation which doubles the amplitude of the UVB while leaving its evolution unchanged. This produces a photo-ionisation rate in reasonable agreement with the results of Becker & Bolton (2013). We assume a temperature-density parameter $\gamma = 1.56$, which matches the Doppler widths of Lyman- α lines (Rudie, Steidel & Pettini 2012; Bolton et al. 2014), and is roughly the same as that produced by the simulations².

The gas in our simulations is enriched by mass return from star particles. The formation rate and yield of AGB stars, type Ia and type II supernovae are computed for each particle using a Chabrier (2003) IMF, as detailed in V13. Metals are distributed into the gas cells surrounding the star using a top-hat kernel with a radius chosen to enclose a total mass equal to 256 times the average mass of a gas element. We have checked that distributing the metal within a radius enclosing 16 times the gas element mass gives the same result.

2.1.4 Initial Conditions

All simulations use the same initial conditions, originally generated for V13. The initial redshift is $z = 127$, and a linear theory power spectrum used was obtained from CAMB with cosmological parameters $\Omega_m = 0.2726$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.0456$, $\sigma_8 = 0.81$, $n_s = 0.963$ and $H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ($h = 0.704$). The box size is $25 h^{-1} \text{ Mpc}$. The initial conditions include 512^3 dark matter particles with a particle mass of $7.33 \times 10^6 h^{-1} M_\odot$, and 512^3 gas elements with an initial mass of $1.56 \times 10^6 h^{-1} M_\odot$. The gas elements are refined and derefined throughout the simulation to keep their mass roughly constant. The dark matter has a fixed comoving gravitational softening length of $1.0 h^{-1} \text{ kpc}$.

2.2 Neutral Hydrogen Fraction

2.2.1 Self-Shielding

At $z < 5$, most gas is in ionisation equilibrium with the UVB and thus highly ionised. However, DLAs are sufficiently dense that they are able to shield themselves from the ionising photons and remain neutral. Self-shielding becomes important at a relatively sharp hydrogen density threshold $n_{\text{H}} \sim 10^{-2} \text{ cm}^{-3}$ (e.g. Haehnelt et al. 1998). Most neutral gas in DLAs has a density significantly above the threshold, so they are only weakly affected by the details of the transition (Katz et al. 1996b). We use the fitting formula provided by Rahmati et al. (2013a) to compute the self-shielding correction to the photo-ionisation rate as a function of the hydrogen density and gas temperature. Star forming gas is assumed to have a temperature of 10^4 K .

Some authors (e.g. Fumagalli et al. 2011; Rahmati et al. 2013b) find that local stellar radiation may be significant for absorbers with $N_{\text{HI}} \gtrsim 10^{21} \text{ cm}^{-2}$. However, Pontzen et al. (2010) find that the effect is negligible. Stellar sources are

² γ is defined by fitting the IGM equation of state with $\rho \propto T^{\gamma-1}$

surrounded by dense photoionised gas with a short recombination time and an uncertain, but small, escape fraction for ionising photons. Calculating this accurately involves physics on parsec scales, far below the resolution of our simulations. We therefore neglect the effect, noting that it can be viewed as part of the unresolved physics included in our feedback prescription.

2.2.2 Molecular Hydrogen

We include a prescription for the formation of molecular hydrogen, following Altay et al. (2011), based on the relationship between the surface density of molecular hydrogen and the mid-plane hydrostatic pressure found by Blitz & Rosolowsky (2006). Here the pressure from the effective equation of state in star-forming gas is used as a proxy for the hydrostatic pressure. Non star-forming gas does not form molecular hydrogen. The molecular fraction in star-forming gas is given by:

$$f_{\text{H}_2} = \frac{\Sigma_{\text{H}_2}}{\Sigma_{\text{H}}} = \frac{1}{1 + (35(0.1/n_{\text{H}})^{5/3})^{0.92}}, \quad (4)$$

where n_{H} is the hydrogen density in cm^{-3} .

Blitz & Rosolowsky (2006) measured molecular fractions for systems with metallicities as low as $0.1Z_{\odot}$, and found no significant dependence on metallicity, although this is expected theoretically (e.g. Schaye 2001). Erkal et al. (2012) suggested that this relation therefore may not hold in the metal-poor dwarf galaxies that make up DLAs in their simulations. We checked the metallicity histogram of the star forming cells, weighted by H_2 mass, in our simulation and found that 85% of the H_2 was in particles with metallicity $Z > 0.1Z_{\odot}$. This suggests that any suppression of H_2 formation at low metallicity would affect the total amount of H_2 formed at the 10% level at most, giving negligible effect on the column density distribution function (CDDF). Furthermore, Fumagalli, Krumholz & Hunt (2010) found that on the scales relevant to our simulations the Blitz & Rosolowsky (2006) relation is valid even for very metal-poor gas.

2.3 Projection and DLA Identification

The rarity of high density columns (due to our small box size, < 1% of our spectra will contain a DLA) necessitates a large number of sightlines to collect a representative sample of DLAs. We project all particles in the volume onto a regular grid covering the entire box, treating each grid cell as a sightline. Each grid cell has a linear size of $1.5 h^{-1} \text{kpc}$, approximately the gravitational softening length, and thus the projected box contains 16384^2 cells.

Gas elements are projected onto the grid using an SPH kernel. The SPH smoothing length is chosen so that the kernel support equals the volume of the moving mesh gas cell. Should multiple DLAs be present in the same cell, but at different depths along the projection axis, they may be unphysically blended together. We checked that for a $25 h^{-1} \text{Mpc}$ box this did not have a significant affect on the statistics of DLAs. However, it did become difficult to distinguish the boundaries of systems visually, as projection effects caused some large systems to appear close together. Therefore, we computed the projected neutral hydrogen column density in

ten projected slices of $2.5 h^{-1} \text{Mpc}$ width. This treats each cell as a single quasar sightline containing ten pixels evenly spaced across the box. We verified that our results were converged with respect to both the grid cell size and the slice width at the 5% level. The column density, for a projection along the x axis, is thus

$$N_{\text{HI}} = \frac{(1+z)^2}{m_{\text{P}}} \int dx \rho_{\text{HI}}(x), \quad (5)$$

where m_{P} is the proton mass and $\rho_{\text{HI}}(x)$ is the neutral hydrogen density. The factor of $(1+z)^2$ enters because ρ_{HI} is in comoving units and N_{HI} is in physical units. A DLA is defined to be an absorber with $N_{\text{HI}} > 10^{20.3} \text{cm}^{-2}$.

2.4 DLA Host Halos

DLAs are typically associated located within a host halo. To identify this halo, we must first find the position of each DLA in the projection direction. This we define as the HI weighted average depth of the particles within each pixel. For reasons of speed and numerical stability, we use only particles with a neutral fraction greater than 10^{-3} . If the DLA lies within the virial radius of a halo, it is assigned to that halo. If a DLA instead lies within the half-mass radius of a subhalo, which is the case for 5 – 10% of DLAs, we assign it to that subhalo's parent. The $\sim 5\%$ of DLAs which are not within the virial radius of any halo or subhalo are classed as field DLAs.

2.5 DLA Bias

The linear DLA bias, b_{DLA} , is defined as

$$P_{\text{DLA}}(k, z) = b_{\text{DLA}}^2(k, z) P_{\text{M}}(k, z), \quad (6)$$

where P_{DLA} is the DLA power spectrum and P_{M} the matter power spectrum. Observationally, b_{DLA} is measured by cross-correlating DLAs with other tracers, such as the Lyman- α forest (Font-Ribera et al. 2012), or Lyman-break galaxies (Cooke et al. 2006). Because strong absorbers are rare, cross-correlating with more common objects increases the statistical power of the measurement. In our simulated results, however, we can produce arbitrarily dense quasar sightlines, so we compute b_{DLA} directly from Eq. (6) (the autocorrelation). P_{M} is the total matter power spectrum computed from the simulation box, rather than the large-scale linear power.

