Moving-mesh cosmology: properties of neutral hydrogen in absorption

Simeon Bird,1,* Mark Vogelsberger,2 Debora Sijacki,2 Matias Zaldarriaga,1 Volker Springel3,4 and Lars Hernquist2

1Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA
2Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
3Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, D-69118 Heidelberg, Germany
4Zentrum für Astronomie der Universität Heidelberg, ARI, Mönchhofstr. 12–14, D-69120 Heidelberg, Germany

ABSTRACT
We examine the distribution of neutral hydrogen in cosmological simulations carried out with the new moving-mesh code AREPO and compare it with the corresponding GADGET simulations based on the smoothed particle hydrodynamics (SPH) technique. The two codes use identical gravity solvers and baryonic physics implementations, but very different methods for solving the Euler equations, allowing us to assess how numerical effects associated with the hydro solver impact the results of simulations. Here we focus on an analysis of the neutral gas, as detected in quasar absorption lines. We find that the high column density regime probed by damped Lyα (DLA) and Lyman limit systems (LLS) exhibits significant differences between the codes. GADGET produces spurious artefacts in large haloes in the form of gaseous clumps, boosting the LLS cross-section. Furthermore, it forms haloes with denser central baryonic cores than AREPO, which leads to a substantially greater DLA cross-section from smaller haloes. AREPO thus produces a significantly lower cumulative abundance of DLAs, which is intriguingly in much closer agreement with observations. The column density function, however, is not altered enough to significantly reduce the discrepancy with the observed value. For the low column density gas probed by the Lyα forest, the codes differ only at the level of a few per cent, suggesting that this regime is quite well described by both methods, a fact that is reassuring for the many Lyα studies carried out with SPH thus far. While the residual differences are smaller than the errors on current Lyα forest data, we note that this will likely change for future precision experiments.

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

1 INTRODUCTION
Absorption features in quasar spectra offer a unique view of structure formation at redshifts z = 2–4. Lyα absorption of neutral hydrogen directly tracks the distribution of gas during the initial stages of galaxy formation. Damped Lyα systems (DLAs) have a neutral hydrogen column density \( N_{\text{HI}} > 10^{20.3} \text{cm}^{-2} \) (Wolfe et al. 1986), and thus can be observed through natural broadening of the Lyα line. DLAs are thought to be high-redshift protogalaxies, sufficiently dense that their interiors are shielded from the ionizing effect of the diffuse radiation background (Katz et al. 1996b; Nagamine et al. 2010). Thus, at \( z = 2–4 \) they are understood to be reservoirs containing most of the neutral hydrogen in the Universe (Gardner et al. 1997), corresponding either to large discs (Prochaska & Wolfe 1997; Maller et al. 2001) or irregular protogalactic clumps (Haehnelt, Steinmetz & Rauch 1998; Okoshi & Nagashima 2005). Recent kinematic data may prefer large discs (Barnes & Haehnelt 2009; Hong et al. 2010). A wide range of quasar surveys have gradually 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. 2009), low-redshift DLAs (Rao & Turnshek 2000; Prochaska et al. 2001; Chen & Lanzetta 2003) and their lower column density cousins, Lyman limit systems (LLS; Péroux et al. 2001; O’Meara et al. 2007; Prochaska & Wolfe 2009). LLS are defined to be absorbers with \( 10^{17} < N_{\text{HI}} < 10^{20.3} \text{cm}^{-2} \). They have been connected to lower density analogues of DLAs (Gardner et al. 2001) or filamentary structures on the outskirts of protogalaxies (Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011), and are affected by shielding (McQuinn, Oh & Faucher-Giguère 2011). At lower column densities, we see the Lyα forest; a complex region of overlapping Lyα lines. The Lyα forest is a probe of the matter distribution in the low-density intergalactic medium (Hernquist et al. 1996; Miralda-Escude et al. 1996; Croft et al. 1998; Gnedin & Hui 1998; Davé et al. 1999). It has been used to constrain the
initial conditions of the Universe (e.g. Seljak et al. 2005; Viel & Haehnelt 2006; Viel, Bolton & Haehnelt 2009; Bird et al. 2011), the processes that govern the thermal state of the gas (e.g. Bolton et al. 2008; Faucher-Giguère et al. 2008b; Lidz et al. 2010) and indirectly helium reionization (Faucher-Giguère et al. 2008a; Becker et al. 2011).

The Lyα forest is produced by diffuse absorbers $\sim 100 h^{-1}$ kpc across, collapsing under gravity. DLAs, on the other hand, are produced in smaller, denser systems and are strongly influenced by gas physics. In both cases, obtaining accurate quantitative results for their properties requires following non-linear gravitational collapse with $N$-body simulation techniques and a method for solving the inviscid Euler equations for the cosmic gas. Historically two main approaches have been used when tackling this problem through direct simulations: grid-based codes that discretize the gas on an adaptively refined mesh (Berger & Colella 1989; Teyssier 2002; O’Shea et al. 2004) and particle-based codes that use the smoothed particle hydrodynamics (SPH) technique (Gingold & Monaghan 1977; Lucy 1977; Monaghan 1992, 2005). The former solves the equations of motion for a fluid on a stationary grid of cells in an Eulerian fashion. The latter is a pseudo-Lagrangian technique, where the fluid is split into a number of discrete mass elements, which are assumed to be indivisible and are then followed as particles. Smoothed fluid quantities are constructed from the particles through kernel interpolation.

