Absorption-line systems in simulated galaxies fed by cold streams

Michele Fumagalli,[1]⋆ J. Xavier Prochaska,[1,2] Daniel Kasen,[3,4] Avishai Dekel,[5] Daniel Ceverino[5] and Joel R. Primack[6]

[1]*Department of Astronomy and Astrophysics, University of California, 1156 High Street, Santa Cruz, CA 95064, USA*
[2]*UCO/Lick Observatory, University of California, 1156 High Street, Santa Cruz, CA 95064, USA*
[3]*Department of Physics, University of California, 366 LeConte, Berkeley, CA 94720, USA*
[4]*Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[5]*Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*
[6]*Department of Physics, University of California, 1156 High Street, Santa Cruz, CA 95064, USA*

ABSTRACT

Hydro-cosmological simulations reveal that massive galaxies at high redshift are fed by long narrow streams of merging galaxies and a smoother component of cold gas. We post-process seven high-resolution simulated galaxies with radiative transfer to study the absorption characteristics of the gas in galaxies and streams, in comparison with the statistics of observed absorption-line systems. We find that much of the stream gas is ionized by UV radiation from background and local stellar sources, but still optically thick ($N_{\mathrm{H\,I}} > 10^{17}$ cm^{-2}) so that the streams appear as Lyman-limit systems (LLSs). At $z > 3$, the fraction of neutral gas in streams becomes non-negligible, giving rise to damped Lyman α absorbers (DLAs) as well. The gas in the central and incoming galaxies remains mostly neutral, responsible for DLAs. Within one (two) virial radii, the covering factor of optically thick gas is <25 per cent (10 per cent) for LLSs and <5 per cent (1 per cent) for DLAs, slowly declining with time following the universal expansion. Nevertheless, galaxies and their cold streams in the studied mass range, $M_{\mathrm{vir}} = 10^{10}$–$10^{12}$ M$_\odot$, account for >30 per cent of the observed absorbers in the foreground of quasars, the rest possibly arising from smaller galaxies or the intergalactic medium. The mean metallicity in the streams is ∼1 per cent solar, much lower than in the galaxies. The simulated galaxies reproduce the Lyα-absorption equivalent widths observed around Lyman-break galaxies, but they severely underpredict the equivalent widths in metal lines, suggesting that the latter may arise from outflows. We conclude that the observed metal-poor LLSs are likely detections of the predicted cold streams. Revised analysis of the observed LLSs kinematics and simulations with more massive outflows in conjunction with the inflows may enable a clearer distinction between the signatures of the various gas modes.

Key words: radiative transfer – galaxies: evolution – galaxies: formation – galaxies: high-redshift – intergalactic medium – quasars: absorption lines.

1 INTRODUCTION

Cold neutral hydrogen that fuels star formation is an important driver of galaxy evolution at all times. Recent progress in observations and simulations has remarkably improved our view of gas in the early Universe. Molecular gas in emission is detected in normal star-forming galaxies to $z \sim 2.3$ (Daddi et al. 2010; Tacconi et al. 2010). Large imaging and spectroscopic surveys have increased by orders of magnitude the samples of quasars and galaxy pairs useful for probing in absorption the interstellar medium (ISM), the circumgalactic medium (CGM) and the intergalactic medium (IGM) (e.g. Adelberger et al. 2003; Hennawi et al. 2006; Noterdaeme et al. 2009; Prochaska & Wolfe 2009; Prochaska, O'Meara & Worseck 2010; Steidel et al. 2010). Similarly, the increased resolution of numerical simulations has made it possible to reproduce the morphology of galaxies especially at high redshift (e.g. Agertz, Teyssier & Moore 2009; Dekel, Sari & Ceverino 2009a; Ceverino, Dekel & Bournaud 2010).

Significant progress is being made on the basic issue of how gas is accreted from the IGM into galaxies. A consistent theoretical picture is emerging from both analytic theory (Birnboim &

⋆E-mail: mfumagalli@ucolick.org

Dekel 2003; Dekel & Birnboim 2006) and numerical simulations (Kereš et al. 2005, 2009; Ocvirk, Pichon & Teyssier 2008; Dekel et al. 2009b). High-redshift galaxies are predicted to acquire the bulk of their gas mass at temperatures of a few 10^4 K via long narrow streams that consist of a smooth component and small merging galaxies. However, indisputable observational evidence for this predicted dominant mode of gas input is still lacking.

Lyα emission due to cooling radiation has been proposed as a signature of cold gas accretion that powers Lyα blobs (e.g. Haiman, Spaans & Quataert 2000; Fardal et al. 2001; Furlanetto et al. 2005; Dijkstra & Loeb 2009; Goerdt et al. 2010; Matsuda et al. 2010), but the resonant nature of the Lyα line and the different physical mechanisms that influence the total Lyα luminosity (Faucher-Giguère et al. 2010, Kasen et al., in preparation) prevent us from unambiguously connecting the observed fluxes with the presence of cold streams. Observations of gas in absorption against background sources are, in principle, ideal to probe the CGM to uncover infalling gas. However, it has been suggested that the low covering factor of the cold filaments reduces the probability of detecting them in absorption (Stewart et al. 2010; Faucher-Giguère & Kereš 2011), especially in the presence of large-scale outflows (Steidel et al. 2010).

The subject of the present paper is the problem of detecting streams of cold gas in absorption against background sources. We formulate detailed theoretical predictions of the covering factor, the cross-section, the column density distribution, the kinematics and the metallicity of the neutral gas in seven galaxies, simulated with a high-resolution adaptive mash refinement (AMR) code in a cosmological context and analysed in the redshift range $z = 1.4$–4. To facilitate the comparison between our theoretical predictions and observations, we adopt the same methodology used in observational studies of absorption-line systems (ALSs) in background quasars and galaxy pairs. This work focuses on the visibility of cold stream, and by comparing and contrasting these predictions with observations, we aim to find the observables that are best suited to probe this mode of accretion. A companion paper (Kasen et al., in preparation) will focus on the Lyα emission properties using the same simulations.

The paper consists of three parts. In the first two sections, we briefly discuss the numerical simulations (Section 2) and the radiative transfer (RT) calculation used to estimate the hydrogen ionization state (Section 3). This part is complemented with Appendix A, where we provide detailed comparisons of the neutral hydrogen masses and column density distributions from three different RT models. In the second part, we quantify the neutral hydrogen cross-section and covering factor as functions of the gas column density and redshift (Section 4). In the third part, we compare more directly model predictions with observations, focusing in particular on the column density distribution and incidence of ALSs (Section 5) as well as their metallicity and kinematics (Section 6). A summary and conclusions follow in Section 7.

Throughout this work, we assume the *Wilkinson Microwave Anisotropy Probe* (*WMAP*5) values of the cosmological parameters (Komatsu et al. 2009), consistent with the cosmology adopted in the simulations ($\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$).

2 THE SIMULATIONS

2.1 Simulation method

A more detailed discussion of the simulations used in this work and comparisons with properties observed in high-redshift galaxies can be found in Ceverino et al. (2010). Here, we briefly summarize the

numerical procedures adopted and review some of the uncertainties associated with these simulated galaxies that are relevant to this analysis.

The simulations have been run with the AMR hydro-gravitational code ART (Kravtsov, Klypin & Khokhlov 1997; Kravtsov 2003). The code incorporates many of the relevant physical processes for galaxy formation, including gas cooling and photoionization heating, star formation, metal enrichment and stellar feedback (Ceverino & Klypin 2009; Ceverino et al. 2010). Cooling rates were tabulated for a given gas density, temperature, metallicity and UV background (UVB) based on CLOUDY (Ferland et al. 1998) and assuming cooling at the centre of a 1 kpc thick cloud, illuminated by a uniform radiation field. Photoheating is self-consistently modelled together with radiative cooling, through the same CLOUDY tables, assuming a uniform UVB (Haardt & Madau 1996). Self-shielding of dense, galactic neutral hydrogen from the UVB is approximated by suppressing the UVB intensity to the prereionization value ($5.9 \times 10^{-26}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}}$) above gas densities $n_H = 0.1\,\mathrm{cm^{-3}}$. This threshold is consistent with the results from RT calculations (see Appendix A).

Stars form according to a stochastic model that is roughly consistent with the Kennicutt (1998) law in cells where the gas temperature is below 10^4K and the density is above the threshold $n_H = 1\,\mathrm{cm^{-3}}$. The code implements a feedback model, in which the combined energy from stellar winds and supernova explosions is released at a constant heating rate over 40 Myr, the typical age of the lightest star that explodes as a Type II supernova (SN II). Energy injection by Type Ia supernovae (SN Ia) is also included. The heating rate from SN Ia assumes an exponentially declining rate with a maximum at 1 Gyr.

The ISM is enriched by metals from SN II and SN Ia. Metals are assumed to be released from each stellar particle by SN II at a constant rate for 40 Myr since its birth, assuming a Miller & Scalo (1979) initial mass function (IMF) and matching the yields of Woosley & Weaver (1995). The metal ejection by SN Ia assumes the same exponentially declining rate. The code treats metal advection self-consistently and metals can diffuse and pollute the medium around the central discs. In these simulations, the dark matter particle mass is $5.5 \times 10^5\,\mathrm{M_\odot}$, the minimum star particle mass is $10^4\,\mathrm{M_\odot}$ and the smallest cell physical size is always between 35 and 70 pc.

Each dark matter halo has been selected from a cosmological dark-matter-only simulation and resimulated with the addition of baryons, using a zoom-in technique that follows the evolution of the selected Lagrangian region of $\sim$1 comoving Mpc^3 (for details on the initial conditions see Ceverino & Klypin 2009). All the simulated galaxies are centrals to their haloes and the analysis has been performed on a box of four times the virial radius (R_{vir}), centred at all times on the main halo. Note that the selected galaxies do not reside in a sizable group or cluster at $z = 0$ because of the difficulties in simulating companion galaxies.

Although these simulations reproduce some of the main properties observed in high-redshift galaxies, like the Tully–Fisher and the mass–SFR (star formation rate) relations (see Ceverino et al. 2010), the use of approximated prescriptions in lieu of the complex interplay of physical processes that operate on scales smaller than our resolution limits the predictive power of these models and of simulations in general. Particularly relevant for our work is the uncertainty associated with stellar feedback, implemented according to the prescription of Ceverino & Klypin (2009). In these simulations, outflows eject hot gas and metals at high velocity outside the central discs into the IGM and CGM (see Ceverino & Klypin

2009). The total outflow mass flux reaches, in some cases, about one-third of the inflow flux. However, these winds may not be as strong as some observations suggest (e.g. Steidel et al. 2010) and active galactic nuclei feedback is missing. As a consequence, a fraction of the gas inside the discs might actually be ejected by stronger outflows, altering the covering fraction of neutral gas and the metallicity distribution in the CGM (see e.g. Faucher-Giguere, Keres & Ma 2011). Unless the low-density winds drastically perturb the shape and kinematics of the gas that inflows in dense narrow streams, these models capture the properties of the cold streams, but caution is advised in generalizing our results.

2.2 Sample properties

For each galaxy, the total dark matter and instantaneous stellar mass are computed summing all the particles within the virial radius. Similarly, we obtain the total SFR within the virial radius by averaging the masses of all the stars formed in the last 60 Myr. This time interval is arbitrarily chosen to be large enough to avoid fluctuations due to the stochastic SFR implemented in the simulation, but the average SFR does not depend on this assumption. The total hydrogen masses are computed from the hydrogen number density in each cell of the AMR grid as $m_H = m_p n_H l^3$, with m_p the proton mass and l the cell size. No correction for heavier elements, including helium, is applied throughout this work, except when we quote a gas fraction where we multiply the hydrogen mass by 1.38 to correct for heavier elements. Finally, we derive virial masses by combining together the baryonic and non-baryonic masses.

The sample is composed of seven massive galaxies that exhibit dark matter and baryonic properties roughly consistent with the observed population of Lyman break galaxies (LBGs). Galaxy properties as a function of time[1] are displayed in Fig. 1 and listed in Tables B1 and B2 in Appendix B. Panels (a) and (b) show the redshift evolution of the virial radius and dark matter mass. These galaxies span roughly more than a decade in mass. As seen in panel (b), they are all well above the characteristic Press–Schechter halo mass (M_*) in the relevant mass range, more so at higher redshifts. Galaxy MW1, and to some extent MW3, exhibits a faster overall growth rate than the other galaxies. The interval in mass and redshifts analysed here is typical for cold gas accretion; only at later times ($z < 2$) haloes approach the critical mass (horizontal dashed line) above which a stable shock develops and cold accretion becomes less efficient (Dekel & Birnboim 2006; Kereš et al. 2009).

The evolution of the SFR within the virial radius is presented in panel (c). While there is a tendency to have higher SFRs in higher mass haloes, the stochastic nature of the simulated SFR and different accretion histories produce a less ordered star formation history (SFH). The star formation in the sample brackets the typical values found in LBGs (shaded area for $z \sim 2$; Reddy et al. 2010), with MW8 and MW9 usually below the commonly observed range. The evolution of the stellar mass is shown in panel (d) and follows the general halo growth. Finally, the gas properties are displayed in panels (e) and (f) where we show the time evolution of the hydrogen mass and of the gas fraction, here defined as the ratio of gas mass to total baryonic mass. The simulated galaxies have a gas fraction comparable to the estimates from observed molecular gas at $z \sim$ 1–2 (Daddi et al. 2010; Tacconi et al. 2010). The baryonic fraction inside the virial radius is at all times below the universal value

$f_b = 0.165$, approaching 0.25–0.3 f_b when only gas and stars are considered inside the central discs. MW2 is an exception, having a particularly high baryon fraction at all redshifts, probably reflecting an extraordinary high gas accretion rate (Ceverino et al. 2010).