Both P_{DLA} and P_{M} are computed using a Fourier mesh with 4096^3 cells. Gas elements are interpolated onto the mesh using an SPH kernel, while dark matter and star particles are interpolated with a cloud-in-cell kernel. The DLA power spectrum is calculated from the interpolated gas density grid by computing

$$\delta_{\text{DLA}} = \text{I(DLA)} / \bar{n}_{\text{DLA}} - 1, \quad (7)$$

where I(DLA) is unity when a cell contains a DLA and zero otherwise. Here \bar{n}_{DLA} is the average density of DLAs per cell and we define P_{DLA} as the power spectrum of δ_{DLA} . This procedure will introduce some smoothing on the scale of the Fourier mesh, causing some DLAs to be misclassified. We checked the size of this effect by comparing with a 2048^3 grid, and found that the DLA bias was converged to $\sim 3\%$

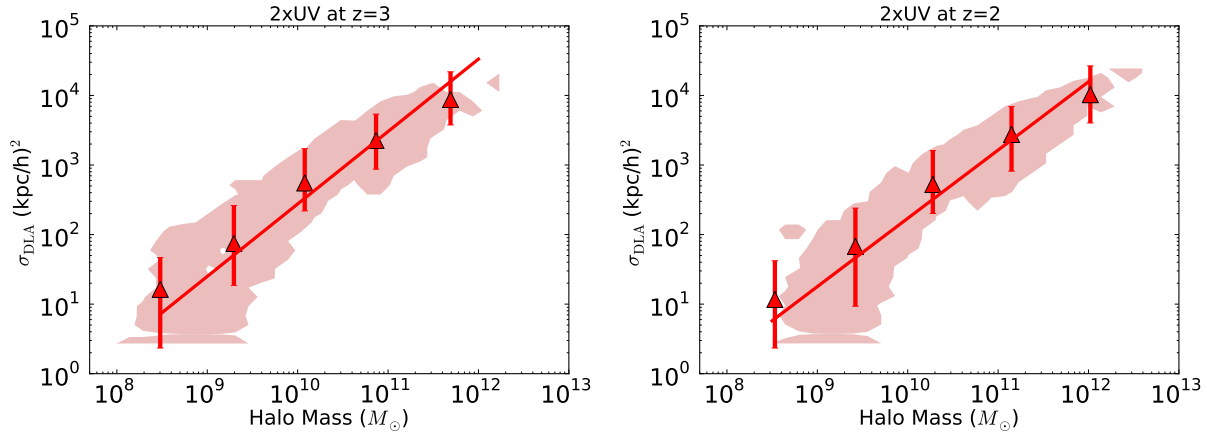


Figure 1. The cross-sectional area of DLAs around an individual halo, σ_{DLA} , as a function of halo mass for the 2xUV simulation. Shaded areas show regions of the plane occupied by at least one halo. Points indicate the median value of σ_{DLA} in comoving $(h^{-1}\text{kpc})^2$ in each mass bin, and error bars give the upper and lower quartiles. The solid line is a power law fit. (Left) At $z = 3$. (Right) At $z = 2$.

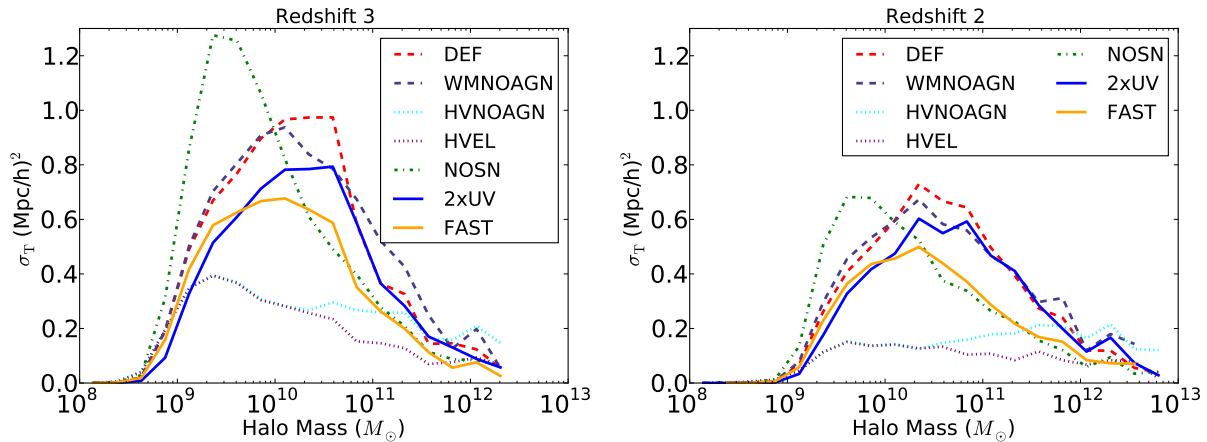


Figure 2. σ_{T} , the total cross-section of DLAs, in comoving units, hosted by all halos in each bin as a function of halo mass. Shown are 20 halo mass bins evenly spaced in $\log(M)$, for a selection of our simulations. (Left) At $z = 3$. (Right) At $z = 2$.

Name	$\sigma_{10} (h^{-1}\text{kpc}^2)$	β
DEF	2.52	1.01
WMNOAGN	2.58	1.06
HVEL	2.27	0.81
HVNOAGN	2.36	0.91
2xUV	2.44	1.04
FAST	2.40	0.93
NOSN	2.53	0.81

Table 2. Table of best-fit parameters to Eq. (8) at $z = 3$.

3 DLA PROPERTIES

3.1 Host Halos of DLAs

Figure 1 shows the cross-sectional area of DLAs in halos, σ_{DLA} , for the 2xUV simulation at $z = 3$ and $z = 2$. There is a cut-off at small halo masses where the DLA cross-section becomes negligible. The position of this cut-off is set by the amplitude of the UVB, which prevents gas in halos with a low virial velocity from collapsing enough to become self-shielded (Okamoto, Gao & Theuns 2008).

We have fit σ_{DLA} with a power law

$$\sigma_{\text{DLA}} = \sigma_{10} \left(\frac{M}{10^{10}} \right)^{\beta}. \quad (8)$$

Table 2 lists the parameters of this fit for all our simulations at $z = 3$. The fit was carried out to the median binned σ_{DLA} in each halo mass bin. We used only bins with $M > 10^{8.5} h^{-1} M_{\odot}$, which roughly corresponds to the cut-off in the DLA cross-section - halo mass relation. At higher halo masses the relation flattens (Barnes & Haehnelt 2009), and so the power law we propose should not be extrapolated to higher halo masses than in our simulation. The param-

percent. We compute the bias at $z = 2.3$ by linearly interpolating $\log b_{\text{DLA}}(k)$ as a function of scale factor, a , between $z = 2.5$ and $z = 2$.

ters of the power law relation for simulation 2xUV at $z = 4$ are 2.86 and 1.02, and at $z = 2$ are 2.19 and 1.00. At lower redshift more concentrated halos and an increased UVB amplitude produce a smaller DLA cross-section per host halo. However, the slope of the relation, which is remarkably linear, exhibits minimal evolution with redshift, although it extends towards higher masses as halos grow.

Figure 2 shows the total cross-section of DLAs in each mass bin, summed over all halos. While this is computed directly from the simulation, it can be thought of as a convolution of the halo mass function with Eq. (8). More massive halos have a greater DLA cross-section per halo, but are less common, and so have a lower total cross-section.

For the NOSN simulation, the lack of outflows means that the slope of the DLA cross-section is shallow and the host halo mass distribution peaks around $10^9 h^{-1} M_{\odot}$, below which the UVB prevents gas in halos from condensing. However, supernova feedback prevents the runaway collapse of small halos into DLAs, significantly reducing their cross-section and increasing the slope of the halo mass- σ_{DLA} relation. The HVEL and HVNOAGN simulations have the flattest DLA distribution; here the high wind velocity suppresses DLAs in host halos with $M < 10^{12} h^{-1} M_{\odot}$. For the other models, which we shall see agree best with observations, the peak halo mass is at $5 \times 10^{10} h^{-1} M_{\odot}$. The higher velocity winds in the FAST simulation suppress the DLA abundance roughly uniformly across the whole range of host halo masses. AGN feedback preferentially suppresses DLAs in high mass halos; WMNOAGN (HVNOAGN) differs from DEF (HVEL) only for halos with $M > 10^{10.2} h^{-1} M_{\odot}$. We verified that it reduces the DLA cross-section in these halos, rather than reducing the number of halos which host DLAs.