Each approach has its own advantages and disadvantages, making it non-trivial to judge their relative accuracy for different applications. For example, while SPH codes have excellent conservation properties, they suffer from large gradient errors and accuracy problems in resolving fluid instabilities (Agertz et al. 2007), something that can make them fail dramatically in some idealized fluid dynamics problems (Sijacki et al. 2012). Eulerian mesh codes on the other hand offer high accuracy for capturing shocks, but many have incorporated relatively inaccurate gravity solvers (see e.g. O’Shea et al. 2005). Furthermore, their truncation error depends on the absolute velocity of the gas, which is problematic for the often highly supersonic flows encountered in cosmological simulations.

Springel (2010) proposed a new technique which aims to combine the strengths of both approaches. In this method, space is tessellated with a moving mesh; cells advect with the fluid flow, and as a result each cell contains approximately the same gas mass. Thus, the resolution automatically increases in areas of higher density, just as in Lagrangian methods like SPH. A moving-mesh method can resolve shocks equally as well as grid methods, since in both cases the Euler equations are solved with a high-accuracy finite-volume method, but without the advection errors present in Eulerian codes. The \texttt{AREPO} code is an implementation of this technique and has been shown to perform well in many idealized simulations of fluid dynamical problems (Springel 2010; Sijacki et al. 2012).

We compare results obtained with \texttt{AREPO} to corresponding simulations with the well-tested SPH code \texttt{GADGET}, last described in Springel (2005). Although we know that SPH cannot accurately describe fluid instabilities and mixing processes in its widely employed `standard' form, it is not obvious a priori to what extent these inaccuracies are important for different aspects of structure formation, or how they affect our interpretation of observational data. This paper therefore examines the important issue of how predictions for the neutral hydrogen distribution in the Universe depend on the employed numerical technique.

Our study is part of a series which compares simulations run using \texttt{AREPO} and \texttt{GADGET} with identical initial conditions, gravity solver and baryonic physics parameters. They thus differ only in the approach to hydrodynamics, offering an unrivalled opportunity to isolate the effects of numerical uncertainties. In previous work, Vogelsberger et al. (2012) examined the global properties of baryons and haloes, as well as performing convergence and numerical tests on \texttt{AREPO}. Sijacki et al. (2012) studied a number of idealized fluid problems to clarify the origins of the observed differences. Keres et al. (2012) looked at the effect on the resulting galaxy properties, and Torrey et al. (2011) focused on the structure of galactic discs. Here we shall look at how the use of a moving-mesh technique affects the properties of DLAs, LLS and the Lyα forest. Neutral hydrogen is particularly relevant as it is a comparatively clean probe of the hydrodynamics. We include star formation, but neglect strong feedback from outflows. Scannapieco et al. (2012) examined the effect of different feedback models and hydrodynamic solvers on a single collapsed object at $z = 0$, and found that it was more strongly affected by the choice of feedback model. Omitting strong feedback thus helps to avoid the possibility that necessary differences in the feedback implementation may affect our results for larger halo samples. Moreover, rather than a full radiative transfer model, we shall use a simple density cut-off to account for self-shielding of the gas. While less sophisticated than many previous works, this simple approach allows us to focus on the effect of the hydro solver and helps to connect our intuition from idealized tests with observations.

We build on a large literature of simulated DLAs: Nagamine, Springel & Hernquist (2004a,b) used SPH simulations with \texttt{GADGET} to examine their abundance and metallicity. They also examined the effects of galactic winds on DLAs, looked at further by Nagamine et al. (2007) and Tescari et al. (2009). Pontzen et al. (2008) attempted to reproduce many observed properties of DLAs with a complete simulation framework incorporating a simple model of radiative transfer and supernova feedback. Radiative transfer was also examined in more detail by Vajima, Choi & Nagamine (2012), Erkal, Gnedin & Kravtsov (2012) and Altay et al. (2011) looked at the effects of molecular hydrogen. Cen (2012) and Fumagalli et al. (2011) both used Eulerian grid codes with adaptive mesh refinement (AMR) to study DLAs and LLS. They included feedback from star formation and metal cooling, which we neglect, but their AMR-based simulations were unable to resolve haloes less massive than $10^{10} h^{-1} M_{\odot}$. This paper is structured as follows. In Section 2, we discuss our methods in more detail. We then present our results in Section 3 and conclude with a summary of our findings in Section 4.

2 METHODS

2.1 Numerical codes and simulation set

Both \texttt{AREPO} and \texttt{GADGET} compute gravitational interactions using the TreePM approach, as described in Springel (2005). Radiative cooling is implemented, following Katz, Weinberg & Hernquist (1996a), using a rate network, including line cooling, free–free emission and inverse Compton cooling off the cosmic microwave background. We assume ionization equilibrium and optically thin gas during the simulation, accounting for gas self-shielding by post-processing the simulation outputs, as described in Section 2.2. The history of the ultraviolet background (UVB) follows the estimates of Faucher-Giguère et al. (2009).