2.3 Satellite identification

Being part of a larger cosmological simulation, each box contains also satellites whose positions and sizes are used during the analysis to disentangle gas that resides in bound structures from gas in the streams. For this reason, we identify the satellites using the AMIGA halo finder (Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009). This code locates the centres of the haloes and iteratively computes their radii so that they define an overdensity Δ_{vir} over the background density ρ_b. For small child haloes embedded within larger parent haloes, this radius is defined instead as the distance at which the density profile reaches a minimum as one moves away from the halo centre. We adopt the internally computed Δ_{vir} from AMIGA as a function of redshift. We consider particles with velocities greater than the escape velocity to be unbound and we select only satellites with more than 182 particles. This sets a lower limit to the satellite halo mass of $\sim 10^8 \, M_\odot$.

While AMIGA outputs several parameters for each of the identified haloes, in the following we compute masses and SFRs considering all the particles that are enclosed in the defining radius, purely based on positions. Although not identical to the mass computed internally by AMIGA, these masses are well correlated with the ones generated by the code. This simple definition allows us to compute self-consistently the H I masses and SFRs. Due to this approximation, our satellite masses can be uncertain up to a factor of 2. Child haloes contribute to the parent halo masses.

3 ATOMIC HYDROGEN NEUTRAL FRACTION

3.1 The radiative transfer post-processing

A reliable estimate of the ionization state of the gas is essential for any study of ALSs. Ideally, the ionization state of the gas should be coupled to the hydrodynamic calculations, but due to the numerical complexity of this problem, codes that solve the (approximated) RT equation (e.g. Gnedin & Abel 2001; Abel & Wandelt 2002) together with the hydrodynamic equations has been developed and used only recently (e.g. Razoumov et al. 2006; Petkova & Springel 2009; Wise & Abel 2011) and they remain expensive for high-resolution simulations. Currently, RT post-processing is the practical way to compute the ionization state of the gas in our high-redshift simulations. The post-processing approximation is justified by the fact that the radiation–diffusion time-scale is short compared to the evolution time-scale.

The neutral fraction $x_{H I}$ of the atomic hydrogen in each AMR cell is determined using a Monte Carlo RT code that includes both collisional ionization and photoionization due to Lyman continuum radiation from both the cosmological UVB and local stellar sources, together with absorption and scattering by dust grains. This calculation improves upon results based on simple ray-tracing schemes. In fact, Monte Carlo transport does not introduce preferred directions in the photon path and takes into account the scattering of ionizing photons. Thus, a diffuse component of the UV radiation field is naturally included (about 32 per cent of the time, photons scatter off of hydrogen atoms rather than being absorbed). Further, radiation from stars and dust absorption is modelled locally for each stellar particle in order to reproduce the anisotropy of the radiation field

[1] Galaxies MW2, MW4 and MW5 were not available at low redshifts when this sample was assembled.
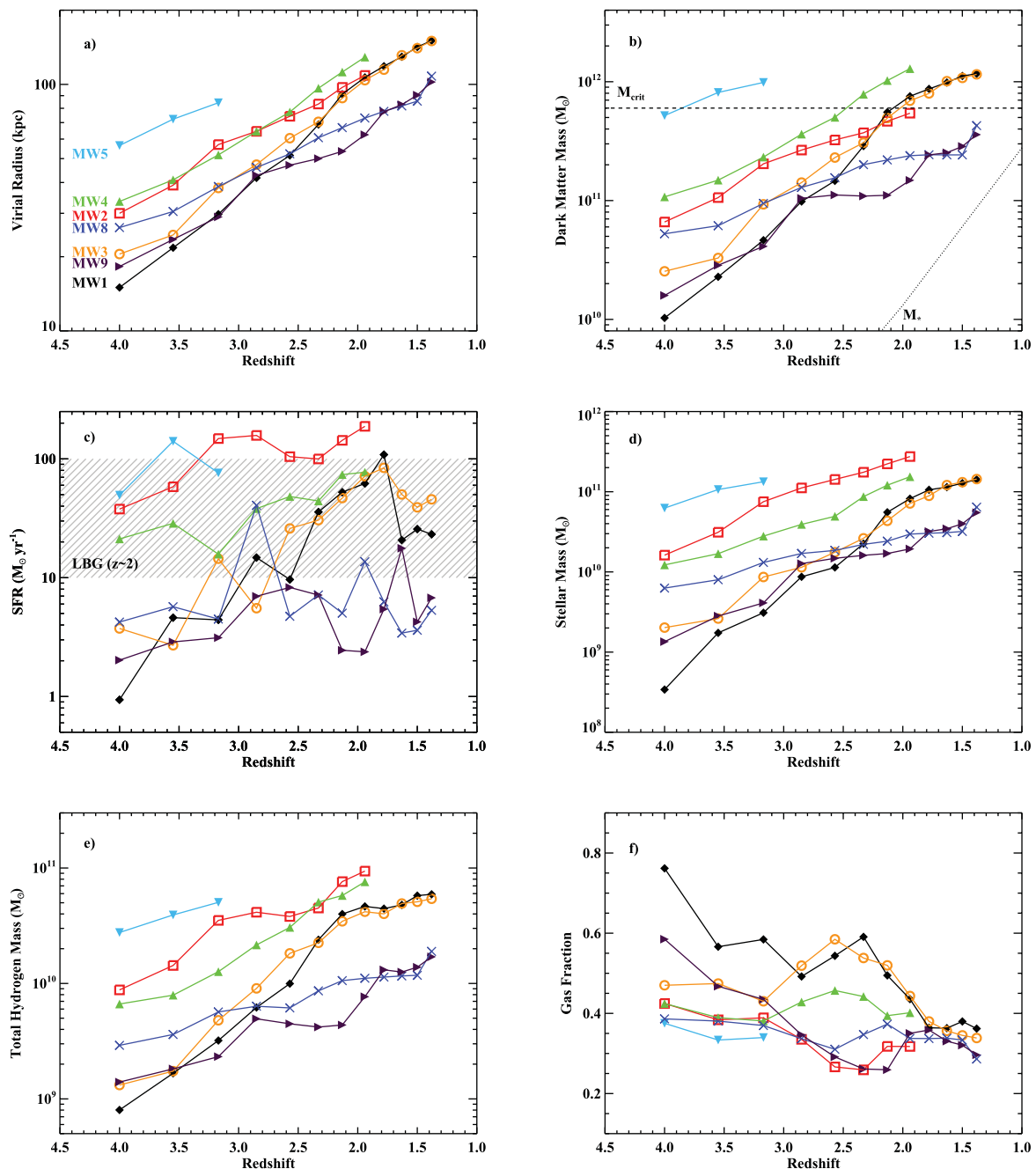
Figure 1. The evolution of the galaxy properties within the virial radius. Each galaxy is identified by a different line connecting different symbols. (a) Virial radii. (b) Dark matter mass, where the dotted line marks the characteristic halo mass, and the horizontal dashed line is the critical mass above which a stable shock can develop (Dekel & Birnboim 2006). (c) SFR, where the shaded area represents the typical SFR of LBGs at $z \sim 2$ (Reddy et al. 2010) to highlight the general overlap. (d) Stellar masses. (e) Hydrogen masses. (f) Gas fraction, defined as the ratio of gas mass to total baryon mass.

within the central discs and satellites galaxies. Additional details on the numerical procedures and validating tests can be found in a companion paper (Kasen et al., in preparation). Here, we briefly summarize the models adopted for the source and sink of ionizing radiation.

3.1.1 Collisional ionization

At temperatures above 10^4 K, hydrogen atoms are ionized by electron collisions. Neglecting cosmic rays that can affect the neutral fraction even at lower temperatures, electron collisions are the only

factor that regulates the ionization state of the gas in the absence of external radiation fields or in regions that are optically thick to UV radiation. In our simplest model, we assume collisional ionization at equilibrium (CIE) under the simplifying assumption of a dust-free gas without external radiation (Gnat & Sternberg 2007). Under this approximation, gas below $T = 10^4$ K is fully neutral, while gas above $T > 2 \times 10^4$ becomes highly ionized. In models that include also UV radiation, collisional ionization is coupled to photoionization. Since our RT calculation does not include photoheating, the final temperature and the resulting ionization by collisions are likely to be underestimated in partially shielded regions. This problem is
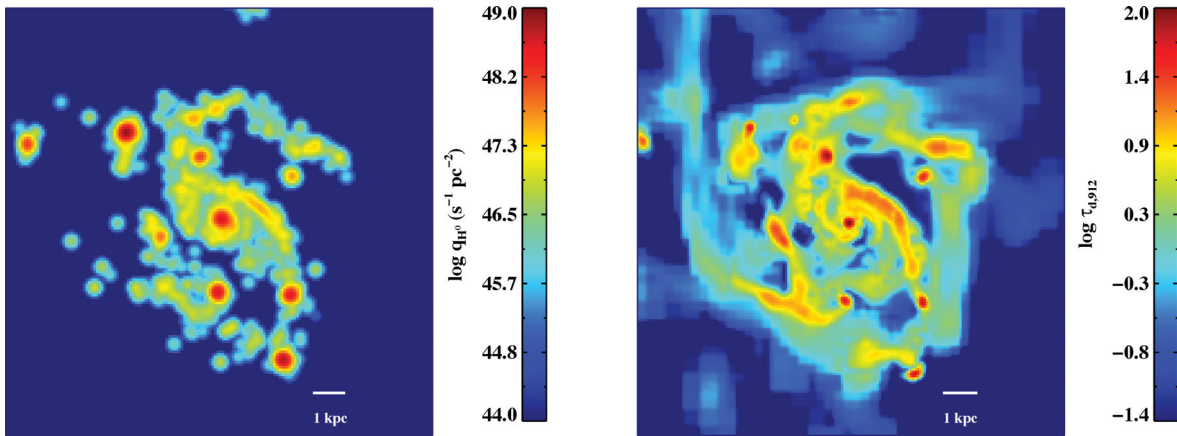
Figure 2. Ionizing radiation and dust in the central disc of MW3 (face on). Left-hand panel: surface density of ionizing radiation smoothed with a Gaussian kernel of 240 pc, showing giant clumps with enhanced star formation. Right panel: integrated optical depth of dust at 912 Å along the entire depth of the box (282 kpc). Star formation occurs primarily in overdense and enriched regions where the dust optical depth is high.

alleviated by the fact that the hydrodynamic code already includes a treatment for photoheating.[2]

3.1.2 UV radiation from cosmological background

The UV radiation from unresolved background sources is the next ingredient we add to the RT calculation. Since the UVB spectrum does not vary much with frequency over the relevant interval for hydrogen ionization, we assume a constant mean intensity $J_{912} = 5 \times 10^{-22}\,\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}}$ (Haardt & Madau 1996) over the energy range 13.6–54.4 eV, corresponding to the H I ionization threshold and the cut-off at the He II threshold, respectively. Further, the modest variation as a function of redshift in the interval $z \sim 2$–4, within the uncertainties of the UVB models (e.g. Faucher-Giguère et al. 2008; Dall'Aglio, Wisotzki & Worseck 2009), justifies the use of a fixed value with time. Our analysis is sensitive to the specific assumptions of the UVB, but it will become evident from the discussion that a decrease in J_{912} from $z = 3$ to $z = 4$ would not significantly affect (and even reinforce) our results.

3.1.3 UV radiation from local sources

An additional and important source of local ionizing radiation is the Lyman continuum from newly born stars. In our model, we estimate the rate of ionizing photons (Q_{H^0}) emitted by a young star using stellar population synthesis models. Q_{H^0} is a function of the metallicity, IMF, SFH and age of the stellar populations. Among those, the most dramatic dependence is on the SFH and age, since the ionizing radiation from a single burst can vary up to several orders of magnitude in a short time interval (1–100 Myr), when the most massive stars that contribute to the Lyman continuum luminosity leave the main sequence.

As a first approximation, we estimate the total number of ionizing photons from a constant mean SFR from the entire simulated box (i.e. from both the central galaxy and satellites) using the proportionality $Q_{\mathrm{H}^0}(\mathrm{s^{-1}}) = \alpha \mathrm{SFR}(\mathrm{M_\odot\,yr^{-1}})$. We compute the normalization $\log \alpha = 53.168$ using STARBURTS99 (Leitherer et al. 1999), with a Kroupa IMF (Kroupa 2001) above $0.08\,\mathrm{M_\odot}$, subsolar metallicity $Z = 0.004$, and stellar models from Lejeune & Schaerer (2001).