Figure 3 shows visually the distribution of DLAs and LLSs around a typical halo in two of our simulations, HVEL and 2xUV. Halos have a central density of neutral hydrogen connected with their central galaxy, surrounded by smaller concentrations associated with subhalos, merging objects or the denser parts of filaments. In 2xUV the slower outflows produce a significantly more extended central component than in HVEL.

3.2 Convergence

A $10 h^{-1} \text{Mpc}$ box was used to check convergence with resolution. This box was too small to include a statistically significant sample of halos with $M \gtrsim 10^{11} h^{-1} M_{\odot}$, which, as shown in Figure 2, contribute significantly to the total DLA density. Thus it did not completely reproduce the DLA population. However, we checked that the DLA cross-section of halos as a function of halo mass was the same in both simulations, and that DLA halos below the mass scale resolved in our $25 h^{-1} \text{Mpc}$ simulation made up less than 1% of the total DLA cross-section. We also checked that the total DLA cross-section in halos with $M = 10^9 - 10^{10} h^{-1} M_{\odot}$ did not change. Note, however, that a $25 h^{-1} \text{Mpc}$ box with 256^3 particles is not converged.

To check convergence with box size, we computed the column density function for the recently completed Illustris simulation (Vogelsberger et al. 2014c,b; Genel et al. 2014); a $75 h^{-1} \text{Mpc}$ box at similar resolution to the simulations used here. Convergence is remarkable; the CDDF was unchanged except for a small region at columns of 4×10^{20} , where it was

larger by about 20%. We attribute this mostly to the slightly reduced grid resolution we were forced by memory limitations to use for the larger box. Evolution due to improved statistics in the most massive halos would be concentrated towards higher columns.

4 COMPARISON WITH OBSERVATIONS

4.1 Column Density Distribution Function

The column density distribution function (CDDF), $f(N_{\text{HI}})$, is defined such that $f(N_{\text{HI}}) dN_{\text{HI}} dX$ is the number of absorbers per sightline with column density in the interval $[N_{\text{HI}}, N_{\text{HI}} + dN_{\text{HI}}]$. We identify sightlines with grid cells and thus count absorbers by computing a histogram of the column densities in each cell. The CDDF is defined by

$$f(N) = \frac{F(N)}{\Delta N} \Delta X(z), \quad (9)$$

where $F(N)$ is the fraction of the total number of grid cells in a given column density bin, and $\Delta X(z)$ is the absorption distance per sightline. The absorption distance is defined to account for evolution in line number density with the Hubble flow. A quantity Q with constant physical cross-section and comoving number density evolving passively with redshift will have constant dQ/dX . The (dimensionless) absorption distance is thus given by

$$X(z) = \int_0^z (1+z')^2 \frac{H_0}{H(z')} dz', \quad (10)$$

and for a box of co-moving length ΔL we have $\Delta X = (H_0/c)(1+z)^2 \Delta L$ (Bahcall & Peebles 1969).

Figure 4 shows the CDDF and its first moment, $N_{\text{HI}} f(N_{\text{HI}})$, which emphasises the differences between simulations, at $z = 3$. Several of our simulations are in reasonable agreement with the data. The DEF simulation overproduces weak DLAs and LLSs by about 1σ . This is corrected in the 2xUV simulation, which uses photo-ionisation rates close to those from Becker & Bolton (2013). This reduction of the LLS abundance in the 2xUV simulation is due to two physical effects. First, the increased UVB amplitude ionises the gas further. Second, there is increased thermal pressure support in small halos, reducing the amount of gas that is able to remain cool. We checked that both effects contribute by changing the UVB amplitude when post-processing the DEF simulation. At low column densities, the gas is beginning to come into photo-ionisation equilibrium with the UVB, and is only weakly affected by self-shielding, hence the difference between 2xUV and DEF tends to a constant. The FAST simulation is similar to 2xUV at columns $N_{\text{HI}} > 10^{18} \text{cm}^{-2}$, showing a degeneracy between the UVB amplitude and the wind velocity. However, at $N_{\text{HI}} < 10^{18} \text{cm}^{-2}$, the FAST CDDF increases until it is larger than DEF, as the faster winds move gas into less dense regions.

A minimum wind velocity decreases the abundance of DLAs, particularly weak DLAs and LLSs. Two effects, both tending to reduce the HI cross-section, contribute. First, an increased minimum wind velocity decreases the mass loading, allowing the gas in equivalent mass halos to become more concentrated, and thus decreasing the spread of neutral gas. Second, when the wind velocity is much greater than the circular velocity of the halo, $v_w \gg v_{\text{circ}}$, out-flowing

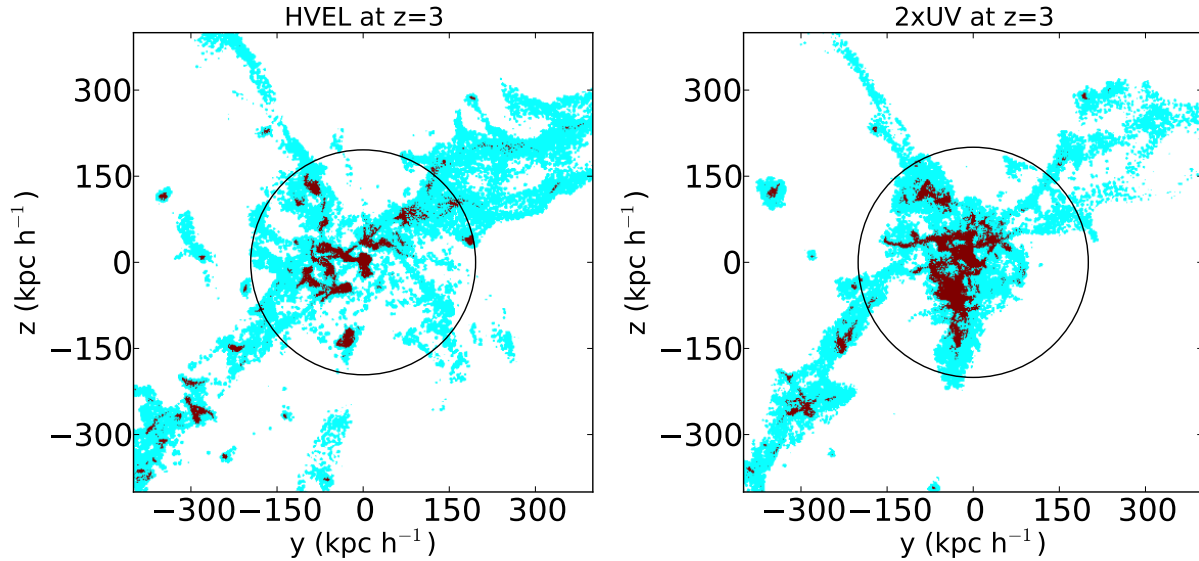


Figure 3. Distribution of DLAs and LLSs around a typical DLA hosting halo, with a central halo mass of $\sim 7 \times 10^{11} h^{-1} M_{\odot}$. DLAs are shown in dark red, LLSs in light blue. Distances are comoving. The black circle shows the virial radius of the central halo. (Left) From the HVEL simulation. (Right) From the 2xUV simulation.