Star formation is implemented with the effective two-phase model of Springel & Hernquist (2003). We use the same parameters as in that paper, giving a threshold density for star formation of $N_h = 0.13 \text{ cm}^{-3}$. The star formation time-scale is assumed to scale
HI absorption on a moving mesh

2.2 Gas self-shielding

The gas in DLAs is self-shielded, with a neutral fraction close to unity. A full treatment requires radiative transfer and has been studied in e.g. Pontzen et al. (2008), Altay et al. (2011), Yajima et al. (2012) and Cen (2012). Yajima et al. (2012), using outputs from an SPH simulation post-processed with radiative transfer, suggested modelling the transition to self-shielded gas with a step function. We considered this, but found that it caused an artificial kink in the column density function, especially prominent for AREPO, although it did not affect any of our other results. Instead we fit the neutral fraction as a function of density to the results of their preferred simulation. Gas is assumed to transition between equilibrium with the UVB for \( \rho < \rho_\text{HI} \) and complete self-shielding for \( \rho > \rho_\text{HI} \), with the transition region given by

\[
n_{\text{HI}} = \frac{n_{\text{UVB}}^{\text{HI}}(\rho_\text{ss} - \rho)^p + (\rho - \rho_\text{ss})^p}{(\rho_\text{ss} - \rho_\text{HI})^p},
\]

where \( n_{\text{UVB}}^{\text{HI}} \) is the neutral fraction when in equilibrium with the UVB. The best-fitting values were \( p = 2.68, \rho_\text{ss} = 4.53 \times 10^{-3} \text{ cm}^{-3} \) and \( \rho_\text{HI} = 1.52 \times 10^{-2} \text{ cm}^{-3} \). Furthermore, we verified that this was consistent with the results of Altay et al. (2011), and that the DLA abundance is not sensitive to the exact position of the transition.

Nagamine et al. (2004a) proposed identifying self-shielding with the onset of star formation, i.e. \( \rho_\text{ss} = 0.1289 \text{ cm}^{-3} \). This is an order of magnitude higher than the value of \( \rho_\text{ss} \) suggested by radiative transfer and produces a slight, probably unphysical, coupling of the neutral fraction to the star formation rate. While we did not use this prescription, we did check its effect and found that, although the absolute value of many DLA properties changed (which is to be expected when changing the density threshold for self-shielding by an order of magnitude), the differences between AREPO and GADGET were largely unaffected. This gives us further confidence that our results are robust to changes in the self-shielding prescription.

Cen (2012) and Altay et al. (2011) included a prescription for the formation of molecular hydrogen. Altay et al. (2011) were motivated by Blitz & Rosolowsky (2006), who found that the surface density of molecular hydrogen in local galaxies was strongly correlated with the hydrostatic pressure. Altay et al. (2011) assumed the same relation held between molecular hydrogen density and pressure in the star-forming phase of the interstellar medium at high redshift. We considered such a prescription, but found that the only noticeable effect was on the column density function for \( N_{\text{HI}} > 10^{22} \text{ atoms cm}^{-2} \) (as found by Erkal et al. 2012) and was not sufficient to bring it into agreement with observations. Furthermore, we compared the shape and amplitude of the molecular hydrogen column density function predicted by the model at \( z = 0 \) to the observed values of Zwaan & Prochaska (2006) and found that they did not match. We believe both these facts are because the lack of feedback in our simulations led to a surplus of over-dense material. We therefore decided not to incorporate molecular hydrogen in our analysis until our simulations include strong feedback processes.

2.3 DLA selection and column density

We use a halo catalogue generated with the friends-of-friends (FOF) algorithm (Davis et al. 1985) and a linking length of 0.2 of the mean interparticle spacing. The FOF algorithm is applied only to the DM particles, whereas baryonic particles/cells are later assigned to their nearest DM particle and included in the halo to which the corresponding DM particle belongs. Self-bound concentrations of mass are identified within each FOF halo using the SUBFIND algorithm (Springel et al. 2001), modified to account correctly for baryons (Dolag et al. 2009). SUBFIND identifies halo substructure by generating an adaptively smoothed density field and searching for gravitationally bound overdensities. Following a common convention in the literature, we use \( M_{200} \) as an estimate of the halo virial mass, \( M_{200} = m_\text{HI} \) where \( m_\text{HI} \) is the mass enclosed within a spherical region of radius \( R_{200} \) within which the mean density is 200 times the critical density.

We consider only resolved haloes with \( M > 400(\Omega_m/\Omega_b) m_\text{HI} \), where \( m_\text{HI} \) is both the SPH gas particle mass and the initial mass of AREPO cells. For our \( 2 \times 512^3 \) simulation, this implies a halo resolution threshold of \( M \geq 2 \times 10^6 \text{ h}^{-1} \text{M}_\odot \). We note that our conclusions are unaffected by the exact placing of this limit. To avoid confusing overlaps, we remove haloes with larger neighbours closer than \( R_{200} \) from our catalogue.