Stellar particles have different degrees of enrichment and the mean metallicity ranges from $Z = 0.004$ to $Z = 0.013$ in our sample, but the variation of Q_{H^0} with metallicity is at most a factor of 2 across the range $Z = 0.001$–0.04 ($\log \alpha = 53.227 - 52.935$), and even less ($\sim 0.1$ dex) with the choice of the available stellar models. We redistribute the total Q_{H^0} among stellar particles which are younger than 20 Myr, in a fraction that is proportional to the stellar mass $M_{i,\mathrm{star}}$

$$Q_{i,\mathrm{H}^0} = Q_{\mathrm{H}^0} \frac{M_{i,\mathrm{star}}(t < 20\,\mathrm{Myr})}{\sum_i M_{i,\mathrm{star}}(t < 20\,\mathrm{Myr})}. \tag{1}$$

Beside the intrinsic uncertainties in the stellar models and metallicity, other reasons make the photoionization rate from stars difficult to estimate. For example, our finite resolution affects the ability to properly resolve the clumpy ISM at scales below 35–70 pc. This enhances the leakage of Lyman continuum photons due to the absence of high-density structures around stellar particles, but at the same time the escape fraction of ionizing photons is decreased by the lack of porosity (Wood et al. 2005). A proper estimate of the error associated with the leakage of UV photons is difficult, but most of the ionizing radiation from stellar sources is absorbed in the proximity of the emitting regions at high hydrogen volume density. In these models, the escape fraction at the virial radius is between 5 and 10 per cent, varying with redshift and from galaxy to galaxy. As a consequence, only a small fraction of the escaping radiation ionizes gas near the central discs with $N_{\mathrm{H\,I}} \sim 10^{18}$–$10^{20}$ cm^{-2} (see the next section). Therefore, the associated uncertainties on the neutral gas and the cumulative cross-section of optically thick hydrogen beyond R_{vir} with $N_{\mathrm{H\,I}} \gtrsim 10^{17}$ cm^{-2} are expected to be minor.

The left-hand panel of Fig. 2 shows the ionizing radiation rate surface density (convolved with a Gaussian kernel of 240 pc only for visualization purposes), for the central disc of MW3 at $z = 2.3$ and in a nearly face-on view. Most of the ionizing radiation is produced in clumps and in spiral density wakes. While the SFR in the entire box approaches 50 $\mathrm{M_\odot\,yr^{-1}}$, the central disc accounts for only half of the total ionizing radiation (with SFR ~ 27 $\mathrm{M_\odot\,yr^{-1}}$). About 50 per cent of the remaining photons come from a large satellite which is forming stars at ~ 10 $\mathrm{M_\odot\,yr^{-1}}$. This example highlights the fact that photons from young stars that escape the surrounding haloes can introduce low-level anisotropies in the UV radiation field.

---

[2] Further discussion on this issue can be found in Appendix A.

3.1.4 Dust opacity

Dust is an important sink of Lyman continuum photons, particularly relevant for radiation from local sources. Using the metallicity in the gas phase from SN II and SN Ia, the dust volume density is given by

$$\rho_d = f_d \mu m_p (Z_{\mathrm{SN\,Ia}} + Z_{\mathrm{SN\,II}})(n_{\mathrm{H\,I}} + 0.01 n_{\mathrm{H\,II}}), \qquad (2)$$

where $f_d = 0.4$ is an estimate of the fraction of metals locked in dust (Dwek 1998) and $\mu = 1.245$ is the mean particle weight, including helium. To mimic grain destruction, we suppress dust formation in ionized regions to only 1 per cent of what is found in the neutral phase. This fraction is the most uncertain quantity in the dust model (see Laursen, Sommer-Larsen & Andersen 2009, section 7).

From the dust density we compute the dust optical depth at 912 Å, i.e. the hydrogen ionization potential, in each AMR cell as $\tau = \alpha l$, with $\alpha = n_d(\sigma_s + \sigma_a) = \kappa \rho_d (1 - A)^{-1}$ and l the cell size. Here, σ_s and σ_a are the scattering and absorption cross-sections, A is the albedo and κ is the frequency-dependent dust absorptive opacity. For the above quantities, we assume $\kappa = 9.37 \times 10^4 \mathrm{\ cm^2\,g^{-1}}$ and $A = \sigma_s/\sigma_d = 0.248$ (Draine 2003). During the RT calculations, at the relevant UV wavelengths, a linear fit to the Li & Draine (2001) data is used: $\kappa(\lambda) = 9.25 \times 10^4 - 91.25 \times (912\,\text{Å} - \lambda)$ and $A(\lambda) = 0.24 + 0.00028(912\,\text{Å} - \lambda)$. A map of the dust optical depth for MW3 at $z = 2.3$ in a nearly face-on view is in the right-hand panel

of Fig. 2. Here the projected dust optical depth is computed along a path of 282 kpc, the size of the entire box.

3.2 Results of the radiative transfer calculation

For each galaxy, we run three different RT models, gradually including additional physical processes. In the first model (hereafter CIE model), we derive the neutral fraction assuming CIE, without any source of radiation. In the second calculation (UVB model), we include the UVB together with dust and collisional ionization. Finally, in our third model (STAR model) we add ionizing radiation from local sources to the UVB model. Fig. 3 presents an output from these calculations. In the top-left panel, we show the projected N_H column density for MW3 at $z = 2.3$. High column density gas is accreting on to the central galaxy through large radial streams, with gas overdensities associated with satellites (two incoming galaxies along the streams and two closer in, near the central disc). In the other three panels, we display the $N_{\mathrm{H\,I}}$ column density from the different RT models.

This figure captures the basic differences that arise from the different physical processes included in the RT calculation. Part of the gas within the filaments has a temperature above $\sim 10^4$ K, and collisional ionization alone (top-right panel) lowers the neutral column density by more than one order of magnitude. A comparison of the H I map for the CIE and UVB models (bottom left) clearly shows that the CIE approximation largely overestimates the neutral
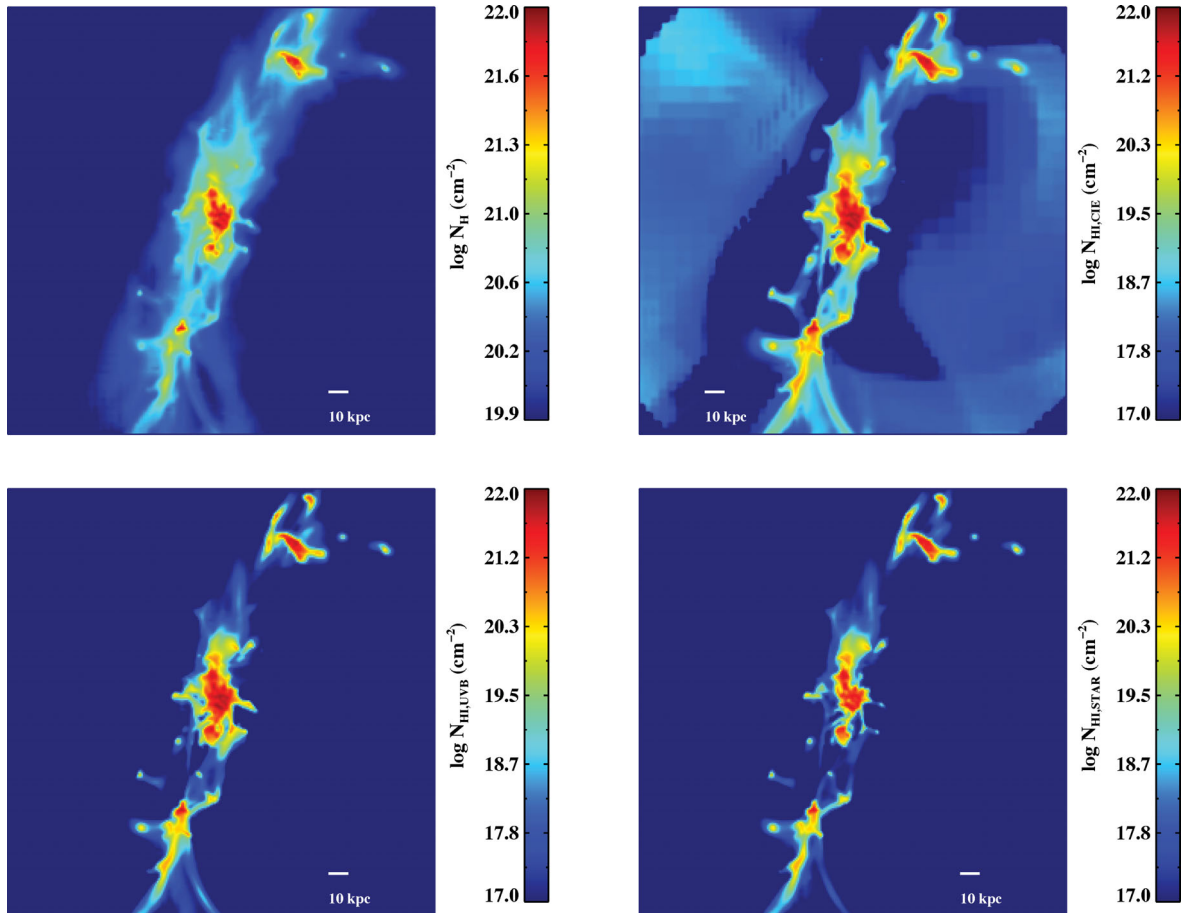

Figure 3. Hydrogen column density for MW3 at $z = 2.3$. Top left: N_H. Top right: $N_{\mathrm{H\,I}}$ from the CIE model. Bottom left: $N_{\mathrm{H\,I}}$ from the UVB model. Bottom right: $N_{\mathrm{H\,I}}$ from the STAR model. Most of the gas that resides in the streams is ionized by electron collisions and the UVB, while photons from newly born stars affect the high column density inside the central and satellite galaxies and their immediate surroundings.

fraction and that photoionization from the UVB affects most of the low-density gas in the streams. Indeed, the filaments are highly ionized, with patches of self-shielded neutral gas that surround the main galaxy and the satellites. Cold streams are not entirely self-shielded. Finally, the inclusion of local sources mostly affects the high column density gas (where stars form) and their immediate surroundings, where the column density is caused to drop below $N_{\rm H_I} \sim 10^{20}$ cm^{-2}. The low escape fraction from the galaxy discs (below 10 per cent at the virial radius) implies a minor effect on the gas in the streams beyond $R_{\rm vir}$ without any appreciable difference compared to the UVB model for column densities below $N_{\rm H_I} \sim 10^{18}$ cm^{-2}.

A more quantitative comparison between the different models is presented in Appendix A. There, we discuss the typical volume density for self-shielding, and the effect of local sources on the column density and mass of neutral hydrogen. We also provide a crude fitting formula to the UVB model useful to improve the CIE approximation.

4 THE PROJECTED AREA OF NEUTRAL GAS

In the previous section, we have shown that the streams are only partially shielded and that both the UVB and local sources are necessary ingredients to study the neutral and ionized gas distribution across a large range of column density. In this section, we quantify the cross-section of neutral gas in and around massive galaxies and the evolution of the covering factor. In the remaining of our analysis, we consider the STAR model as our fiducial RT calculation, providing from time-to-time a comparison with results from the UVB model.

4.1 Neutral hydrogen cross-section

We derive the cross-section that these simulations offers in projection as a function of the H I column density by collapsing the AMR 3D grid on to a 2D regular grid with cell size equal to the size of the finest level in the AMR structure. We measure in each galaxy the covered area in $N_{\rm H_I}$ bins of 0.2 dex, averaged in three orthogonal directions along the axes of the AMR grid. An example for MW3 at $z = 2.3$ (STAR model[3]) is given in Fig. 4. For relative comparisons, we introduce in the right-hand side axis the covering factor, defined as the cross-section normalized to the area inside a circle of radius $2R_{\rm vir}$. The solid line represents the cross-section measured in the entire box, while the dot–dashed line highlights the contribution of the cold streams alone, obtained by masking all the gas cells inside one quarter of the virial radius of the main haloes and satellites. Further, to isolate the contribution of the central galaxy, we display the cross-section within two concentric cylinders of sizes $0.2R_{\rm vir}$ (dotted line) and $R_{\rm vir}$ (dashed lines). In the latter case, we display the minimum and maximum cross-section among the three projection axis with a hatched region. Obviously, different orientations have different cross-section distributions, but the dispersion along the mean appears not to exceed the amplitude of the features that are visible in the distribution. The radius $0.25R_{\rm vir}$ is well suited to separate streams and galaxies and it encompasses the interphase region where the flowing gas is reconnecting with the central disc. However, the fraction of neutral gas in the column density interval
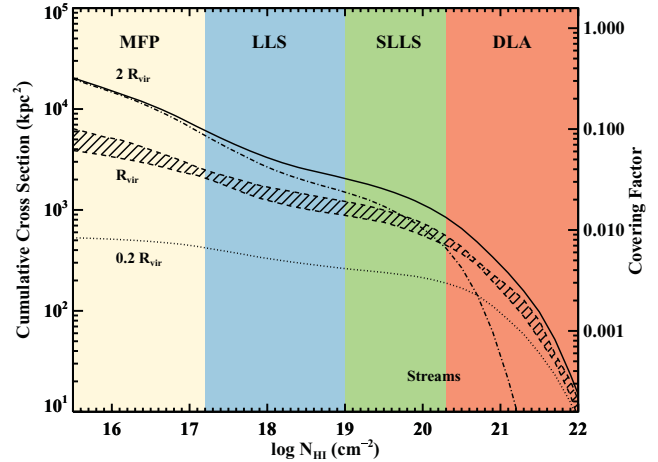
---

³ We highlight how the three RT calculations affect the cross-section at various column densities in Fig. A1.

Figure 4. Cumulative cross-section for MW3 at $z \sim 2.3$ (STAR model) averaged along three orthogonal directions. The different curves are for a cylinder of radius $2R_{\rm vir}$ (solid line), a cylinder of radius $0.2R_{\rm vir}$ (dotted line) and for the streams only within $2R_{\rm vir}$ (dash–dotted line). The hatched region between dashed lines highlights the dispersion along the three projection axis for a cylinder of radius $1R_{\rm vir}$. On the right axis, the physical area is translated into a fractional covering factor, normalized to the area within $2R_{\rm vir}$. Colours represent intervals of column density for DLAs, SLLSs, LLSs and MFP gas. Gas within galaxies contributes to half of the DLA cross-section, while LLSs and MFP gas are mainly associated with cold streams. Only $\sim$10 per cent of the projected area is covered by optically thick gas and $\sim$1 per cent of the cross-section is covered by primarily neutral gas at $N_{\rm H_I} \gtrsim 10^{20}$ cm^{-2}.