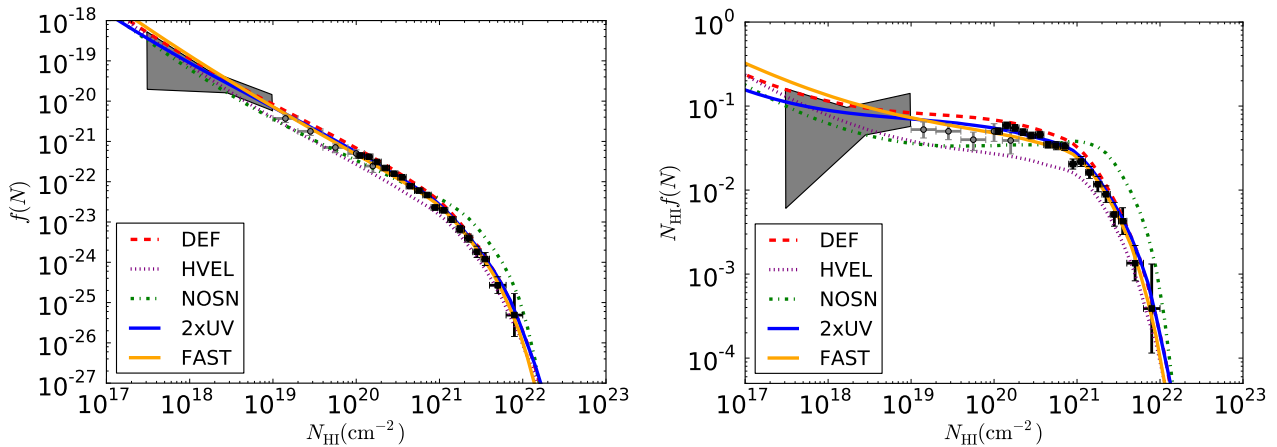


Figure 4. The column density distribution function for our simulations at $z = 3$, compared to observational data from Noterdaeme et al. (2009) at $z = 2.9$ (black squares), Zafar et al. (2013) (grey circles) and Prochaska et al. (2010) (grey box). (Left) $f(N_{\text{HI}})$ as a function of column density. (Right) The first moment of the column density, $N_{\text{HI}} f(N_{\text{HI}})$. N_{HI} is in physical cm^{-2} .

gas can escape the halo entirely and reach lower density regions where it is ionised by the UVB. The minimum wind velocity of 600 km s^{-1} in simulation HVEL is large enough to suppress the DLA abundance over the halo mass range which dominates the DLA cross-section. It thus behaves similarly to the REF OWLS model (Altay et al. 2013; Rahmati et al. 2013a), and yields a CDDF with a shape in good agreement with the data, but a normalisation a factor of two too low.

Simulation NOSN, which does not have significant feedback, produces too few LLSs, but too many absorbers with column densities $N_{\text{HI}} > 10^{21} \text{ cm}^{-2}$. Without feedback, halos undergo runaway gravitational collapse and form dense objects hosting both stars and DLAs. Figure 5 shows the consequences of including H_2 formation for the 2xUV simulation. The effect is negligible for $N_{\text{HI}} < 10^{21.5} \text{ cm}^{-2}$, but at

higher column densities decreases the CDDF by roughly 1σ . Note that even without molecular hydrogen our simulations are in reasonable agreement with the data, so any inaccuracies in our prescription for H_2 formation will not strongly affect our conclusions.

Figure 6 shows the differences between the simulations at $z = 4$ and $z = 2$ in more detail. The effect of AGN feedback is small, but most important at $z = 2$, where there are more large halos. AGN feedback is more effective in HVEL than in DEF, as a larger fraction of the DLAs in HVEL reside in halos massive enough to host AGN. Furthermore, the stronger supernova feedback in HVEL produces lower density gas on the outskirts of halos with a longer cooling time and thus a larger response to AGN heating. Note that the property of HVEL which increases the effect of AGN also causes it to disagree with observations.

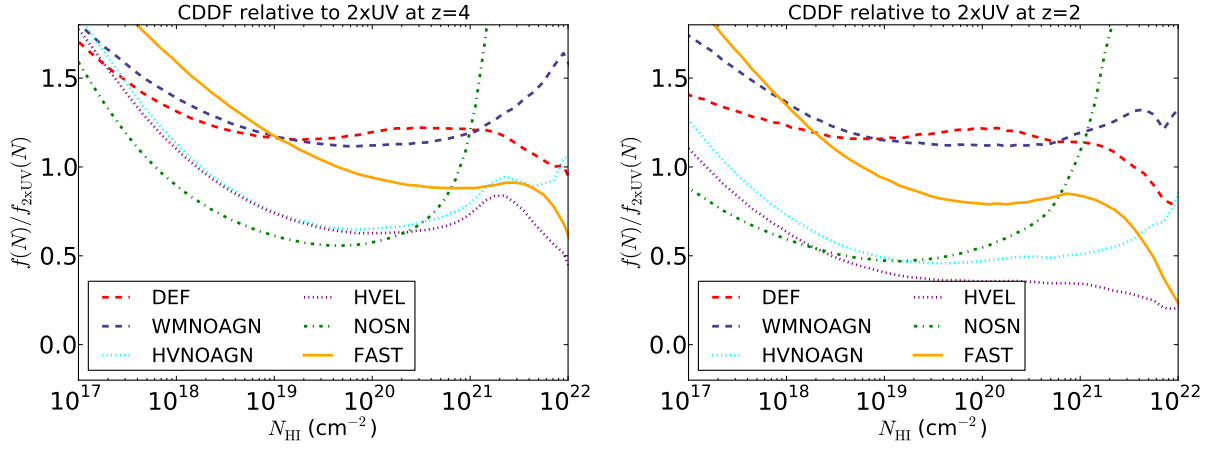


Figure 6. The ratio of the CDDF in our simulations to the CDDF of simulation 2xUV. (Left) At $z = 4$. (Right) At $z = 2$.

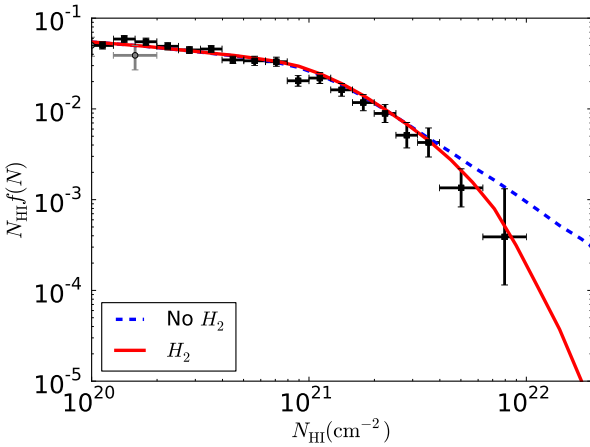


Figure 5. The first moment of the CDDF for the 2xUV simulation with and without H_2 formation at $z = 3$, compared to observational data. N_{HI} is in physical cm^{-2} .

It is interesting to compare Figure 6 to Figure 2 of Altay et al. (2013), who considered the effect of a range of feedback and star formation parameters on DLAs. The variation between their models is bounded by 0.2 dex in the LLS range, with the exception of one model where the Universe is not re-ionised (and thus, freed of the effect of photo-ionisation, much smaller halos can form DLAs) and one model with a different cosmology. This bound is saturated by their WV-CIRC and REF models, whose wind mass loading in small halos is different by a factor of ~ 3 . By contrast, the wind mass loading in small halos in our DEF simulation differs from that in the HVEL simulation by ~ 16 , producing a variation in the LLS abundance nearer 0.4 dex. This emphasises the importance of considering a wide feedback parameter space.

4.2 Redshift Evolution of Neutral Hydrogen

Figure 7 shows the redshift evolution of the incident rate of DLA systems, $\frac{dN}{dX}$, and the total HI density in DLAs, Ω_{DLA} .

$\frac{dN}{dX}$ is the integral of the CDDF

$$\frac{dN}{dX} = \int_{10^{20.3}}^{\infty} f(N, X) dN, \quad (11)$$

while Ω_{DLA} is the integral of its first moment

$$\Omega_{\text{DLA}} = \frac{m_{\text{P}} H_0}{c \rho_c} \int_{10^{20.3}}^{\infty} N f(N, X) dN. \quad (12)$$

Here ρ_c is the critical density at $z = 0$ and m_{P} is the proton mass. Note that this definition differs by a factor of $X_{\text{H}} = 0.76$, the primordial hydrogen mass fraction, from the quantity quoted by Noterdaeme et al. (2012). The difference comes from the fact that they divide the total mass of neutral hydrogen by the hydrogen mass fraction in order to obtain the total gas mass in DLAs. However, this assumes the neutral fraction of hydrogen is unity, while in fact a significant fraction of the hydrogen in DLAs may be molecular, implying that total gas mass in DLAs is greater than $\Omega_{\text{DLA}}/0.76$ by an uncertain amount. We therefore quote the total HI mass. Because the CDDF falls off sharply above the DLA threshold, $\frac{dN}{dX}$ is dominated by systems with $N_{\text{HI}} \sim 5 \times 10^{20} \text{ cm}^{-2}$, while Ω_{DLA} has a stronger contribution from higher column densities. This explains why NOSN and DEF have similar values for $\frac{dN}{dX}$; Figure 6 shows that their CDDFs cross at $N_{\text{HI}} \sim 5 \times 10^{20} \text{ cm}^{-2}$. This is a coincidence, which would not be true for a different column density threshold.