To find the projected neutral hydrogen column density around each halo, we consider a grid of size \( 2R_{200} \) centred on each halo,
divided into equal-sized cells with side length equal to the gravitational softening length (for us \(\sim 1 \text{ h}^{-1} \text{kpc}\)). Gas elements in the halo are projected on to this grid, using an SPH kernel. In the case of AREPO, the SPH smoothing length is chosen so that the volume covered by the SPH kernel is identical to the volume of the moving-mesh cell. We also considered a cloud-in-cell kernel and, like Nagamine et al. (2004a), found that it made negligible difference to our results.

We calculated the column density along a line of sight by projecting our interpolated density field along a single direction, here the x-axis. The column density is given by

\[
N_{\text{HI}} = \sum_i \frac{\rho_\text{HI}(x)}{m_p} \epsilon (1 + z)^2,
\]

where \(m_p\) is the proton mass, \(\epsilon\) is the side length of a single grid cell and \(\rho_\text{HI}(x)\) is the average neutral hydrogen density inside the cell. The factor of \((1 + z)^2\) enters because \(\rho_\text{HI}\) is in comoving units and \(N_{\text{HI}}\) is in physical units. We checked different projection directions and found that the overall statistical properties of DLAs were unchanged. However, since galaxies in AREPO are more disc-shaped, they tended to look somewhat different when viewed edge-on. We also verified that our results are robust to changes in the resolution of the grid used for map making.

### 2.4 Column density function

The column density function, \(f(N_{\text{HI}})\), is defined observationally 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 density on the grid. Equating grid cells with sightlines assumes that two simulated DLAs will rarely be found along the same sightline, which is a good assumption given the small size of our box. More explicitly, we define the column density function by

\[
f(N) = \frac{F(N)}{\Delta N} \Delta X(z),
\]

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. As described by Bahcall & Peebles (1969) and Nagamine et al. (2004a), the (dimensionless) absorption distance is given by

\[
X(z) = \int_0^z (1 + z')^2 \frac{H_0}{H(z')} \, dz',
\]

and, for a box of comoving length \(\Delta L\), we have \(\Delta X = (H_0/c)(1 + z)^2 \Delta L\).

### 3 RESULTS

In this section, we present our results for the comparison between AREPO and GADGET. In Sections 3.1 and 3.2, we look at the effects on large and small haloes, respectively. Then we examine various statistical properties of the DLAs: Section 3.3 considers the DLA and LLS cross-sections, Section 3.4 considers the observed DLA abundance and Section 3.5 considers the column density function. Finally, in Section 3.6 we discuss our results for the Ly\(\alpha\) forest. We emphasize that we have checked explicitly that our results are unchanged when using simulations with \(2 \times 256^3\) particles – a factor of 8 lower mass resolution – at least for haloes above the resolution limit of the lower resolution simulations.

#### 3.1 Large haloes

Fig. 1 shows the distribution of neutral hydrogen in the largest halo in our simulation for \(z = 4 - 2\). For column densities with \(N_{\text{HI}} > 10^{19}\) atoms cm\(^{-2}\), gas is almost entirely self-shielded and thus is traced extremely well by the neutral hydrogen. We can see that GADGET produces a large number of small, circular, gaseous artefacts, which are largely absent in AREPO. Similar ‘blobs’ have been seen in other cosmological SPH simulations (including calculations with different SPH codes), and we interpret their existence as a numerical artefact due to the suppression of fluid instabilities and mixing in SPH; see Torrey et al. (2011) for further details. In AREPO, these SPH blobs make up the bulk of the LLS, especially in the column density range \(10^{17} < N_{\text{HI}} < 10^{19}\) atoms cm\(^{-2}\). AREPO reveals that LLS are more commonly produced from distinct filamentary structures. Fig. 2 shows this more explicitly. Here we have discretized the column density map; any cell with \(N_{\text{HI}} < 10^{17}\) cm\(^{-2}\) is shown in white, cells producing LLS are orange, while DLAs are red.

There is a more subtle change in the DLA cross-section. DLAs in AREPO are more concentrated in the centre of the halo, but the overall cross-section is not significantly changed; although there are fewer substructures, this is partially compensated by the higher accretion rate of the central halo.

#### 3.2 Small haloes

Fig. 3 shows the radial profile obtained by stacking all haloes with mass \(3 \times 10^9 < M < 3.5 \times 10^9\) \(\text{M}_\odot\) at \(z = 3\). We stacked all haloes in the given mass range. These haloes are approximately spherical, so the effect of substructure is relatively small. The column density in the central \(5\ h^{-1}\) kpc of the halo is larger by a factor of about 7 in GADGET. This effect is more pronounced at redshift \(z = 4\) (a factor of 10) than at \(z = 2\) (a factor of 3). This does not appear to be due to resolution effects; the density of the innermost \(5\ h^{-1}\) kpc is larger in GADGET for all haloes with mass up to \(M = 5 \times 10^8\) \(\text{h}^{-1}\) \(\text{M}_\odot\), in both the \(2 \times 512^3\) and \(2 \times 256^3\) particle simulations. However, for \(M > 10^{10}\h^{-1}\) \(\text{M}_\odot\) (and at \(z = 2\)), the characteristic size of a DLA is much larger than \(5\ h^{-1}\) kpc, so changes in the central density have a much smaller effect.