10^{19}–10^{20} cm^{-2} that is associated with the streams depends on the location of this boundary.

We mark in this figure four $N_{\rm H_I}$ intervals that define the ionization state of the atomic hydrogen and that are associated with different classifications of observed ALSs. This is done to highlight at what intervals of column densities cold streams are predicted to dominate the neutral gas cross-section. Specifically, at $N_{\rm H_I} \geq 2 \times 10^{20}$ cm^{-2}, the hydrogen is neutral, highly optically thick ($\tau \gtrsim 1000$), and gives rise to damped Lyman α absorbers (DLAs). The interval 10^{19} cm$^{-2} \leq N_{\rm H_I} < 2 \times 10^{20}$ cm^{-2} marks the transition between fully neutral to ionized hydrogen and defines the super-Lyman-limit systems (SLLSs) or sub-damped Lyman α absorbers. The interval 1.6×10^{17} cm$^{-2} \leq N_{\rm H_I} < 10^{19}$ cm^{-2} defines Lyman-limit systems (LLSs) which are highly ionized, but retain enough neutral hydrogen to be optically thick ($1 \lesssim \tau \lesssim 60$). Finally, below $N_{\rm H_I} = 1.6 \times 10^{17}$ cm^{-2}, the hydrogen becomes optically thin ($\tau \lesssim 1$) and it is highly ionized. However, this gas is estimated to contribute to the mean-free path (MFP) of Lyman photons in the Universe at $z > 3$ (Prochaska, Worseck & O'Meara 2009; Prochaska, O'Meara & Worseck 2010). In this work, we consider the interval 3.2×10^{15} cm$^{-2} \leq N_{\rm H_I} < 1.6 \times 10^{17}$ cm^{-2}, which we dub as 'MFP' gas.

The cross-section distribution within $R_{\rm vir}$ steeply rises from the high end towards $N_{\rm H_I} \sim 10^{20}$ cm^{-2} and then flattens between $N_{\rm H_I} \sim 10^{18}$ and 10^{20} cm^{-2} where gas is ionized. At lower column density there is a mild increase. Gas within $2R_{\rm vir}$ exhibits a similar shape in its cross-section, but the separation between the two distributions increases systematically moving towards low column densities. From the example provided in Fig. 4 we find that at $2R_{\rm vir}$, $\sim$10 per cent of the projected area is covered by optically thick gas and $\sim$1 per cent is covered by predominantly neutral gas. Note that despite the lower covering factor, this gas may still contribute to the integrated opacity of the ionizing radiation (see below). About 30 per cent of the area is occupied by gas with $N_{\rm H_I} \gtrsim$

10^{15} cm^{-2}. Most of the MFP gas and of the LLSs are associated with the streams alone, and half of the area that is covered by SLLSs resides outside the virial halo of the massive central galaxy, with streams accounting for two third of the covering factor. A significant fraction of the neutral gas is located within haloes with half of the DLA cross-section inside the streams. Above $N_{\rm H\,I} \sim 10^{21}$ cm^{-2}, the fraction of DLAs in the streams drops below 20 per cent and becomes negligible at even higher column densities. Only 20 per cent of the DLA cross-section is found within $0.2R_{\rm vir}$, where it is likely associated with the central disc, indicating that a significant fraction of DLAs can be found within clumps and satellites inside the virial radius of the main halo, but outside the central disc.

That a non-negligible fraction of the DLA cross-section is found within the virial radius but outside the central discs appears to be a typical property of these simulations, as shown in Fig. 5. In each snapshot (STAR model), we identify an independent structure as an ensemble of contiguous cells above $N_{\rm H\,I} = 2 \times 10^{20}$ cm^{-2}. We define the central galaxy to be the closest structure to the centre of the main halo and we consider satellites or clumps in the streams to be all of the remaining groups of contiguous cells. The average cross-section as a function of the halo mass is shown with blue squares for the central galaxies and black circles for the central galaxies and satellites. The amplitude of the halo mass intervals, chosen to have an equal number of objects, is indicated by the horizontal error bar, while the vertical error bar is for the standard deviation along the mean. Values are listed in Table B6.

In the entire sample, the central galaxy contributes $\sim$35–65 per cent of the DLA cross-section, with galaxies residing in more massive haloes showing smaller fractions. We conclude that in these simulations the satellites of massive haloes and clumps that reside in the streams are at least as important to the DLA cross-section of massive dark matter haloes as the central galaxy (see also Maller et al. 2001; Razoumov et al. 2006). Note that the number of objects
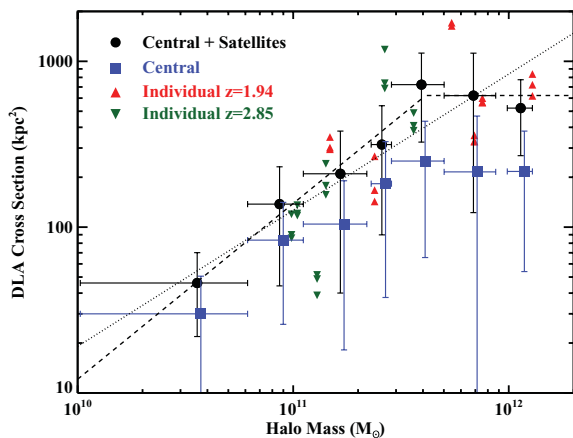
included in this study is limited and the cross-sections listed here can be affected by sample variance. Also, the cross-section of DLAs is particularly sensitive to the disc sizes that are notoriously difficult to correctly reproduce in simulations, although our simulations show extended discs with 3–5 kpc radii at high redshift (Ceverino et al. 2010). Further, these simulations have been selected to study cold gas accretion in massive galaxies and are not representative of the entire population of galaxies giving rise to ALSs. Nevertheless, it is reassuring that the mean DLA cross-section within $R_{\rm vir}$ ranges between 50 and 800 kpc^2 for haloes of 10^{10}–10^{11} M$_\odot$, in good agreement with the cross-sections found at similar masses in the larger cosmological simulation by Pontzen et al. (2008) and Tescari et al. (2009). Our predictions are towards the lower end of the size distribution from Cen (2010), who finds a larger area as a consequence of strong feedback. This difference emphasizes how various feedback prescriptions affect the final cross-sections.

The total area is nearly proportional to the halo mass below 4×10^{11} M$_\odot$ (slope $\gamma = 1.06$ and normalization $\beta = 3.16 \times 10^{10}$ for a log–log linear regression; dashed line in Fig. 5) and roughly constant at higher masses. Across the whole available mass range, an approximate fit to the points in Fig. 5 is given by the mass dependence of the area within the virial radius at a fixed time: $A_{\rm vir} \propto R_{\rm vir}^2 \propto M_{\rm vir}^{2/3}$. This is consistent with DLAs originating from regions where the surface density is above a fixed threshold, provided that the distribution of densities within the virial radius (or $2R_{\rm vir}$) is self-similar among the haloes.

Note that the relation found in Fig. 5 is based on only seven independent haloes at all times and redshift is partially degenerate with the halo mass, since the most massive haloes are found only at late times. It is therefore useful to inspect values for single galaxies, superimposed in this figure at $z \sim 2.8$ and $z \sim 1.9$ in three projections. While a mild evolution with redshift is visible (see also the next section), the mean cross-sections previously derived appear to be a fair representation of this sample as a whole, more so considering the large intrinsic scatter from galaxy to galaxy.

4.2 Redshift evolution of the covering factor

A qualitative description of the covering factor evolution within $2R_{\rm vir}$ is provided in Fig. 6, where we show a gallery of projected H I column densities (STAR model) for a selected number of galaxies. Redshift decreases from the top to the bottom and mass at fixed redshift increases from the left to the right. DLAs are marked in red, SLLSs in green, LLSs in blue and MFP gas is in orange. In each panel, the galaxy centre and virial radius are indicated by a dotted circle. Gas that surrounds the galaxies is non-uniform and the cold streams are 'patchy' with neutral regions that are embedded in a widespread ionized medium.

An evolution in redshift appears from this figure. The gas coverage is larger at early times than at late times. Evolution in the gas concentration is evident for DLAs. At highest redshifts, there are large regions of neutral gas within the streams. These become more confined to the central discs and satellites at later times. A trend in mass is also visible, although less pronounced, with larger gas coverage in lower mass haloes. Also, the topology of the streams seems to evolve with time and mass. In low-mass systems and at early times, wide streams penetrate inside the virial radius to deliver cold gas to the forming galaxies. At lower redshifts and in more massive systems, the streams break into a series of narrow filaments (see Kereš & Hernquist 2009; Stewart et al. 2010; Faucher-Giguère & Kereš 2011). In our simulations, numerical resolution could be responsible for the increased smoothness in the smallest galaxies

Figure 5. Average DLA cross-section within $R_{\rm vir}$ as a function of halo mass for the central galaxies only (blue squares) and central galaxies plus satellites (black circles). All the galaxies in our sample, at all sampled redshifts, are included. The vertical error bar indicates the standard deviation about the mean, while the horizontal error bar indicates the interval of halo mass used in the average. Also shown are individual galaxies at $z \sim 2.8$ (green downward triangle) and $z \sim 1.9$ (red upward triangle). A linear regression (dashed line) shows that the total cross-section is nearly proportional to the halo mass below 4×10^{11} M$_\odot$, while it is roughly constant at higher masses. A shallower dependence is found over the entire available mass range (dotted line), consistent with DLAs originating above a fixed surface density threshold for a self-similar gas density distribution. Central galaxies typically account for less than 50 per cent of the DLA mean cross-section.
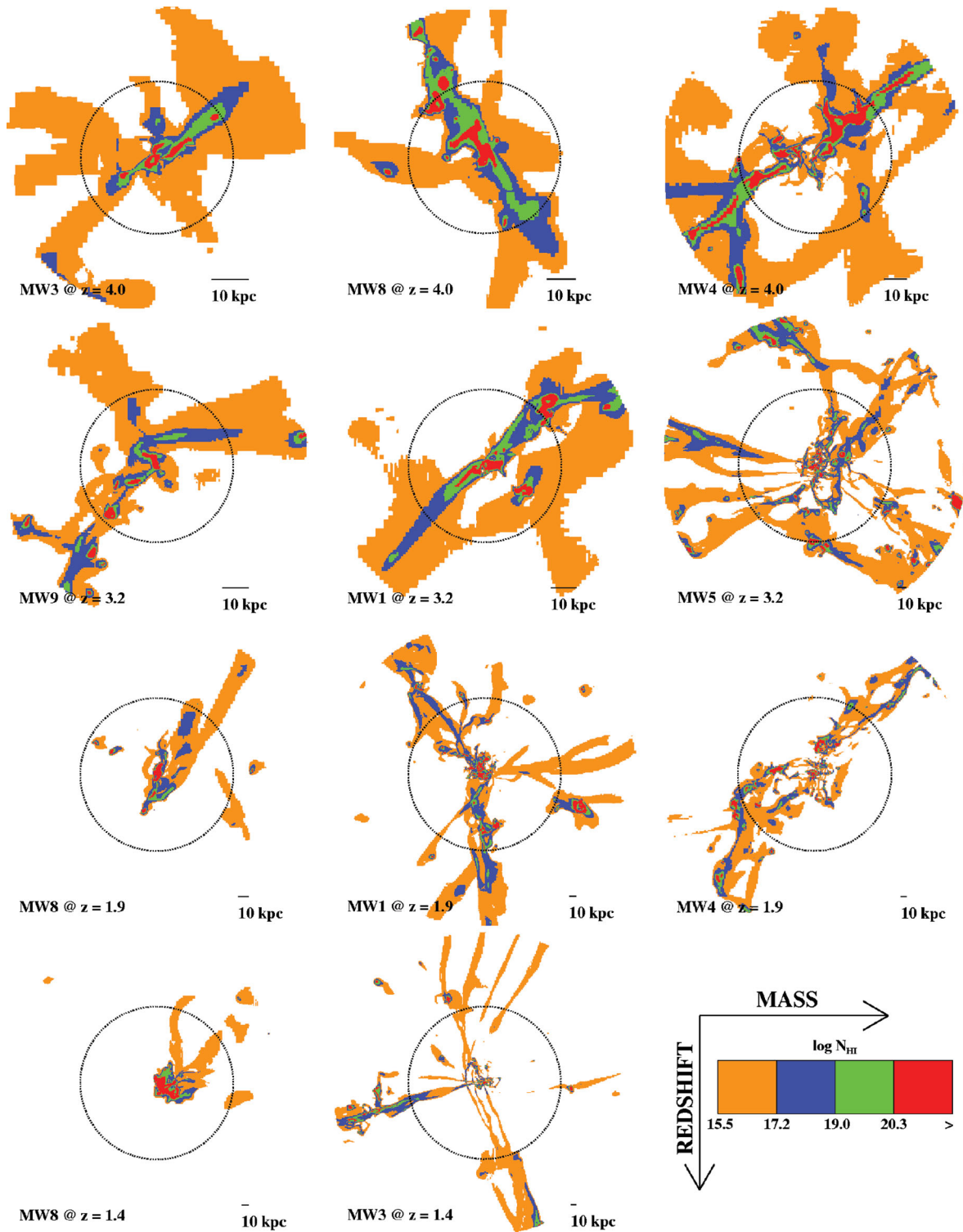
Figure 6. A gallery of projected H I column density (STAR model) in selected galaxies and redshifts. Four intervals of column density are marked with different colours (DLAs in red, SLLSs in green, LLSs in blue and MFP gas in orange, similar to Fig. 4). Redshift is decreasing from top to bottom, and virial mass is increasing at constant redshift from left to right. The dotted circles mark the virial radius. Cold streams are 'patchy', with neutral pockets of gas embedded in a more widespread ionized medium. A difference is seen between $z > 3$ and $z < 2$, and a marginal mass dependence is hinted, but the latter could be an artefact of numerical resolution.

at the highest redshifts and higher resolution is required to confirm whether this trend is real.