Quoted errors in Figure 7 are statistical. Systematic errors certainly dominate for $\frac{dN}{dX}$ (P. Noterdaeme, private communication), and are probably significant for Ω_{DLA} .³ At $z = 3$, Figure 7 shows the trends from Figure 4. DEF, 2xUV and FAST are in good agreement with the data. HVEL predicts too little neutral hydrogen and too little DLA cross-section, while NOSN predicts too high a value for Ω_{DLA} .

The three simulations with a large mass loading in small halos, 2xUV, DEF and FAST, all have very weak redshift evolution between $z = 4$ and $z = 3$, reproducing observational trends and reflecting a balance between gravitational collapse and supernova feedback. However, between

³ Note that identifying DLAs in low-resolution SDSS spectra is significantly more difficult than for the high resolution spectra used to measure the DLA metallicity.

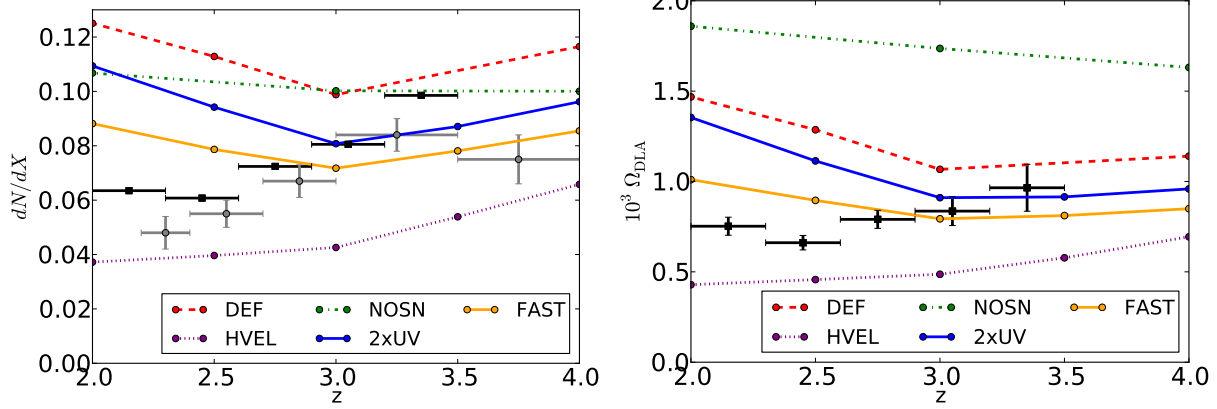


Figure 7. (Left) Redshift evolution of the total incident rate of DLAs, $\frac{dN}{dX}$. (Right) Total neutral hydrogen density in DLAs, Ω_{DLA} . Black squares show the observational data of Noterdaeme et al. (2012). For Ω_{DLA} , errors shown are statistical. For $\frac{dN}{dX}$, systematic errors in the DLA detection pipeline dominate and so no error is quoted in Noterdaeme et al. (2012). Thus we show the independent estimate from Prochaska & Wolfe (2009) (grey circles); the difference between them is a reasonable estimate of systematic uncertainty. Ω_{DLA} is dominated by stronger absorbers, and so is less affected.

$z = 2$ and $z = 3$, both $\frac{dN}{dX}$ and Ω_{DLA} begin to increase, a trend not seen in the observations. Figure 8 shows our results at $z = 2.5$, compared to the CDDF from Noterdaeme et al. (2012) at that redshift. The excess HI is concentrated in systems with column densities between 5×10^{20} and $5 \times 10^{21} \text{ cm}^{-2}$, and thus is only weakly affected by the change in the photo-ionisation rate between DEF and 2xUV. The FAST and HVEL simulations are closest to the observed Ω_{HI} at $z = 2.5$, each within 50% of the observed value. Figure 8 reveals that they do not match the observed shape of the CDDF at $z = 2.5$, as they suppress its amplitude relative to DEF, where what is required is an alteration of the shape to suppress a particular column density range. An interpolation between them could potentially match Ω_{HI} , but would produce too many columns with $N_{\text{HI}} \sim 10^{21} \text{ cm}^{-2}$ and too few with $N_{\text{HI}} \sim 10^{20} \text{ cm}^{-2}$.

Figure 9 shows the contribution to Ω_{DLA} from different halo masses, revealing that the increase at $z = 2.5$ is coming predominantly from the formation of halos with $M > 10^{10} h^{-1} M_{\odot}$. Thus a form of feedback able to suppress the amount of HI in these more massive halos is required. V13 found that their wind model over-produces the galaxy stellar mass function in halos of this size at $z = 0$, compared to the observations of Baldry, Glazebrook & Driver (2008). There is thus also some evidence for stronger feedback in the galactic stellar component. Alternatively, the gas may be more ionised than we have assumed, perhaps by a thermal component to the feedback. We have checked that one cannot achieve agreement with observations by choosing a minimum wind velocity intermediate between HVEL and DEF. This suppresses DLAs in small halos, but has a reduced effect on the more massive halos which are responsible for the increase in Ω_{DLA} at $z < 3$. We also considered adding a feedback component in which the wind mass loading scales as the inverse of the wind velocity (momentum driven winds, following Davé et al. (2013)), but found that a momentum driven component large enough to affect Ω_{DLA} completely suppressed columns with $N_{\text{HI}} > 10^{21} \text{ cm}^{-2}$, while leaving lower columns unchanged, and so was strongly discrepant with the observed column density function.

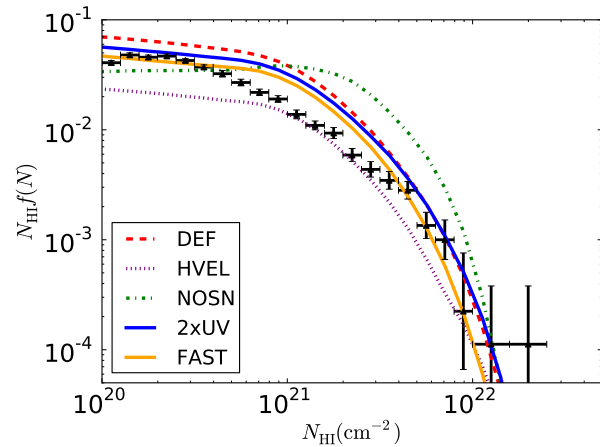


Figure 8. The first moment of the CDDF at $z = 2.5$, compared to observational data from Noterdaeme et al. (2012) (triangles), for our simulations. N_{HI} is in physical cm^{-2} .

4.3 DLA Bias

We define the DLA bias in Eq. (6) as the ratio of the DLA power spectrum to the non-linear matter power spectrum. Figure 10 shows the results of our simulations and the measurements of Font-Ribera et al. (2012) (henceforth F12). The 2xUV simulation is in good agreement with the observations on the scales where the two overlap. The DLA bias in DEF and FAST is a little lower than is observed, as both simulations are less efficient than 2xUV at suppressing the formation of DLAs in low-mass halos. The HVEL simulation yields a strongly scale-dependent DLA bias, which comes close to the observations on scales of the simulation box, but is significantly larger on smaller scales. Recently, Kuhlen et al. (2012); Kuhlen, Madau & Krumholz (2013) proposed a model without stellar feedback, where star formation is instead suppressed by raising the star formation threshold density. Such a model would produce DLAs with a distribution similar to our NOSN simulation, and thus have a DLA bias significantly smaller than is observed.