#### 3.3 DLA and LLS cross-section

We define the DLA cross-section, \(\sigma_{\text{DLA}}\), of a halo to be the area covered by all grid cells with column density \(N_{\text{HI}} > 10^{20.3}\) atoms cm\(^{-2}\). The LLS cross-section, \(\sigma_{\text{LLS}}\), is defined similarly, but with \(10^{20.3} > N_{\text{HI}} > 10^{17}\) atoms cm\(^{-2}\).

Fig. 4 shows the DLA cross-section, \(\sigma_{\text{DLA}}\), as a function of halo mass. The shaded region delineates the area in the \(M - \sigma_{\text{DLA}}\) plane populated by haloes. Symbols with error bars show the median, upper and lower quartiles of \(\sigma_{\text{DLA}}\) for haloes in seven mass bins. In order to fit the features at small and large halo masses, we modelled \(\sigma_{\text{DLA}}\) and \(\sigma_{\text{LLS}}\) with a three-component power law

\[
\sigma = \left(\frac{M}{M_0}\right)^a + \frac{d}{10^{8.5}} \left(\frac{M}{M_0}\right)^b 10^{(b-N)/5} - c,
\]

where \(M_0 = 10^{10.5}\h^{-1}\) \(\text{M}_\odot\) and \(N = 20.3\) for DLAs or \(N = 17\) for LLS. The free parameters \(a, b, c, d\) and \(e\) are found by a simultaneous
fit to \( \sigma_{\text{DLA}} \) and \( \sigma_{\text{LLS}} \),\(^1\) using \textsc{mpfit} (Markwardt 2009).\(^2\) The resulting numerical values are listed in Table 1 and the fit is shown by the lines in Fig. 4. This procedure is similar to that outlined in Nagamine et al. (2004a), except that we have excluded haloes which do not form DLAs from our fit, instead of assigning them an arbitrary small cross-section as in that work.

\(^1\) A consequence of fitting to both quantities together is that at \( z = 3 \), the best-fitting line matches \( \sigma_{\text{LLS}} \) but is slightly above the median \( \sigma_{\text{DLA}} \).

\(^2\) As ported to \textsc{python} by Mark Rivers and Sergey Koposov.

Qualitatively, our results for \( \sigma_{\text{DLA}} \) in \textsc{gadget} are in good agreement with those of Nagamine et al. (2004a) for overlapping halo masses. They found that, for SPH, a turnover in the \( \sigma_{\text{DLA}} \)-halo mass relation occurred for \( M \approx 10^{8.5} h^{-1} M_\odot \), below the resolution limit of our simulations. They also found a slight excess in \( \sigma_{\text{DLA}} \) over their bare power law for large haloes, which we have fit for explicitly. Our \( \sigma_{\text{DLA}} \) has a somewhat smaller amplitude than the O3 model of Nagamine et al. (2004a) and the \( \sigma_{\text{DLA}} \)-halo mass relation is somewhat shallower in slope. This is in agreement with the results of Tescari et al. (2009) and is probably due to a difference in cosmological parameters; Nagamine et al. (2004a) used \( \sigma_8 = 0.9 \), while we have \( \sigma_8 = 0.8 \).
Haloes in the mass range $10^{11} > M > 10^{10} h^{-1} \, M_\odot$ show similar cross-sections in both AREPO and GADGET. The extra substructure discussed in Section 3.1 leads to much larger $\sigma_{\text{LLS}}$ in GADGET for haloes of mass larger than $10^{11} h^{-1} \, M_\odot$, especially at low redshift, because of the increased number of massive haloes. This effect is smaller for $\sigma_{\text{DLA}}$, as discussed above, but still present.

Due to the effect discussed in Section 3.2, $\sigma_{\text{DLA}}$ is significantly reduced in AREPO for small haloes, producing a turnover in the power-law relation between mass and DLA cross-section. This turnover occurs because haloes below a certain size have a central column density which is on average less than the DLA cut-off, $N_{\text{HI}} = 10^{20.3} \, \text{cm}^{-2}$. $R$ is comoving, but the radial density is in physical units.

3.4 Incidence rate of DLA systems

Following Nagamine et al. (2004a), we calculate the incidence of DLAs by convolving the halo mass function with our fit to $\sigma_{\text{DLA}}$. This allows us to account for DM haloes of smaller mass than the resolution limit of the simulation. The cumulative number of DLAs per unit redshift is then defined as

$$\frac{dN_{\text{DLA}}}{dz}(> M, z) = \frac{dr}{dz} \int_{M}^{\infty} \frac{dn_h}{dM} \sigma_{\text{DLA}} dM,$$

where $dn_h/dM$ is the Sheth–Tormen DM halo mass function, $\sigma_{\text{DLA}}$ is given by equation (5), truncated at $M = 10^{12.5} h^{-1} \, M_\odot$ and

$$\frac{dr}{dz} = \frac{c}{H_0} \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$$

defines the line element.