A more quantitative analysis of the redshift dependence of the covering factor is given in Fig. 7. Each bar represents the mean values for DLAs (red), SLLSs (green), LLSs (blue) and MFP gas (orange) at one redshift and within $2R_{vir}$. The central column is for the galaxies and streams in the STAR model, the left-hand column is for the streams alone in the STAR model and the
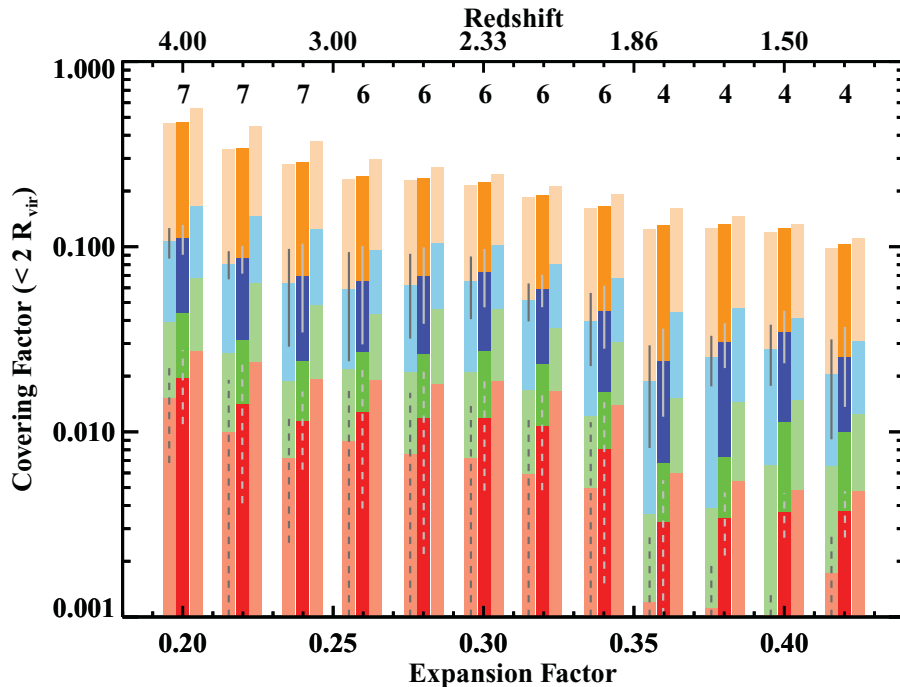
Figure 7. Time evolution of the cumulative covering factor within $2R_{\rm vir}$. DLAs are in red, SLLSs in green, LLSs in blue and MFP gas in orange. The central columns are for the galaxies and streams in the STAR model, the left-hand columns are for the streams alone in the STAR model and the right-hand columns are for the UVB model, galaxies and streams. The numbers above the bars indicate the number of galaxies in each bin of redshift. Grey dashed and solid lines indicate the standard deviation for the mean covering factor of DLAs and LLSs, respectively. The covering factor slowly decreases with redshift at all column densities, reflecting the decline of cosmological density.

right-hand column is for the UVB model (streams and galaxies). Values for DLAs and LLSs in individual galaxies are listed in Table B4. All covering factors are given as cumulative, i.e. integrated from the lowest column density to infinity (numerically approximated to $N_{\rm H\,I} = 10^{25}$ cm^{-2}). The numbers at the top of the bars indicate the number of galaxies included in each redshift bin. The standard deviation of the sample is shown for DLAs and LLSs with grey dashed or solid lines to highlight the wide scatter from galaxy to galaxy, partially associated with the mass (see Fig. 9).

Fig. 7 confirms the impression that the covering factor within $2R_{\rm vir}$ is slowly declining with time. At the highest redshifts, 30–50 per cent of the area is covered by MFP gas, 7–10 per cent is covered by LLSs and 1–2 per cent is occupied by DLAs (compare with Faucher-Giguère & Kereš 2011). By $z \sim 2$, the covering factor becomes 1 per cent for DLAs, 5 per cent for LLSs and $\sim$20 per cent for MFP gas. In the range z=2–1.3, the covering factors are $\sim$10 per cent, 3–4 per cent and 0.3–0.4 per cent, respectively. Note that the already limited statistics are dominated by lower mass galaxies below $z \sim 1.9$, while the highest redshift bins may be more prone to resolution effects.

In the redshift interval $2 < z < 3$, where the sample is more homogeneous, the time evolution of the covering factor can be parametrized by $\sim a^{-d}$, with a the expansion factor. This trend can be explained by the universal expansion for a self-similar density distribution of neutral hydrogen. In these simulations, the cumulative H I column density probability distribution function (PDF) can be described by $\propto N_{\rm H\,I}^{-c}$, with $c \sim 0.2$ for $N_{\rm H\,I} \lesssim 10^{20}$ cm^{-2} and $c \sim 1$ at higher column densities. In our sample, the viral mass and the virial radius grow with time as $\sim a^3$ and $\sim a^2$, and the virial column density decreases as $\sim a^{-1}$. For a self-similar density distribution of neutral hydrogen, it follows that $c \sim d$, in agreement with the evolution in Fig. 7.

Comparing next the UVB model (right bars) with the STAR model, a systematic offset towards higher values is apparent from the former. The inclusion of local sources does not drastically alter the covering factor (a factor of 2 for DLAs and LLSs), but it is clear that ionizing radiation from stars has appreciable effects especially at intermediate column densities and it is required for a consistent comparison with observations (Schaye 2006). The redshift evolution is still present in the UVB model, reassuring us that it is not driven by local sources. Furthermore, this trend cannot be artificially induced by our choice of a constant UVB, since the intensity of the UVB decreases towards the highest redshifts considered here.

In Fig. 7, we finally compare the covering factor from the streams and galaxies with the streams alone. Overall, the covering factor of highly optically thick gas in the streams is decreasing faster with time than that of galaxies. For the MFP gas, this effect is absent, with streams responsible for the totality of the covering factor at all times. Streams contribute to more than 80 per cent of the LLSs covering factor over the full redshift interval. Instead, a rapid evolution is seen for SLLSs and DLAs. Streams are responsible for as much as 70 per cent of the covering factor at the highest redshifts, but less than 30 per cent of the covering factor at later times as neutral gas accumulates in proximity of the central galaxies an satellites.

Being a normalized quantity, the covering factor within $2R_{\rm vir}$ is suitable for comparisons among different galaxies and redshifts. However, it does not provide direct information on the physical cross-section and its evolution. Further, observations probe either the gas cross-section or the covering factor within a given projected distance from a source. In the latter case, the most natural quantity to compute is the covering factor within a fixed angular distance from the halo centre. In Fig. 8, we show the mean covering factor within $R_{\rm vir}$ (right columns), within $2R_{\rm vir}$ (left columns) and within a circular aperture of 10 arcsec in radius (central column). Again, the
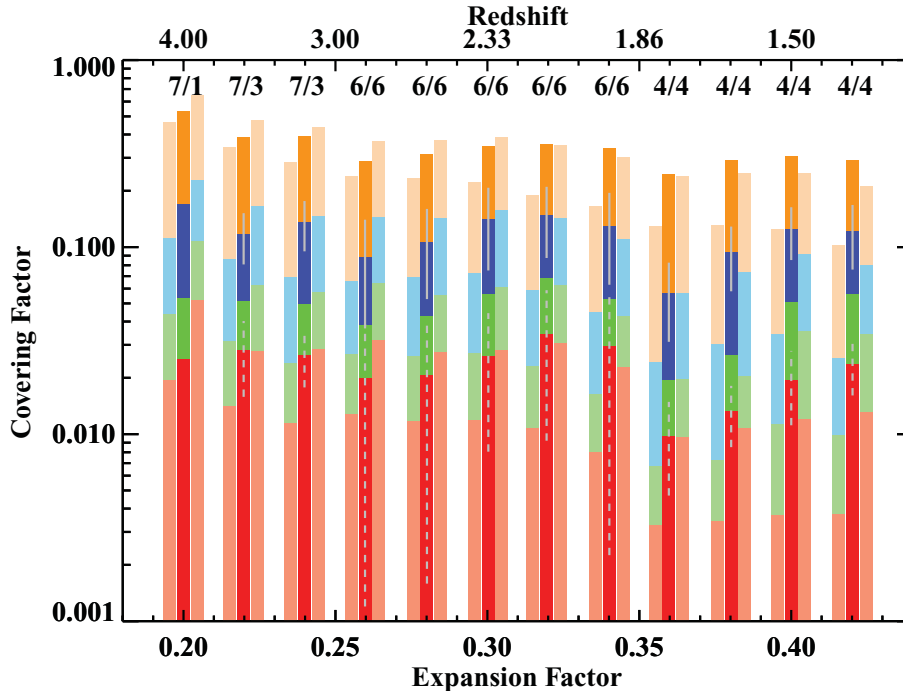
Figure 8. Same as Fig. 7, comparing covering factors that are measured in different areas. In the centre, the covering factor is measured within a fixed aperture of 10 arcsec in radius. Shown for comparisons are the covering factors defined within R_{vir} and $2R_{vir}$, on the right and left, respectively. The number of galaxies included in each bin is shown above the columns (variable/fixed apertures) and the standard deviation around the mean is highlighted with solid and dashed lines. The variation with time is much weaker when the aperture is fixed, reflecting the decrease of covering area with distance from the halo centre.

number at the top indicates the number of galaxies included in each bin with the second number for the fixed aperture. The variance about the mean for LLSs and DLAs is shown with a solid and dashed line, respectively. At the redshifts examined here, an angular distance of 10 arcsec corresponds to 70–85 proper kpc, equivalent to approximately three to five times the virial radius at the highest redshifts and to ∼0.5–1 times the virial radius at the lowest redshifts. The snapshots in which the extraction box exceeds the simulated region are not included. Further, for a consistent comparison, we consider a path-length through the grid of twice the virial radius in all three cases.

The covering factor measured within R_{vir} is higher than the one inside $2R_{vir}$, particularly for DLAs. This is trivially expected for a gas distribution with declining column density. Considering now a fixed angular distance, the covering factor exhibits a much weaker evolution with a modest increase in the interval $2 < z < 3$, where the sample is more homogeneous. As the universe expands, the mean column density within $2R_{vir}$ drops and the fraction of the area within $2R_{vir}$ subtended by gas above a fixed column density threshold decreases accordingly. However, the virial radius is increasing with respect to the size of the extraction aperture, resulting in a nearly constant or slightly increasing covering factor.

Within $\sim 100\,h^{-1}$ kpc, the covering factor of absorbers in the transverse direction of quasi-stellar object (QSO) host galaxies is significantly higher than our estimates, being unity for $N_{H\,I} > 1.6 \times 10^{17}$ cm^{-2} (Hennawi et al. 2006) and 30 per cent for $N_{H\,I} > 10^{19}$ cm^{-2} (Hennawi & Prochaska 2007). This discrepancy may originate in the environment around QSOs, more massive than the galaxies in our sample. Further, these simulations do not resolve overdensities on very small scales and consequently the cross-section distribution suffers from an additional uncertainty. For example, a population of neutral clouds of $\lesssim 100$ pc in size that

are pressure confined by the hot halo gas (see e.g. Dekel & Birnboim 2008; Prochaska & Hennawi 2009; Birnboim & Dekel 2011) may be missed in these simulations and the cross-section of optically thick gas would be underestimated. Similarly, shells or bubbles associated with radiative shocks on small scales may not be properly resolved (cf. Simcoe et al. 2006).

The weak redshift evolution seen in Figs 7 and 8 may be the result of having combined galaxies at different mass. To explicitly test this hypothesis, in Fig. 9 we display the LLS covering factor within $2R_{vir}$ (bottom panel) and within 10 arcsec (top panel) for individual galaxies. Although with significant scatter from object to object, the covering factor within $2R_{vir}$ appears to be roughly constant or slowly declining regardless to the galaxy mass. Similarly, the covering factor within a 10 arcsec aperture is slowly increasing at all masses.