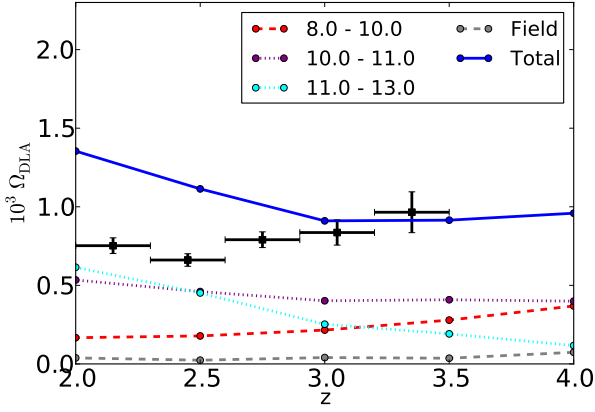


Figure 9. Contributions to Ω_{DLA} from different halo mass bins as a function of redshift for the 2xUV simulation. The increase at $z = 2.5$ is due to the formation of more massive halos. The field category denotes DLAs we could not associate with a halo. Black squares show the observational data of Noterdaeme et al. (2012).

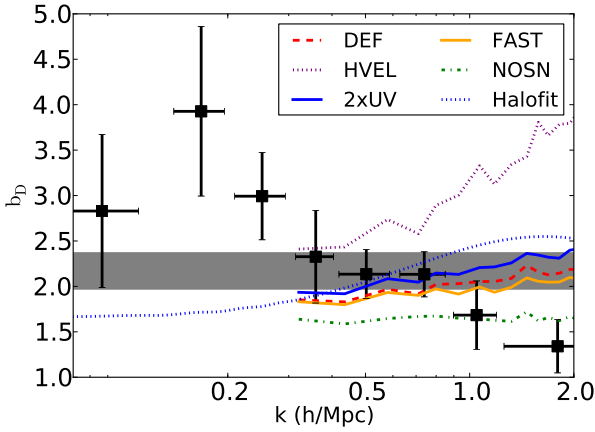


Figure 10. The bias of the DLA power spectrum to the matter power spectrum at $z = 2.3$, as measured from our simulations. Points with error bars show the bias measured by Font-Ribera et al. (2012) as a function of scale, while the grey band shows a 1σ confidence interval for the average bias on scales $k < 1.2 h\text{Mpc}^{-1}$. The dotted line labelled Halofit shows the results of applying Eq. (14) to the 2xUV simulation.

F12 used their measurements to deduce the mass of DLA host halos. To do this, they assumed scale independence of the DLA bias in the large-scale limit, a halo mass function and a model for the abundance of DLAs in a halo. The DLA bias in this linear model is then

$$b_{\text{DLA}} = \frac{\int b(M)\sigma_{\text{T}}(M)dM}{\int \sigma_{\text{T}}(M)dM}, \quad (13)$$

where $b(M)$ is the linear halo bias (Tinker et al. 2010), and σ_{T} is the total DLA cross-section in a halo mass bin. Note that $\sigma_{\text{T}}(M) = n(M)\sigma_{\text{DLA}}(M)$, where $n(M)$ is the halo mass function (Tinker et al. 2008), and σ_{DLA} the DLA cross-section for a given halo mass. F12 described the DLA cross-section using the results of Pontzen et al. (2008), which has $\sigma_{\text{DLA}} \sim M^{0.5}$ for $M > 10^{10} h^{-1} M_{\odot}$. As discussed in Section 3.1, our 2xUV simulation has $\sigma_{\text{DLA}} \propto M$, which yields

a significantly increased DLA cross-section in more massive halos, and thus a larger DLA bias.

Applying Eq. (13) to 2xUV gives $b_{\text{DLA}} = 1.72$ at $z = 2.3$, about 15% less than the mean bias shown in Figure 10. This discrepancy arises due to non-linear growth in the DLA power spectrum. For our cosmology at $z = 2.3$, the non-linear scale (defined as the scale where the variance of the fluctuations is unity) is $k \approx 2.4 h\text{Mpc}^{-1}$. However, for objects which trace the matter power spectrum with a bias of two, the non-linear scale is four times larger, $k \sim 0.6 h^{-1}\text{Mpc}$, and non-linear growth can affect the power spectrum at the 10% level for $k \sim 0.2 h^{-1}\text{Mpc}$.

As the rate of structure growth is higher in the non-linear regime, non-linear growth leads to an increase in bias on small scales. Since more massive halos are more biased and thus more non-linear, non-linear growth increases their fractional contribution to the DLA bias. We shall account for these effects by making the (crude) assumption that the non-linear power spectrum of a biased object, P_{NL} , can be described by applying the HALOFIT model, \mathcal{H} , (Smith et al. 2003) to the biased linear power spectrum. That is, we assume $P_{\text{NL}} = \mathcal{H}[b^2 P(k)]$. The modified form of Eq. (13) is then

$$b_{\text{DLA}}^2(k) = \frac{\int \mathcal{H}[b^2(M)P(k)] \sigma_{\text{T}}^2 dM}{\int \mathcal{H}[P(k)] \sigma_{\text{T}}^2 dM}. \quad (14)$$

Figure 10 shows Eq. (14) applied to the 2xUV simulation. While this model overestimates non-linear effects on small scales, it can reconcile Eq. (13) with the direct calculation. It does not, however, agree with the scale-dependence of the observed bias, which decreases for $k > 0.8 h\text{Mpc}^{-1}$. This may be due to the effects of noise, or to differences between the DLA power spectrum and the cross-correlation with the Lyman- α forest. Computing the cross-correlation directly is beyond the scope of this paper, but we plan to examine it further in future work. For now, we note only that the physical mechanism causing this effect may conceivably impact larger scales and thus change our agreement with the observed data.

Since the good agreement of the 2xUV simulation with observations results in part from non-linear growth, we expect the bias to decrease on larger scales, and asymptote to the linear theory value by $k = 0.1 h\text{Mpc}^{-1}$. This will reduce the mean bias derived from the simulations; a simple estimate of the effect using Eq. (14) suggests that the simulation would be around 1.5σ below the observations. However, since most of the statistical power for the F12 measurement is derived from scales probed by our simulation box, the discrepancy in a full likelihood analysis would probably be somewhat less.

4.4 Metallicity

Figure 11 shows the probability density function of the log of mass-weighted DLA metallicity, normalised so that the total area under the curve is unity. We define the metallicity of each DLA grid cell to be

$$Z = \frac{\sum M_Z}{\sum M_H}, \quad (15)$$

where M_Z is the mass in metals of a pixel and M_H the total gas mass. All metallicities are quoted in units of the Solar

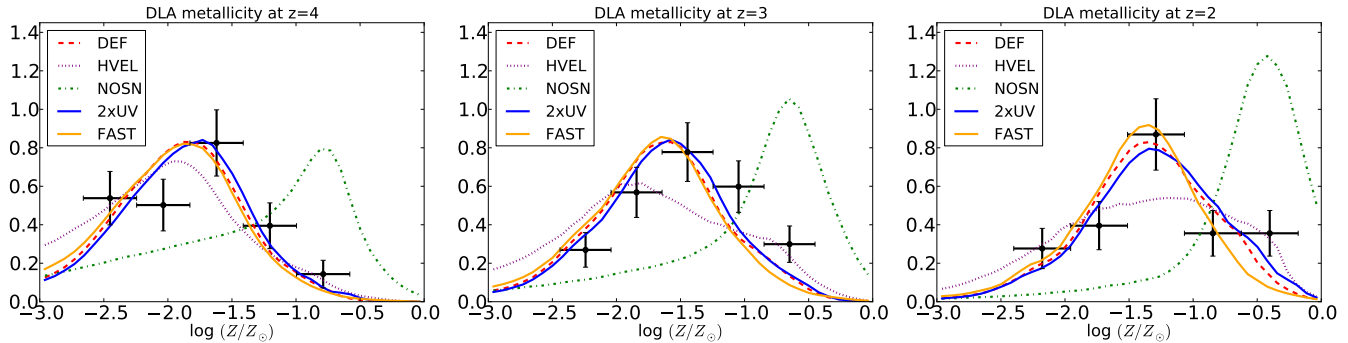


Figure 11. The probability density function of the log of DLA metallicity normalised so that the total area under the curve is unity. (Left) $z = 4$ (Centre) $z = 3$ (Right) $z = 2$. Data is from the compendium found in Rafelski et al. (2012). $Z_{\odot} = 0.0182 = 0.0134/0.7381$.

metallicity, which we take to be $M_{Z_{\odot}}/M_{H_{\odot}} = 0.0134/0.7381$ (Asplund et al. 2009).⁴ The total mass in metals and the total gas mass in each cell are computed separately using the same method used to compute HI columns and explained in Section 2.3.