Fig. 5 shows $dN_{\text{DLA}}/dz$ for the AREPO and GADGET simulations. The yellow band marks an observational estimate of the total DLA abundance,

$$\log_{10} \left( \frac{dN_{\text{DLA}}}{dz} \right) = -0.6 \pm 0.1,$$

recovered by Nagamine et al. (2007) from the data of Prochaska et al. (2005). Our results for the GADGET simulation are similar to the weak wind scenario of Tescari et al. (2009) and produce more DLAs than are observed. They showed that this discrepancy could be reduced with the addition of feedback processes. We find that the lower DLA cross-section of small and large mass haloes in AREPO combine to reduce the cumulative DLA abundance by a factor of 2, bringing it almost into agreement with the upper limit from observations.

3.5 Column density function

Fig. 6 shows the neutral hydrogen column density function, defined in Section 2.4. Our results suggest a significantly shallower column density function than preferred by observations. This was also seen by Nagamine et al. (2004a) and Tescari et al. (2009), who found that adding feedback to the simulations produced better agreement.
HI absorption on a moving mesh

Figure 4. Left-hand panels: the comoving DLA cross-section, $\sigma_{\text{DLA}}$. Right-hand panels: the LLS cross-section, $\sigma_{\text{LLS}}$. Each row shows a different redshift: $z = 4$ (top), $z = 3$ (middle) and $z = 2$ (bottom). Regions containing at least one GADGET halo are shown in light blue, while regions containing at least one AREPO halo are shown in red. The symbols with error bars show the median DLA cross-section in seven evenly spaced mass bins; the error bars show the upper and lower quartiles. Red triangles are for AREPO, while blue squares are for GADGET. The symbols for GADGET have been offset horizontally by 10 per cent for visibility. The lines denote the results of our fit; red solid lines are for AREPO, while blue dashed lines are for GADGET.

We will address this question in future work based on forthcoming AREPO simulations with stronger wind feedback.

Overall, the differences in the column density function between the codes are fairly limited; a change of $\sim 30$ per cent of the total. At first glance this may seem puzzling; why have the generally smaller AREPO haloes made so little difference to the column density function? Fig. 7 shows the fraction of the column density function contributed by haloes in three mass ranges, corresponding to those in Sections 3.1 and 3.2. The column density function has been calculated only for those haloes with the desired mass and then normalized using the total GADGET column density function. This allows us to clearly see the differences between the codes. The reduced substructure in $M > 10^{11} \, h^{-1} M_\odot$ haloes does lead to a reduced column density for these haloes at $N_{\text{HI}} \sim 10^{20}$ atoms cm$^{-2}$,
Table 1. Numerical parameters of the fit described in equation (5).

<table>
<thead>
<tr>
<th>Redshift</th>
<th>Code</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>AREPO</td>
<td>0.593</td>
<td>33.1</td>
<td>74.0</td>
<td>1500</td>
<td>1.07</td>
</tr>
<tr>
<td>4</td>
<td>GADGET</td>
<td>0.429</td>
<td>33.5</td>
<td>79.8</td>
<td>1070</td>
<td>2.97</td>
</tr>
<tr>
<td>3</td>
<td>AREPO</td>
<td>0.496</td>
<td>33.8</td>
<td>101</td>
<td>-474</td>
<td>0.529</td>
</tr>
<tr>
<td>3</td>
<td>GADGET</td>
<td>0.625</td>
<td>32.7</td>
<td>-27.3</td>
<td>316</td>
<td>2.73</td>
</tr>
<tr>
<td>2</td>
<td>AREPO</td>
<td>0.518</td>
<td>33.6</td>
<td>66.7</td>
<td>-1120</td>
<td>0.787</td>
</tr>
<tr>
<td>2</td>
<td>GADGET</td>
<td>0.849</td>
<td>31.5</td>
<td>-50.0</td>
<td>20.1</td>
<td>2.51</td>
</tr>
</tbody>
</table>

Figure 5. Cumulative abundance of DLAs per unit redshift as a function of total halo mass at \( z = 3 \) for AREPO (red solid line) and GADGET (blue dashed line). The yellow bar shows the observed total cumulative DLA abundance of Prochaska et al. (2005).

but their rarity means that they make up only for a small fraction of the total column density function. The largest change is produced by the reduction in cross-section of \( M < 10^{10} h^{-1} M_\odot \) haloes, although this too is diluted because \( 10^{10} < M < 10^{11} h^{-1} M_\odot \) haloes, which make up almost as great a fraction of the total column density, are mostly unchanged. This explains why differences become larger with increasing redshift; we found in Section 3.2 that the effect on \( M < 10^{10} h^{-1} M_\odot \) haloes became larger at early times.

Both codes produce a slight kink at \( 10^{19} - 10^{20} \) atoms cm\(^{-2}\). This feature is more prominent in AREPO, and is a consequence of the transition between self-shielding and UVB ionization equilibrium, made more prominent by the smoother gas distribution in AREPO.