5 PREDICTED ABSORPTION-LINE STATISTICS

In the previous section we discussed the redshift evolution of the covering factor in streams and galaxies, where we found that only a small fraction of the virial area is covered by optically thick gas. In this section, we wish to translate the cross-sections and covering factors previously measured into more direct observables that can be used in comparison to real data to uncover signatures of cold gas accretion. We start by introducing some of the formalism adopted in typical measurements of ALSs. Readers familiar with this notation can skip the next section and continue from Section 5.2, where we derive the column density distribution function for cold-stream fed galaxies and discuss their visibility in absorption.
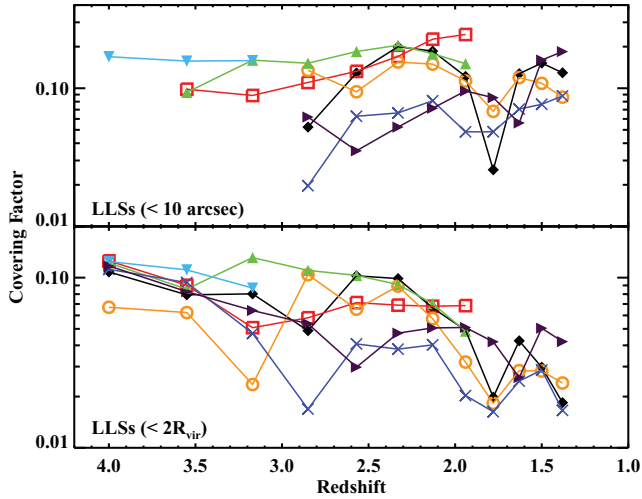
Figure 9. Redshift evolution of the covering factor for LLSs in individual galaxies (colour coded as in Fig. 1) within $2R_{vir}$ (bottom panel) and for a fixed aperture of 10 arcsec (top panel). Trends similar to those inferred from Figs 7 and 8 are visible when inspecting individual galaxies, regardless of their mass.

5.1 Common formalism

To probe any form of gas that lies along the line of sight, observers collect large spectroscopic samples of QSOs that serve as randomly distributed background sources (e.g. Prochaska, Herbert-Fort & Wolfe 2005; O'Meara et al. 2007). In this experiment, there is no a priori knowledge of the foreground system (see however Rafelski, Wolfe & Chen 2010) and one can directly measure the probability to intersect gas along a random sightline. This quantity translates to a sky covering factor that is the percentage of total observed area covered by gas.

One can define the incidence

$$\ell(z) = \frac{N_{abs}}{\Delta z} \qquad (3)$$

as the number of systems N_{abs} detected across the total searched redshift Δz. This quantity is related to the probability of intersecting a system along the line of sight at redshift z. In analogy with the idealized experiment of a body moving in a sea of particles with number density n and cross-section σ, the following equality holds

$$\ell(z)\mathrm{d}z = n_{abs}\sigma_{abs}\mathrm{d}l_c \qquad (4)$$

with n_{abs} the comoving number density of absorbers, σ_{abs} their comoving cross-section and $\mathrm{d}l_c$ the infinitesimal comoving path-length.

Within a given cosmology $H(z) = H_0[\Omega_\Lambda + (1+z)^3\Omega_m]^{1/2}$ and $\mathrm{d}l_c = c/H(z)\,\mathrm{d}z$. Replacing z by an 'absorption length' X

$$\mathrm{d}X = \frac{H_0}{H(z)}(1+z)^2\,\mathrm{d}z \qquad (5)$$

and using the identity $\ell(z)\,\mathrm{d}z = \ell(X)\,\mathrm{d}X$, the incidence can be rewritten as the number of systems per absorption length (Bahcall & Peebles 1969)

$$\ell(X) = \frac{c}{H_0}n_{abs}\phi_{abs} \qquad (6)$$

where the cross-section ϕ_{abs} now represents the physical area of the absorber.

Equation (6) is the zeroth moment of the column density distribution function $f(N_{HI}, X)$ that characterizes the number of ALSs

per absorption length as a function of the H I column density. From the definition of $\ell(X)$, it follows that

$$f(N_{HI}, X) = \frac{c}{H_0}n_{abs}(N_{HI}, X)A_{abs}(N_{HI}, X), \qquad (7)$$

with A_{abs} the physical area per unit N_{HI}. Equation (7) reveals that absorption-line statistics directly probe the shape and the evolution of the physical cross-section times the number density of absorbers rather than the covering factor of gas around these systems.

5.2 The column density distribution function

To understand the nature of the gas probed in absorption, it is common practice to reproduce the $f(N_{HI}, X)$ in its shape and normalization from simulations of large volumes, required for statistical significance. Here we take a different approach where we derive an $f(N_{HI}, X)$ that, by construction, is due to gas associated only with massive galaxies and the streams with merging galaxies that feed them. Any similarity or discrepancy with the observed distribution function informs of what features of $f(N_{HI}, X)$ can be related to the gas distribution in these particular class of systems as predicted by these simulations. The reader should keep in mind that our sample and our $f(N_{HI}, X)$ are not intended to be representative of the full ALS population.

Our model for $f(N_{HI}, X)$ follows the definition in equation (7). The physical area per unit N_{HI} is constructed from the average cross-section distribution function over all the galaxies at a common redshift

$$\overline{A_{gal}(N_{HI}, z)} = \sum_{i}^{N_{gal}} \frac{\phi_{gal,i}(N_{HI}, z)}{N_{gal}\Delta N_{HI}}. \qquad (8)$$

We consider the virial radius to define the gas that is physically associated with a galaxy, but we construct also a second model considering the cross-section within $2R_{vir}$ to stress the effects of the cold streams. For the comoving number density of absorbers, we use the cumulative number density[4] of galaxies that are more massive than the least massive galaxy at a given redshift. In practice, we construct an average cross-section that we attribute to all the haloes of equal mass or more massive than the galaxies in our sample. This procedure implies an extrapolation towards higher masses, since systems much above 10^{12} M$_\odot$ are not represented in our sample. Due to the scatter from galaxy to galaxy in the cross-section and to its mass dependence, $\overline{A_{gal}(N_{HI}, z)}$ and the final $f(N_{HI}, X)$ should be regarded as approximations.

In Fig. 10 we compare the functional form of the $f(N_{HI}, X)$ derived for the STAR model at $z \sim 3.5$ within the virial radius (solid black line), within twice the virial radius (dash–dotted line) and in the streams alone (dashed line). The observed distribution at $z \sim 3.7$ from Prochaska et al. (2010) is superimposed with colours. For DLAs and SLLSs (i.e. $N_{HI} > 10^{19}$ cm^{-2} where one resolves the damping wings of Lyα), $f(N_{HI}, X)$ is derived from observations of individual absorbers, but at lower column densities observations provide only integrated constraints on the distribution function. For this reason, the shape of $f(N_{HI}, X)$ in the LLSs and MFP region is more uncertain. In Fig. 10, we report the preferred model from Prochaska et al. (2010).

---

[4] The comoving number density is computed following the approximations described in Dekel & Birnboim (2006, appendix A). Recent simulations (e.g. Klypin, Trujillo-Gomez & Primack 2010) reveal deviations from these approximations, but they become relevant only above $z \sim 4$.
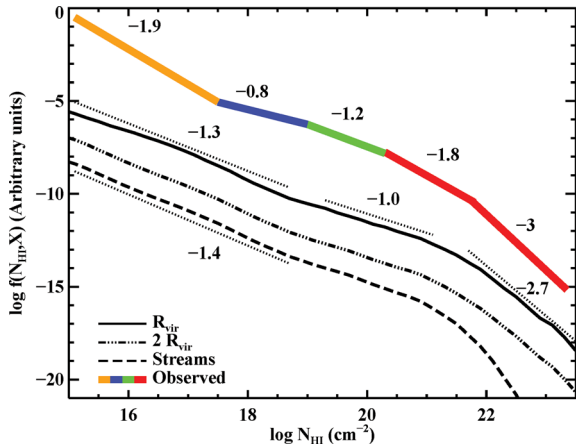
Figure 10. Comparison of the shape of the simulated and observed column density function $f(N_{H_I}, X)$. The normalizations of the different curves are arbitrary, to highlight the shapes only. The simulated $f(N_{H_I}, X)$ from the STAR model at $z \sim 3.5$ is shown for gas inside the virial radius (solid line), for gas within $2R_{vir}$ (dash–dotted line) and for the streams only (dashed line). The observed $f(N_{H_I}, X)$ at $z \sim 3.7$ from Prochaska et al. (2010) is shown as a coloured curve. Power-law regressions to the simulation results in intervals of column density are shown as dotted lines. The numbers denote the power-law indexes. Simulations of massive galaxies at high resolution reproduce the observed cut-off at high column density, but predict an $f(N_{H_I}, X)$ that is shallower than observed around $N_{H_I} \sim 10^{20}$ cm^{-2} and for SLLSs. Cold streams dominate the $f(N_{H_I}, X)$ below $N_{H_I} \sim 10^{18}$ cm^{-2}, but a larger contribution from the IGM is required at lower N_{H_I} to satisfy the constraints imposed by the observed MFP of ionizing radiation.

Although not a perfect match, these models reproduce some of the features of the observed $f(N_{H_I}, X)$. Three different intervals can be identified. First, there is a steep decrease at high column density with an evident break around $N_{H_I} = 3 \times 10^{21}$ cm^{-2}. From our model, we measure a power law ~ -2.7 for log $N_{H_I} \gtrsim 21.7$ (in cm^{-2}; dotted line), similar to the observed slope ~ -3. This cut-off naturally emerges from our high-resolution simulations of massive discs, while it is hard to reproduce in cosmological simulations at lower resolution (e.g. Pontzen et al. 2008; Tescari et al. 2009; Cen 2010, but see Altay et al. 2010; McQuinn, Oh & Faucher-Giguere 2011). The inclusion of molecules would suppress even further the number of systems with H I above 10^{22} cm^{-2} (see Fig. A1), bringing the model to even better agreement with observations.

Moving towards lower column densities, the model exhibits a plateau with slope ~ -1 between $N_{H_I} = 3 \times 10^{18}$ and 10^{21} cm^{-2}. The observed $f(N_{H_I}, X)$ distribution also flattens below 10^{20} cm^{-2} (O'Meara et al. 2007; Prochaska et al. 2010) which has been interpreted as the result of transitioning from predominantly neutral gas to predominantly ionized (Zheng & Miralda-Escudé 2002; McQuinn et al. 2011). In fact, our model flattens at somewhat high N_{H_I} and underpredicts the observed incidence of SLLSs relative to DLAs. Provided that SLLSs arise from galaxies, perhaps this discrepancy would be resolved by including less massive galaxies than those modelled here. As previously noted for the cross-section, we might be missing a population of small clouds of neutral hydrogen due to limited resolution. Finally, below $N_{H_I} = 10^{18}$ cm^{-2} the distribution function becomes moderately steeper with slope ~ -1.3. This qualitatively follows the observed distribution, but it is not sufficiently steep to reproduce the integrated MFP to ionizing radiation at $z \sim 4$ (Prochaska et al. 2009). We infer, therefore, that gas beyond the virial radius must contribute significantly at these column densities (see also Kohler & Gnedin 2007). As already seen in Fig. 4,

due to the stream contribution (dashed line) the $f(N_{H_I}, X)$ steepens mildly below $N_{H_I} \sim 10^{18}$ cm^{-2}, but gas within $2R_{vir}$ is not enough to reproduce the observed slope (dash–dotted line). Contribution from the IGM at larger radii may be required.

Fig. 11 compares again the observed $f(N_{H_I}, X)$ with the model inside the virial radius, including at this time the absolute normalization. The observed $f(N_{H_I}, X)$ has been corrected to match our adopted cosmological parameters. Values are listed in Table B5. This figure reveals an overall satisfactory agreement between the observations and the models for high column density DLAs and at $N_{H_I} \sim 10^{18}$ cm^{-2}. As already noted, the two distributions diverge entering the optically thin regime, suggesting a missing IGM contribution in the simulations. Also missing are SLLSs and high column density LLSs that are not found in large enough numbers within the haloes of massive galaxies. The UVB model, superimposed as a dashed line, has a larger fraction of LLSs and SLLSs, but not enough to account for the missing systems. It is interesting to note that in these models SLLSs originate from gas at the outskirts of the discs and are ionized by the leaking UV photons from stars that are forming in the disc (Schaye 2006). While most of the $f(N_{H_I}, X)$ is shaped by the UVB (Razoumov et al. 2006; Nagamine, Choi & Yajima 2010), RT effects from local sources produce non-negligible effects. Albeit with all the uncertainties discussed, from this simple comparison we infer that massive discs shape the $f(N_{H_I}, X)$ at the highest column densities and cold streams contribute non-negligibly to the LLS population, particularly between $N_{H_I} \sim 10^{17}$ and 10^{18} cm^{-2}.

Finally, in grey, we superimpose different $f(N_{H_I}, X)$ for the red-shift interval $z=4$–1.4. These models are remarkably invariant with redshift and this limited evolution in the $f(N_{H_I}, X)$ is consistent with the observation that the column density distribution function of DLAs preserves its shape over 12 Gyr (Prochaska & Wolfe 2009). This is also consistent with the present-day distribution function from 21-cm observations in H I rich discs. Approaching the optically thin regime, a larger degree of evolution is observed (Prochaska, O'Meara & Worseck 2010; Ribaudo, Lehner & Howk 2011a), which is attributed to RT effects. This variation is not captured by our model, probably because massive galaxies are less sensitive to the effect of the UVB than the IGM and low-mass galaxies which are needed in order to fully match the observed distribution.

The limited evolution in the models presented in Fig. 11 can be qualitatively understood from visual inspection of Fig. 6. As the virial radius grows with time, the gas preserves its distribution with centrally concentrated neutral regions surrounded by SLLSs. In turn, this column density interval marks the transition to more filamentary structures that dominate the LLSs and MFP cross-section. More quantitatively, this non-evolution is related to the fact that the neutral hydrogen column density PDF is found to be nearly self-similar. Therefore, in our simulations, the lack of evolution in $f(N_{H_I}, X)$ is driven by the distribution of gas within each halo (i.e. the cross-section) more than by the cosmological evolution (i.e. the number density of systems).