DEF, 2xUV and FAST give very similar results in very good agreement with the data at all three redshifts. NOSN is a very poor fit to the data, giving a mean DLA metallicity much higher than observed. While HVEL is in good agreement at $z = 4$, at $z = 3$ it does not reproduce the peak in the metallicity, instead yielding a plateau. Figure 1 shows that this plateau appears also in the distribution of DLA host halos, and we have checked that the differences in metallicity between HVEL and DEF are dominated by changes in the DLA host halos, rather than variations in the DLA mass-metallicity relation.

The mean metallicity evolves with redshift, increasing at later times. This mirrors evolution in the host halos of DLAs shown in Figure 2. At later times DLAs trace more massive halos, which are more metal-rich. The DLA metallicity is well described by a power law in halo mass, whose parameters show little evolution with redshift, confirming that the evolution in the average DLA metallicity is driven by evolution in the host halo distribution.

We have checked that the metal loading factor of the winds does not affect the DLA metallicity, even at redshift four, where re-accretion will not be significant. Thus DLAs are primarily enriched directly by the stellar mass in the host halo, rather than through outflows. Our good agreement with the observed DLA metallicity implies that our DLA halo hosts have the correct stellar mass. Note however that these DLAs are hosted in small halos, so their stellar components are hard to observe directly.

5 CONCLUSIONS

We have studied Damped Lyman- α systems (DLAs) using a suite of hydrodynamic simulations which vary the parameters of the stellar feedback model around those used in Vogelsberger et al. (2013) (V13). The predictions of this model for DLAs have not previously been examined, and

so our work serves as an important verification of its ability to match observed data. By using a suite of simulations, we are able to show the extent to which each observation constrains our model, something which has only previously been done for the column density function.

Turning to the DLA metallicity distribution, most of our simulations gave similar results, in good agreement with the observations at $z = 2 - 4$. This suggests that the DLA metallicity is not particularly sensitive to the detailed parameters of the feedback model. The exceptions were a simulation without feedback and one including outflows with a constant velocity in small halos. Similar models tend to over-produce the faint end of the galaxy stellar mass function at low redshift (Puchwein & Springel 2013). Since the metallicity distribution traces the stellar mass of DLA hosts, it provides a valuable indirect probe of the faint end of the stellar mass function at redshifts where it cannot yet be observed directly. That the two simulations mentioned above do not match the DLA metallicity therefore demonstrates the need, even at $z = 2 - 4$, for a form of feedback which suppresses star formation very efficiently in small halos.

The observed abundance of neutral hydrogen in DLAs is more constraining. The parameters chosen in V13 over-produce the number of weak DLAs and LLS relative to observations at $z = 3$. We demonstrate that by increasing the photo-ionisation rate from the UV background to match constraints from improved measurements of the IGM temperature (Becker & Bolton 2013), we are able to alleviate the problem, both directly by ionising the gas and indirectly by preventing the accretion of small halos. With this modification, we match the observed column density function at $z = 3$ and its integrals, the total density of neutral hydrogen and the incident rate of DLAs, at $z = 3 - 4$.

However, for $z < 3$ the total HI density in our simulations increases, and by $z = 2.5$ all our simulations are in strong tension with observations. By comparing to the column density function (CDDF) at $z = 2.5$, we demonstrate that this tension arises from growth in the amount of HI in halos of mass $10^{11} - 10^{13} h^{-1} M_{\odot}$. We considered a minimum wind velocity of 600 km s^{-1} , which makes the wind velocity constant in most DLA host halos. However, we found that, although we now reproduce the observed evolution in Ω_{DLA} at $z > 2$, the stronger winds over-suppress the DLA cross-section, producing a normalisation of Ω_{DLA} which is too low. We considered other variations, including a simulation with

⁴ Note that in Vogelsberger et al. (2014c) we incorrectly used a Solar metallicity of 0.0134.

a 50% higher wind velocity. This simulation does produce somewhat better agreement with Ω_{DLA} at $z = 2.5$, but it does not match the detailed shape of the CDDF, producing too few weak DLAs and too many stronger ones. Resolving this tension could involve winds which behave like our model in smaller halos, but are effective at reducing the density of gas in more massive halos. Alternatively, the feedback could become yet stronger in small halos at low redshift.

We compare recent measurements of the DLA bias from the cross-correlation of DLAs with the Lyman- α forest to our simulations. Due to our strong feedback, the DLA cross-section in our preferred models is dominated by systems hosted in halos of mass $10^{10} - 10^{11} h^{-1} M_{\odot}$, and we find that our preferred model produces simulations and observations in good agreement on scales where they overlap. However, a simple linear bias model applied to our simulation suggests a bias around $2 - \sigma$ lower than observed. We show that this discrepancy is due to non-linear effects, which on the scales concerned can increase the DLA bias by 15%. Structure formation is faster in the non-linear regime; hence the larger and more massive halos make a larger contribution to the total power spectrum, increasing the DLA bias. Non-linear growth should therefore be taken into account when interpreting observational results.

The bias measurements, unlike our simulated results, appear to decrease substantially for $k > 1 h\text{Mpc}^{-1}$. The reason for this is currently unclear, but may be related to the fact that we have derived the DLA bias directly from the DLA power spectrum, while the observed bias is derived from the cross-correlation of DLAs with the Lyman- α forest.

We have deferred discussion of the velocity structure of DLAs, which provide an important extra constraint on the model, to a companion paper (Bird et al. 2014). There we will show that we are able to use the same models which are preferred in this paper to match the velocity width distribution significantly better than previous work, corroborating the need for strong feedback in low mass haloes at $z = 2 - 4$.

To summarise, our detailed comparison of observed DLA properties with hydrodynamic simulations has shown that the observed incidence rate and bias provide particularly powerful constraints on the parameters of the stellar feedback model in the simulations. The default feedback model presented in V13 (with the UVB amplitude doubled to match the most recent UVB measurements) produces DLAs in good agreement with a wide range of observed DLA properties. Some discrepancies remain, particularly at $z < 3$, but overall our results confirm a picture where DLAs form in a wide range of halos with a relatively modest mean mass, $10^{10} - 10^{11} h^{-1} M_{\odot}$. We have shown explicitly that this mass range is not in disagreement with the recently reported value of the DLA bias parameter, despite the indications of simple linear calculations, due to non-linear growth of the biased DLA host haloes in the simulations.

ACKNOWLEDGEMENTS

SB thanks Pasquier Noterdaeme for useful discussions regarding the data, J. Xavier Prochaska for useful comments and proof-reading an earlier version of the paper and Ryan Cooke for noticing a units error in an earlier draft. SB is supported by the National Science Foundation grant num-

ber AST-0907969, the W.M. Keck Foundation and the Institute for Advanced Study. MGH acknowledges support from the FP7 ERC Advanced Grant Emergence-320596. VS acknowledges support from the European Research Council under ERC-StG grant EXAGAL-308037. LH is supported by NASA ATP Award NNX12AC67G and NSF grant AST-1312095.