3.6 The Ly\( \alpha \) forest

For an analysis of the Ly\( \alpha \) forest, we generated 16 000 simulated Ly\( \alpha \) spectra with 1024 pixel each from our simulations. The positions of the spectra were chosen at random, and particles/cells were interpolated to the lines of sight using an SPH kernel. We verified that using cloud-in-cell interpolation for AREPO did not affect our results. The optical depth from the absorption due to each particle was calculated as described in detail in Bird et al. (2011). In order to avoid contaminating the spectra with strong absorbers, we did not apply a self-shielding correction here. Bolton et al. (2008) found evidence for an inverted temperature–density relation in the Ly\( \alpha \) forest, so that lower density regions have a higher temperature. Mechanisms proposed for reproducing this include helium reionization (McQuinn et al. 2009) or volumetric heating from blazars (Puchwein et al. 2012). As our purpose in this paper is a code comparison, we did not attempt to model this in our simulations. Thus, our model produces a temperature–density relation \( T \propto \rho^{\gamma} \), where \( \gamma \) asymptotes towards \( 1/1.7 \) at low redshift (Hui & Gnedin 1997) rather than the observed value of \( \gamma \sim 1 \). At \( z = 3 \) we have \( \gamma = 1.55 \).

We define the transmitted flux as \( \mathcal{F} = \exp(-\tau) \), where \( \tau \) is the optical depth. A completely transparent Universe has \( \mathcal{F} = 1 \). Observations have determined the effective optical depth, \( \tau_{\text{eff}} = \mathcal{F} \), where \( \mathcal{F} \) is the mean transmitted flux, the one-dimensional flux probability
distribution function (PDF; a histogram of the flux from each spectral pixel) and the flux power spectrum. We extracted all three of these quantities from our simulations and compared the results of AREPO and GADGET. We checked convergence using the simulations with 256³ particles. Fig. 8 shows the effect of a changing particle number on τ_{eff}. GADGET converged at the 2 per cent level for τ_{eff} and AREPO to ~4 per cent, but this convergence becomes significantly poorer for z > 3.5 in both codes. Bolton & Becker (2009) found that this occurs because at high redshift the Lyα transmission is dominated by progressively less dense regions, which are less well resolved.

In more detail, the effective optical depth is slightly lower in AREPO, by 4 per cent at z = 4, 2 per cent at z = 3 and 1 per cent at z = 2. Note that τ_{eff} is known observationally to ~4 per cent at z = 3 (Viel et al. 2009). This small difference is due to a slight increase in the mean temperature, T_{eff}, of the Lyα absorbers, defined to be gas with T < 10⁴ K and ρ < ρ_{c}, where ρ_{c} is the critical density. T_{eff} in AREPO is ~240 K (2 per cent) higher than that in GADGET at z = 2–4. The Lyα forest is assumed to be in ionization equilibrium with the UVB, so a higher temperature produces a greater ionization fraction, thus decreasing the effective optical depth. These changes are somewhat larger for the lower resolution simulations and thus are likely to be reduced further with higher numerical resolution. It is the standard procedure, when comparing to data, to rescale simulated spectra to have the same mean flux as is observed. As we are performing a code comparison, we do not rescale our spectra for the presented results, but we checked that it did not significantly affect our conclusions.

Fig. 9 shows the flux PDF extracted from our simulations in three redshift bins, chosen to match observations. There are small differences between the codes; for F < 0.6 the flux PDF is larger in AREPO, but for higher transmission regions at z = 3, the trend is reversed and AREPO produces a lower flux PDF. Fig. 10 shows a similarly sized effect on the power spectrum of the flux; AREPO predicts a slightly lower power spectrum. For comparison, the typical uncertainty in the observed flux PDF is 5–7 and 5–10 per cent in the flux power spectrum. These are of the same order as the differences we find here between the numerical codes. We attribute these changes in the flux PDF and flux power primarily to subtle shifts in the temperature–density distribution. Fig. 11 shows a mass-weighted histogram of the temperature and density of gas elements. We can see that although the particles follow a similar temperature–density relation overall, AREPO produces a wider spread in temperatures for gas near the critical density; this trend continues for higher density gas until ρ ~ 10ρ_{c}.

This difference was somewhat larger for lower resolution simulations with 256³ particles, which might naively suggest that AREPO is converging more slowly in underdense regions than GADGET. However, it is not completely clear that the two codes are converging to the same result because the differences we see could well be the effect of weak structure formation shocks. While these are followed accurately in AREPO, they are largely lost in GADGET, due to SPH’s poorer shock-resolving capability (see e.g. Keshet et al. 2003). We would thus expect that a fully resolved AREPO simulation would still produce more gas elements scattered off the mean temperature–density relation than GADGET, consistent with our results. We checked, using a simulation with a reduced Courant factor, that Fig. 11 was not affected by the time step.

We recall that we have not included feedback in our simulations; this is expected to affect the Lyα flux PDF at the 5 per cent level (Bolton et al. 2008; Viel, Schaye & Booth 2012) at z = 2–2.5. Feedback processes such as galactic winds may interact with the gas dynamics differently in AREPO than in GADGET, potentially...
adding further systematic differences between the codes of a similar magnitude.

4 CONCLUSIONS

We compared the statistics of neutral hydrogen absorption in the SPH code GADGET with that in the moving-mesh code AREPO. Our aim was to understand how known differences in the treatment of the fluid equations manifest themselves as changes in the observable properties of DLAs, LLS and the Lyα forest. There were significant qualitative differences: in GADGET, DLAs and LLS in large haloes were primarily produced by small spherical objects, and the only difference between the two classes of absorbers was the central density. However, in the AREPO simulations, DLAs and LLS came from quite different sources. The bulk of the DLA cross-section in a large halo was from a central disc, while LLS were produced in filamentary structures. This made little quantitative difference to the DLA cross-section, but it suggests a different interpretation of high column density systems and could potentially be reflected kinematically in detailed line profile shapes. We would suggest that any future studies sensitive to the detailed distribution of neutral hydrogen inside massive haloes, especially LLS, should avoid using SPH.