5.3 Are cold streams visible in absorption?

By integrating the column density distribution function, one can compare predictions from the simulations with the observed incidence of absorbers. Fig. 11 shows that the normalization of the simulated $f(N_{H_I}, X)$ is not too far from the observed one. In these simulations, massive galaxies ($M_{vir} \sim 10^{10}$–10^{12} M$_\odot$) alone can account for $\sim$20–30 per cent of the observed DLA and LLS population. The inclusion of satellites and streams up to twice the
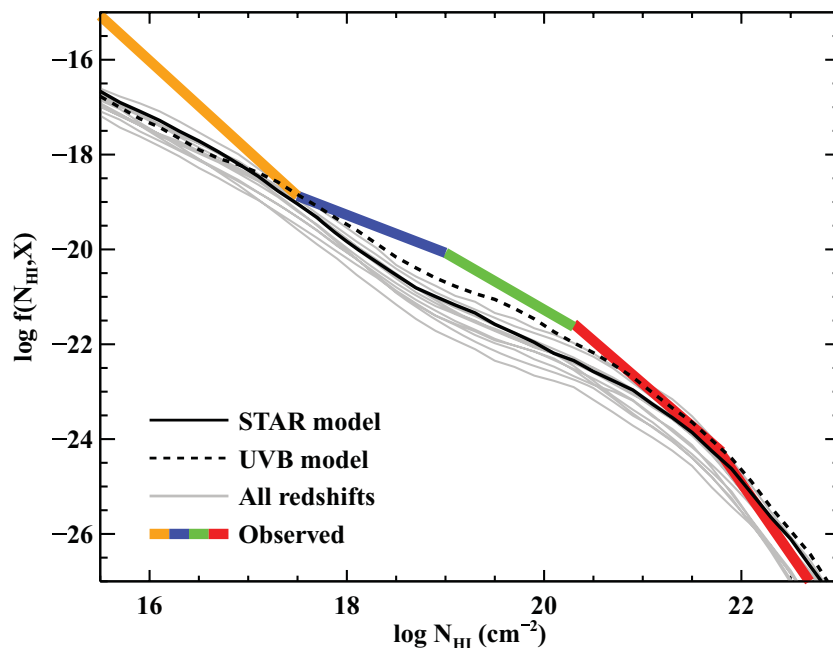
Figure 11. Column density distribution function from observations at $z \sim 3.7$ (colours; Prochaska et al. 2010) and simulations for the STAR (solid line) and the UVB (dashed line) models at $z = 3.5$ within the virial radius. Shown in grey are simulated $f(N_{H_1}, X)$ for different redshifts in the interval $z = 4$–1.4. There is a satisfactory agreement between theory and observations for DLAs and at $N_{H_1} \sim 10^{18}$ cm^{-2}, while the two distributions diverge at lower column densities. SLLSs and high-density LLSs are underrepresented by these models. The $f(N_{H_1}, X)$ exhibits modest redshift evolution in shape, in agreement with observations.

virial radius further increases this contribution (cf. Maller et al. 2003). Although uncertain up to a factor of 2, these numbers suggest that a non-negligible fraction of observed ALSs originate in galaxies and streams as represented in these simulations.

This result is not at odds with the low covering factor for optically thick gas discussed in the previous sections. This is because the incidence of absorbers with background QSOs is sensitive to the physical cross-section times the number density of haloes. Given a large enough number density of haloes, even galaxies with relatively small covering factor can offer enough total cross-section to account for a significant fraction of the observed incidence. Most of the DLAs are estimated by theory to originate in the mass interval 10^9–10^{11} M$_\odot$ (e.g. Pontzen et al. 2008; Tescari et al. 2009), with a peak around $\sim 5 \times 10^{10}$ M$_\odot$. In models with strong feedback (e.g. Nagamine et al. 2007; Cen 2010), the peak shifts to even higher masses. It is not surprising that our sample that includes galaxies with $5 \times 10^{10} - 3 \times 10^{11}$ M$_\odot$ at $z \sim 3$ is within a factor of a few of the observed incidence.

Combining the observed incidence for DLAs and LLSs with the cross-section from our simulations, we can infer the minimum halo mass required to fully match observations. At $z \sim 3.5$ ($z \sim 2.8$), the average cross-section within $2R_{vir}$ is ~ 304 kpc^2 (~ 530 kpc^2) for DLAs and ~ 1548 kpc^2 (~ 2010 kpc^2) for LLSs. The halo number density required to match the observed incidence of DLAs and LLSs (Prochaska & Wolfe 2009; Prochaska et al. 2010; Ribaudo et al. 2011a) is ~ 0.062 Mpc^{-3} (~ 0.032 Mpc^{-3}), corresponding to a minimum halo mass of 1.6×10^{10} M$_\odot$ (3.6×10^{10} M$_\odot$). This is a factor of 1.4 (2.7) below the minimum mass included in this sample. This calculation is very crude since, for example, at lower masses the mean cross-section is expected to decrease (e.g. Fig. 5). Further, these results are sensitive to the radius adopted for the cross-section determination, here chosen to be $2R_{vir}$, i.e. the size of the re-simulated box. Since ideally one would adopt the mean

separation between haloes, larger than $2R_{vir}$ at these masses, our numbers may be even underestimated. With these limitations, our calculation suggests that only a small extrapolation towards lower masses is needed to fully account for the observed incidence of ALSs.

In summary, despite their small covering factor in each halo, the cold streams should have already been detected in absorption in large spectroscopic surveys. Since most of the cross-section in cold streams is predicted to be at $N_{H_1} \lesssim 10^{19}$ cm^{-2} (see Figs 4 and 10), LLSs are the best candidates for the cold gas in the inflowing streams. If there are other components of cold gas in the halo that may not exist in proper abundances in our current simulations (e.g. high-velocity cloud analogues or massive outflows), then other indicators are required to distinguish between them and the cold streams.

6 OBSERVABLE PROPERTIES OF THE CGM

In a second class of experiments designed to probe cold gas in absorption, observers select an object or a class of objects of interest and use background sources to probe the gas distribution around the foreground systems. Using a background quasar (or galaxy) and knowing the angular separation between the probe and the object of interest, this experiment yields a direct measurement of the angular H I distribution (e.g. Christensen et al. 2007; Monier, Turnshek & Rao 2009; Fumagalli et al. 2010b) about the object of interest or of the gas covering factor and kinematics (e.g. Steidel et al. 2010). In the next section, we provide a summary of the formalism adopted in the analysis of absorption lines as well as a more technical description on how line profiles are derived from these simulations. Starting from Section 6.2, we discuss predictions for the metallicity and kinematics of absorption lines, providing a comparison with recent observations.

6.1 Common formalism and numerical procedures

6.1.1 The equivalent width

The normalized profile of an absorption line is characterized by the equivalent width

$$W(\lambda_0) = \frac{\lambda_0}{c} \int_{-\infty}^{+\infty} \left[1 - \bar{I}(v)\right] dv, \qquad (9)$$

with $\bar{I}$ the normalized intensity and λ_0 the rest-frame wavelength at the line centre. The equivalent width provides a direct measurement of the light absorbed along the line of sight. Through the RT equation, $W(\lambda_0)$ can be related to the gas optical depth that in turn depends on the gas column density and kinematics and, via the oscillator strength, on the quantum mechanical properties of the transition.

Three different regimes are typically identified in the so-called curve of growth that describes the variation of the equivalent width with both column density and oscillator strength. At very high column densities, in the damped regime, all the light at the line centre is absorbed and the equivalent width is a solely function of the column density. At intermediate column densities, on the saturated portion of the curve of growth, the line centre is again fully black, but the equivalent width is most sensitive to the line of sight component of the gas velocity that regulates the linewidth. Finally, at low column densities, in the linear regime, the equivalent width becomes proportional to the column density in the observed direction and the absorption profile is no longer zero at the line centre. Averaging along different sightlines, one can further probe the mean covering factor around a class of sources by looking at the fraction of transmitted light close to the line centre.

6.1.2 Techniques for the simulated line profiles

To reproduce the absorption-line profile for Lyα and Si $\text{II}\lambda1260$, we compute the transmitted intensity I_v of a background source with intensity I_0 in a random projected sightline i

$$I_{v,i} = I_0 \exp(-\tau_{v,i}), \qquad (10)$$

where the optical depth $\tau_{v,i}$ is given by

$$\tau_{v,i} = \sum_{k|i} N_k s \phi_k(v). \qquad (11)$$

Here, the sum is over all the AMR cells aligned along the ith line of sight, and N_k is the column density in each cell. The frequency-integrated absorption cross-section s is given by

$$s = \frac{\pi e^2}{m_e c} f, \qquad (12)$$

with f the oscillator strength, e the electron charge, m_e the electron mass and c the speed of light. The frequency-dependent line profile is

$$\phi_k(v) = \frac{H_k(u, a)}{\Delta v_{D,k} \sqrt{\pi}}, \qquad (13)$$

with the Voigt function

$$H(u, a) = \frac{a}{\pi} \int_{-\infty}^{+\infty} dy \frac{\exp(-y^2)}{(u - y)^2 + a^2}. \qquad (14)$$

Here, $a = \gamma/(4\pi\Delta v_D)$, $u = (v - v_l)/\Delta v_D$, $y = v/b$ and $\Delta v_D = (v_0 b)/c$. In the above equations, $b^2 = (2kT/m_X + \zeta^2)$ is the broadening parameter given by that gas temperature and element mass, γ

is the sum over the spontaneous emission coefficients, v is the observed spectral frequency, and $\zeta = 10 \, \text{km s}^{-1}$ the turbulent velocity. We assume the central galaxy is at zero velocity (corresponding to the rest-frame frequency v_0) and we compute for each gas cell the line centre frequency v_l corrected for the gas velocity along the line of sight. We integrate equation (14) using the formalism provided in Zaghloul (2007).

All the spectra are computed at a resolution of $4 \, \text{km s}^{-1}$. For Lyα, we assume $f = 0.4164$ and $\gamma = 6.265 \times 10^8$. For Si $\text{II}\lambda1260$, we use $f = 1.007$, and $\gamma = 2.533 \times 10^9$ and we derive the Si II density from the neutral hydrogen density and metallicity, assuming $\log(\text{Si/H})_{\odot} + 12 = 7.51$ (Asplund et al. 2009). To compensate for the different ionization potential of silicon and hydrogen, we boost the Si II volume density by a factor of 10 in cells with $x_{\text{H\,I}} < 0.1$. The amplitude of this boost depends on the gas and radiation field properties, but it will become clear from our analysis that a variation of several order of magnitudes in the Si II column density is necessary to significantly alter our results. In this simple model, we also neglect Si II recombination at low temperature and a lower collisional rate in shielded regions. These effects alter the Si II column density, but only in regions of small covering factor with minor consequences on the average absorption lines. A complete analysis of metal lines requires the inclusion of collisional ionization and photoionization, which is beyond the scope of this paper.

6.2 Metallicity

The metal enrichment of the gas may help distinguish between gas that is accreting from the cosmic web and the more enriched material that is located in proximity to a galaxy or is outflowing from regions with intense star formation. Production and diffusion of metals are known to be problematic in numerical simulations (Wadsley, Veeravalli & Couchman 2008), but the AMR code used for our simulations should minimize problems of metal transport typical of smoothed particle hydrodynamic simulations (but see Shen, Wadsley & Stinson 2010). Caution is advised in generalizing these comparisons of metallicity between our simulations and observations because our current simulations are limited to supernova-driven winds which do not reproduce outflows as massive as suggested by some observations (e.g. Steidel et al. 2010), and also because the specific IMF (Miller & Scalo 1979) is assumed in the simulations.

In Fig. 12, the column density weighted metallicity in solar units ($Z_{\odot} = 0.0134$; Asplund et al. 2009) for a projection of MW3 at $z \sim 2.3$ is shown. On the top, we display metals in both streams and galaxies for optically thick gas; on the bottom, we present the metallicity in the streams alone. Contours are for $N_{\text{H\,I}} = 1.6 \times 10^{17} \, \text{cm}^{-2}$ (black line) and $N_{\text{H\,I}} = 2 \times 10^{20} \, \text{cm}^{-2}$ (white line). A spatial gradient in metallicity is visible. DLA gas inside the main galaxy and satellites is significantly enriched, with values close to 1/10 solar (in reasonable agreement with DLA observations; e.g. Prochaska et al. 2003). The fraction of metals decreases moving far from the central discs and at lower column densities. The bulk of the LLSs is enriched at 1/10–1/100 solar, with the streams skewed at the lowest metallicity.