REFERENCES

- Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, *ApJL*, 737, L37, arXiv:1012.4014,
 Altay G., Theuns T., Schaye J., Booth C. M., Dalla Vecchia C., 2013, *MNRAS*, 436, 2689, arXiv:1307.6879,
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481, arXiv:0909.0948,
 Bahcall J. N., Peebles P. J. E., 1969, *ApJL*, 156, L7,
 Baldry I. K., Glazebrook K., Driver S. P., 2008, *MNRAS*, 388, 945, arXiv:0804.2892,
 Barnes L. A., Haehnelt M. G., 2009, *MNRAS*, 397, 511, arXiv:0809.5056,
 Becker G. D., Bolton J. S., 2013, *MNRAS*, 436, 1023, arXiv:1307.2259,
 Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, *MNRAS*, 410, 1096, arXiv:1008.2622,
 Bird S., Vogelsberger M., Sijacki D., Zaldarriaga M., Springel V., Hernquist L., 2013, *MNRAS*, 429, 3341, arXiv:1209.2118,
 Bird S., Haehnelt M., Neeleman M., Genel S., Vogelsberger M., Hernquist L., 2014, *ArXiv e-prints*, arXiv:1407.7858
 Blitz L., Rosolowsky E., 2006, *ApJ*, 650, 933, arXiv:astro-ph/0605035,
 Bolton J. S., Becker G. D., Haehnelt M. G., Viel M., 2014, *MNRAS*, arXiv:1308.4411,
 Cen R., 2012, *ApJ*, 748, 121, arXiv:1010.5014,
 Chabrier G., 2003, *PASP*, 115, 763, arXiv:astro-ph/0304382,
 Cooke J., Wolfe A. M., Gawiser E., Prochaska J. X., 2006, *ApJ*, 652, 994, arXiv:astro-ph/0607149,
 Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, *MNRAS*, 434, 2645, arXiv:1302.3631,
 Erkal D., Gnedin N. Y., Kravtsov A. V., 2012, *ApJ*, 761, 54, arXiv:1201.3653,
 Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008a, *ApJ*, 681, 831, arXiv:0709.2382,
 Faucher-Giguère C.-A., Lidz A., Hernquist L., Zaldarriaga M., 2008b, *ApJ*, 688, 85, arXiv:0807.4177,
 Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416, arXiv:0901.4554,
 Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, *ApJ*, 725, 633, arXiv:1005.3041,
 Font-Ribera A. et al., 2012, *JCAP*, 11, 59, arXiv:1209.4596,
 Fumagalli M., Krumholz M. R., Hunt L. K., 2010, *ApJ*, 722, 919, arXiv:1008.4137,
 Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, *MNRAS*, 418, 1796, arXiv:1103.2130,
 Gardner J. P., Katz N., Weinberg D. H., Hernquist L., 1997, *ApJ*, 486, 42, arXiv:astro-ph/9705118,

- Genel S. et al., 2014, ArXiv e-prints, arXiv:1405.3749
- Haehnelt M. G., Steinmetz M., Rauch M., 1998, ApJ, 495, 647, arXiv:astro-ph/9706201,
- Jedamzik K., Prochaska J. X., 1998, MNRAS, 296, 430, arXiv:astro-ph/9706290,
- Katz N., Weinberg D. H., Hernquist L., 1996a, ApJS, 105, 19, arXiv:astro-ph/9509107,
- Katz N., Weinberg D. H., Hernquist L., Miralda-Escude J., 1996b, ApJL, 457, L57, arXiv:astro-ph/9509106,
- Kim T.-S., Partl A. M., Carswell R. F., Müller V., 2013, A&A, 552, A77, arXiv:1302.6622,
- Kuhlen M., Krumholz M. R., Madau P., Smith B. D., Wise J., 2012, ApJ, 749, 36, arXiv:1105.2376,
- Kuhlen M., Madau P., Krumholz M. R., 2013, ApJ, 776, 34, arXiv:1305.5538,
- Maller A. H., Prochaska J. X., Somerville R. S., Primack J. R., 2001, MNRAS, 326, 1475,
- Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750, arXiv:1305.5360,
- Nagamine K., Springel V., Hernquist L., 2004a, MNRAS, 348, 421, arXiv:astro-ph/0302187,
- Nagamine K., Springel V., Hernquist L., 2004b, MNRAS, 348, 435, arXiv:astro-ph/0305409,
- Nagamine K., Wolfe A. M., Hernquist L., Springel V., 2007, ApJ, 660, 945, arXiv:astro-ph/0510729,
- Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353, arXiv:1301.6753,
- Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087, arXiv:0908.1574,
- Noterdaeme P. et al., 2012, A&A, 547, L1, arXiv:1210.1213,
- O’Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666, arXiv:astro-ph/0610726,
- Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920, arXiv:0806.0378,
- Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208, arXiv:0909.0265,
- Okoshi K., Nagashima M., 2005, ApJ, 623, 99, arXiv:astro-ph/0412561,
- Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577, arXiv:0712.1827,
- Pontzen A. et al., 2008, MNRAS, 390, 1349, arXiv:0804.4474,
- Pontzen A. et al., 2010, MNRAS, 402, 1523, arXiv:0909.1321,
- Prochaska J. X., Wolfe A. M., 1997, ApJ, 487, 73, arXiv:astro-ph/9704169,
- Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543, arXiv:0811.2003,
- Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123, arXiv:astro-ph/0508361,
- Prochaska J. X., O’Meara J. M., Worseck G., 2010, ApJ, 718, 392, arXiv:0912.0292,
- Puchwein E., Springel V., 2013, MNRAS, 428, 2966, arXiv:1205.2694,
- Puchwein E., Pfrommer C., Springel V., Broderick A. E., Chang P., 2012, MNRAS, 423, 149, arXiv:1107.3837,
- Péroux C., Storr-Lombardi L. J., McMahon R. G., Irwin M., Hook I. M., 2001, AJ, 121, 1799, arXiv:astro-ph/0101179,
- Péroux C., Dessauges-Zavadsky M., D’Odorico S., Sun Kim T., McMahon R. G., 2005, MNRAS, 363, 479, arXiv:astro-ph/0507353,
- Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89, arXiv:1205.5047,
- Rahmati A., Schaye J., 2014, MNRAS, 438, 529, arXiv:1310.3317,
- Rahmati A., Pawlik A. H., Raicevic M., Schaye J., 2013a, MNRAS, 430, 2427, arXiv:1210.7808,
- Rahmati A., Schaye J., Pawlik A. H., Raicevic M., 2013b, MNRAS, 431, 2261, arXiv:1301.1978,
- Razoumov A. O., 2009, ApJ, 707, 738, arXiv:0808.2645,
- Rudie G. C., Steidel C. C., Pettini M., 2012, ApJL, 757, L30, arXiv:1209.0005,
- Rudie G. C., Steidel C. C., Shapley A. E., Pettini M., 2013, ApJ, 769, 146, arXiv:1304.6719,
- Schaye J., 2001, ApJL, 562, L95, arXiv:astro-ph/0109280,
- Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877, arXiv:0705.2238,
- Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999, arXiv:1109.3468,
- Smith R. E. et al., 2003, MNRAS, 341, 1311, arXiv:astro-ph/0207664,
- Springel V., 2005, MNRAS, 364, 1105, arXiv:astro-ph/0505010,
- Springel V., 2010, MNRAS, 401, 791, arXiv:0901.4107,
- Springel V., Hernquist L., 2003, MNRAS, 339, 289, arXiv:astro-ph/0206393,
- Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776, arXiv:astro-ph/0411108,
- Storr-Lombardi L. J., Wolfe A. M., 2000, ApJ, 543, 552, arXiv:astro-ph/0006044,
- Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411, arXiv:0904.3545,
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709, arXiv:0803.2706,
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878, arXiv:1001.3162,
- Torrey P., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 427, 2224, arXiv:1110.5635,
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, arXiv:1305.4931,
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031, arXiv:1305.2913,
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2014a, MNRAS, 438, 3607,
- Vogelsberger M. et al., 2014b, MNRAS, 444, 1518, arXiv:1405.2921,
- Vogelsberger M. et al., 2014c, Nature, 509, 177, arXiv:1405.1418,
- Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D., 1986, ApJS, 61, 249,
- Wolfe A. M., Lanzetta K. M., Foltz C. B., Chaffee F. H., 1995, ApJ, 454, 698,
- Yajima H., Choi J.-H., Nagamine K., 2012, MNRAS, 427, 2889, arXiv:1112.5691,
- Zafar T., Péroux C., Popping A., Milliard B., Deharveng J.-M., Frank S., 2013, A&A, 556, A141, arXiv:1307.0602,