We found that GADGET produced clouds of clumpy substructure around haloes with $M > 10^{11}$ $h^{-1}$ $M_\odot$, which were essentially absent in AREPO. These GADGET clumps had a peak column density of $10^{18}$–$10^{19}$ cm$^{-2}$ and so boosted the LLS cross-section of these haloes significantly. There was a similar, but somewhat smaller, effect on the DLA cross-section. Furthermore, haloes in AREPO had central densities which were lower than their counterparts in GADGET. This significantly lowered the DLA cross-section in AREPO for objects sensitive to the density in the central $5$ $h^{-1}$ kpc (in practice, haloes with $M < 10^{10}$ $h^{-1}$ $M_\odot$ at $z > 2$), but did not greatly affect the LLS.

These systematic changes made the halo mass–cross-section relation shallower where it was dominated by the changes to large haloes and steeper where it was dominated by changes to small haloes. The former occurred for LLS, and DLAs at redshift $z = 2$, and the latter for DLAs at $z = 3$ and 4. Furthermore, both changes acted to reduce the overall abundance of DLAs in AREPO. The DLA abundance for GADGET simulations is somewhat in excess of the observed value. This discrepancy can be removed with the incorporation of galactic winds into the simulation (Nagamine et al. 2004a; Tescari et al. 2009), but we found that our AREPO simulations already substantially reduce it, even without strong feedback. This was mostly due to the reduced cross-section of small haloes; large haloes are sufficiently rare that changes in their cross-section do not have a large effect on the total abundance. It also shows that a calibration of feedback strength from the DLA abundance would be compromised by numerical effects in SPH.

The column density function shows a similar pattern: a modest reduction in amplitude for $N_{\text{HI}} = 10^{20}$–$10^{22}$ cm$^{-2}$ and $z > 2$, driven by the lower central density of small haloes. We found that the amplitude of the column density function for $N_{\text{HI}} > 10^{21}$ cm$^{-2}$ is still much larger than observed. This discrepancy is much larger than the differences between our two simulations, which can be interpreted as evidence that some feedback in the form of outflows is needed. Although the column density function for large haloes was significantly reduced in AREPO by the changes in their substructure, the relative rarity of these haloes meant that they did not have a strong impact on the total column density function. Note, however, that feedback processes reduce the amplitude of the high-end column density function by removing preferentially the high-column density cells in small haloes; the more massive the halo, the less it is affected by winds (Tescari et al. 2009). Thus, although it appears as if the reduced substructure around large haloes only produces subtle changes in this statistic, once galactic winds have blown the small haloes away, what remains may be very greatly affected.

Finally, we examined basic Lyα forest statistics and found that differences between the codes were typically at the few per cent level. This was due partly to a slightly changed thermal history (which is typically marginalized out in cosmological studies of the Lyα forest) and partly to an increased width in the temperature–density relation of the gas in AREPO. These differences are small when compared to changes in the high column density systems. This is not unexpected; the evolution of the Lyα forest is dominated by gravitational effects, hence issues such as the accuracy problems of SPH for fluid instabilities do not play an important role (a similar result was found by Regan, Haehnelt & Viel 2007). They are, however, comparable to the statistical uncertainties of current data.

Figure 11. Two-dimensional mass-weighted histograms of the total gas mass as a function of temperature and density, at $z = 3$, with $2 \times 512^3$ resolution. The colour scale shows the total mass of gas elements found in a given temperature–density bin.
and thus may bias derived parameters. Although a full cosmological analysis is beyond the scope of this paper, it seems likely that the largest difference will occur in the derived thermal history of the IGM, since the change in the flux power spectrum came predominantly from an alteration in the temperature–density relation. These effects will also have to be taken into account when analysing upcoming Lyα experiments such as the Baryon Oscillation Sky Survey (Slosar et al. 2011), which are expected to have statistical errors an order of magnitude smaller than those of current data.

There are many questions about high column density systems which remain unanswered by the present work. Although our AREPO simulations reduce the discrepancy between DLA simulations and observations, completely eliminating it still seems to require stronger feedback processes. Furthermore, SPH simulations and observations, completely eliminating it still seems to require stronger feedback processes. Further studies are needed to clarify the role of feedback in these systems. 

ACKNOWLEDGMENTS

SB would like to thank Yajima Hidenobu for sending us his collection of observational data on the column density function. SB is supported by the National Science Foundation grant number AST-0907969 and the Institute for Advanced Study. DS acknowledges financial support from the Klaus Tschira Foundation and the Deutsche Forschungsgemeinschaft through SFB 881, ‘The Milky Way System’. LH is supported by NASA ATP Award NNX12AC67G.

REFERENCES

Komatsu E. et al., 2011, ApJS, 192, 18

HI absorption on a moving mesh 3351
Seljak U. et al., 2005, Phys. Rev. D, 71, 103515

This paper has been typeset from a TeX/LaTeX file prepared by the author.