A similar metal content characterizes the entire sample at $z = 1.3$–4, almost independently of redshift and mass. Fig. 13 shows the distribution of column density weighted metallicity ψ in DLAs (red dashed line), SLLSs (green dotted lines) and LLSs (blue dash–dotted line). The metal distributions for streams alone (only for $N_{\text{H\,I}} > 1.6 \times 10^{17} \, \text{cm}^{-2}$; solid line) and for satellites without centrals (orange long-dashed line) are superimposed. This
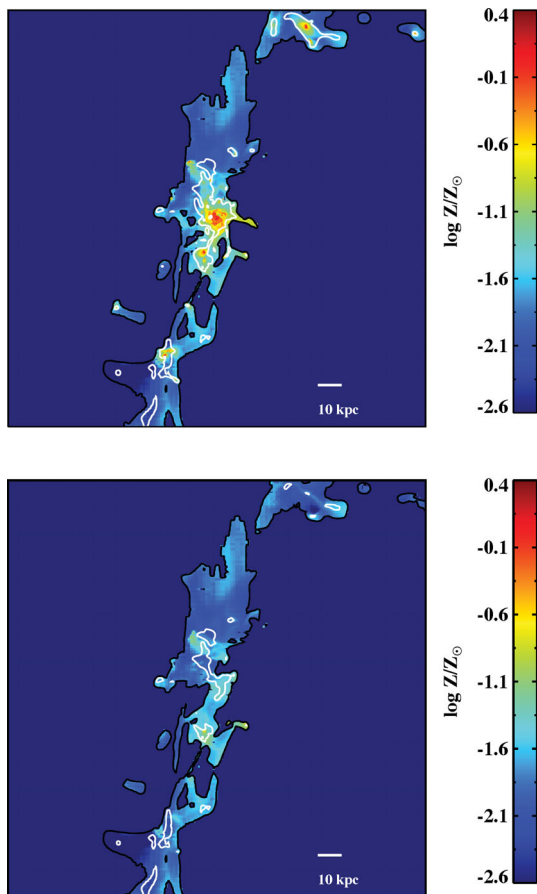
Figure 12. Metallicity in the optically thick gas of MW3 at $z = 2.3$, weighted by hydrogen column density. Top: galaxies and streams. Bottom: streams alone. The black and white contours denote $\log(N_{H_I}/cm^{-2}) = 17.2$ and $\log(N_{H_I}/cm^{-2}) = 20.3$, respectively. A metallicity gradient is visible. Gas in and around galaxies is more enriched than in the streams alone.

distribution has been normalized so that the integral over the metallicity gives the average covering factor in the entire sample for a fixed range of column density

$$\psi(\Delta N_{H_I}, Z) = \frac{\phi(\Delta N_{H_I}, \Delta Z)}{\Delta z \phi_{tot}} . \tag{15}$$

Here, ϕ is the physical cross-section.

From Fig. 13, we see that most of the area that is covered by optically thick gas is occupied by metal-poor cold filaments, with a distribution centred at $Z \sim 10^{-2} Z_{\odot}$ and a spread of over ~ 1 dex. Cold streams are responsible for a metal-poor population of LLSs, while optically thick absorbers that originate within $0.25 R_{vir}$ of centrals and satellites can be enriched above $\log Z/Z_{\odot} = -1$. DLAs are characterized by a broad distribution of metals centred at $\log Z/Z_{\odot} \sim -1$ with a tail to lower metallicity. Satellites alone have a distribution similar to the one of DLAs. SLLSs originate from ionized material in the surroundings of the discs and have an equally broad distribution, centred around $\log Z/Z_{\odot} \sim -1.5$. Outflows more vigorous than the ones generated in our simulations are expected to produce even higher metallicity around the star-forming regions. Since more primordial gas that replenishes galaxy discs with fresh fuel for star formation is predicted to have one order of magnitude less metals than gas in galaxies, we conclude that metal-poor LLSs in proximity to galaxies are the best candidates for cold gas in inflowing streams. The metallicity distribution in LLSs
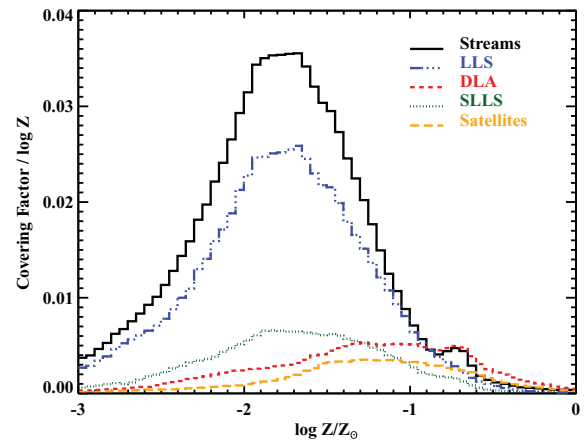

Figure 13. Distribution of metallicity, column density weighted, in the entire sample for DLAs (red dashed line), SLLSs (green dotted line) and LLSs (blue dot–dashed line). Also shown are the streams alone (black solid line) and satellites without centrals (orange long dashed line). The smooth stream component is more metal poor than the galaxies, with typical $Z \sim 0.01 Z_{\odot}$. However, gas in streams is not pristine, but already enriched by previous episodes of star formation.

(e.g. Prochter et al. 2010) may offer a crucial test to distinguish between cold inflowing streams and gas that already resides in the galaxies or is outflowing (see also Giavalisco et al. 2011).

6.3 Absorption-line profiles

The information on the metallicity and covering factor can be combined together with kinematics to formulate predictions on the strength and shape of absorption-line profiles at different projected distances from the galaxy centres. These models can then be compared with the population of ALSs observed in the CGM of LBGs at $z \sim 2$–3 (Steidel et al. 2010). We perform this calculation by stacking several spectra generated along multiple sightlines in a subset of our sample (MW1, MW2, MW3 and MW8 in two projections) at $z \sim 2.3$ and $z \sim 3.2$. The absorption profiles at $z \sim 2.3$ for Lyα (top panel) and Si IIλ1260 (bottom panel) are shown in Fig. 14, where each spectrum is a composite of independent sightlines in five intervals of impact parameter, chosen to match the observations presented in Steidel et al. (2010).

Despite the fact that Lyα saturates at low H I column densities, none of the spectra is black at the line centre due to non-unity covering factor of neutral gas. The transmitted normalized intensity at the line centre ranges in the interval 0.5–0.6, similar to what is observed. Damping wings are visible within the innermost 40 kpc, where most of the highly optically thick gas resides. The line profiles are characterized by a full width at half-maximum (FWHM) velocity of ~ 400–600 km s^{-1} for $b < 41$ kpc, related to the kinematics of the central galaxies and the inner satellite relative motions, as well as the bulk velocity of the incoming streams. Conversely, at larger impact parameters ($b > 41$ kpc), a velocity of 200–300 km s^{-1} FWHM is most likely associated with the inflowing gas that is streaming at ~ 200 km s^{-1} or more (Dekel et al. 2009b). The line profiles are symmetric, despite the fact that the gas is infalling towards the central galaxies, indicating that there are no preferred directions in the velocity field in a stack of multiple sightlines. Inspecting the metal lines, it is evident that the low covering factor and the intrinsic low Si/H abundance, together with the low metallicity
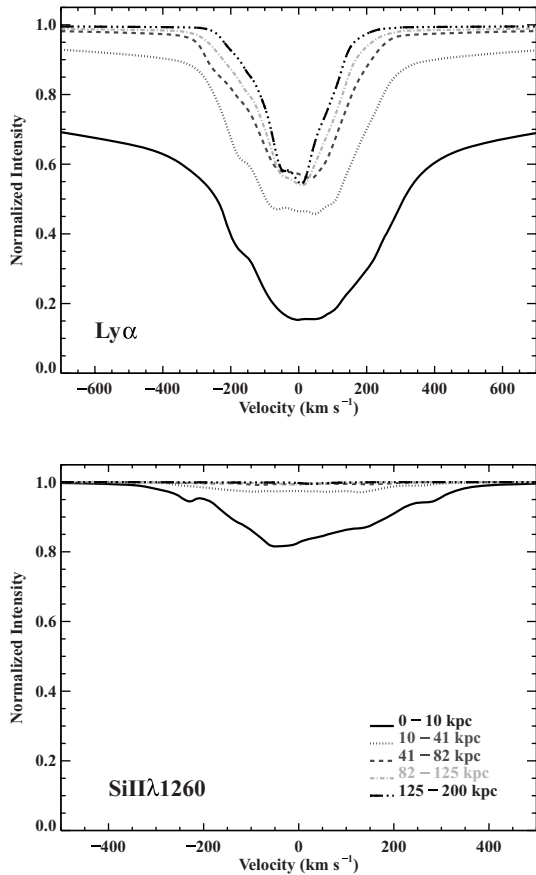
Figure 14. Averaged Lyα (top) and Si ɪɪλ1260 (bottom) absorption lines for background sources at different impact parameters from the central galaxy. The Lyα profiles are never saturated due to the low covering factor and are characterized by an FWHM of ~200–300 km s^{-1} for $b > 41$ kpc. The low covering factor, low intrinsic Si/H abundance and low metallicity produce very weak average Si ɪɪλ1260 absorption.

of the streams, produce absorption lines that are weaker than the observed values at all impact parameters $b > 10$ kpc.

In Fig. 15 we present a comparison between the line strength from simulations and observations. For each radial bin, the rest-frame equivalent width of Lyα at $z \sim 2.3$ and $z \sim 3.2$ is shown with red open squares and crosses, respectively. Similarly, we superimpose the equivalent width of Si ɪɪλ1260 at $z \sim 2.3$ (blue open circles) and $z \sim 3.2$ (blue x). Values are listed in Table B7. The black dotted and dashed lines are for the observed rest-frame equivalent width from spectra of LBGs at redshift $z \sim 2$–3 (Steidel et al. 2010). As expected from direct inspection of the absorption profiles, massive galaxies that are accreting gas through cold streams have a large enough covering factor and sufficiently complex kinematics in H ɪ gas to match observed Lyα values. We emphasize that this is achieved through the gas associated with satellites and streams alone; the supernova-driven winds in our current simulations are hot and not as strong as the outflows suggested by Steidel et al. (2010). While winds, commonly observed in star-forming galaxies, are essential ingredients for realistic models, the fact that the same observations can be reproduced by two extreme scenarios stands as a reminder that evidence for outflowing/inflowing gas are subtle and more refined models are now required.

That these simulations fail to reproduce the strength of Si ɪɪ at all impact parameters $b < 100$ kpc is not surprising since, due to
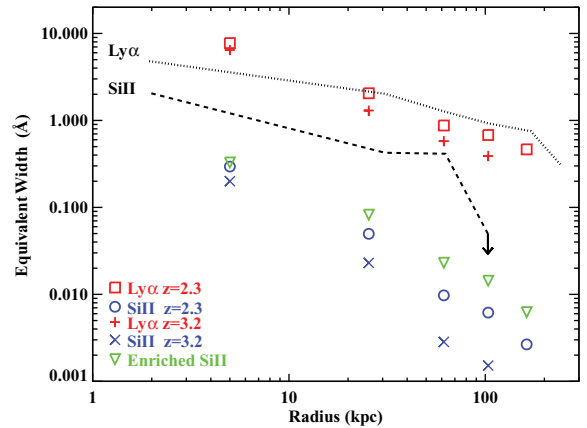

Figure 15. Equivalent width profile at $z \sim 2.3$ and $z \sim 3.2$ for Lyα (red open squares and crosses) and Si ɪɪλ1260 (blue open circles and x). The black dotted and dashed lines show the observed rest-frame equivalent widths in LBGs at redshift $z \sim 2$–3 (Steidel et al. 2010). Massive galaxies that are accreting gas through cold streams have a large enough covering factor and line-of-sight velocity to account for the observed strength of the Lyα absorption. Conversely, the metal-poor streams cannot reproduce the observed equivalent width in metal lines. Boosting the degree of enrichment by introducing an artificial metallicity floor at $0.1 \, Z_\odot$ (green downward triangles) is not enough for matching the observations.

the lack of strong outflows of cold gas, low-ion metals are tracing the distribution of hydrogen in the galaxies and streams. We can consider whether, for example, the discrepancy for Si ɪɪλ1260 might be attributable to the low enrichment of the cold streams. In Fig. 16, we show the ratio of the Si ɪɪ to Lyα optical depth, integrated in the velocity interval ±250 km s^{-1}. Only gas with $N_{\mathrm{H\,I}} > 10^{13}$ cm^{-2} is displayed. The low abundance of silicon relative to hydrogen at solar metallicity suppresses the metal line opacity by five orders of magnitude and the low metallicity of cold streams decreases the Si ɪɪ opacity even further, by a factor of ~10–100. For these reasons, metal lines that originate in cold streams are weak and likely to be undetected. A higher degree of enrichment is not enough to boost the equivalent width by one order of magnitude, as demonstrated by a simple experiment (see also Kimm et al. 2011). If we artificially
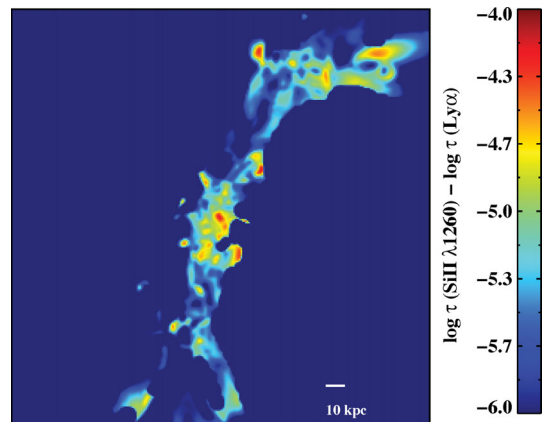
Figure 16. Optical depth integrated over velocities in the interval ±250 km s^{-1} for Si ɪɪλ1260 relative to Lyα for gas with $N_{\mathrm{H\,I}} > 10^{13}$ cm^{-2}. The low intrinsic Si/H abundance and metal content of cold streams suppress the metal line opacity by more than five orders of magnitude compared to Lyα.

impose a floor in metallicity at $Z = 0.1Z_\odot$ for all the gas with $N_{H\,I} > 10^{15}$ cm^{-2}, we obtain the equivalent widths shown by the green downwards triangles in Fig. 15. A much higher degree of enrichment increases the line strength only by a modest value.

We conclude that cold streams are unlikely to produce the large equivalent widths of low-ion metal absorption around massive galaxies. If the average equivalent widths reported by Steidel et al. (2010) are confirmed (e.g. with higher resolution observations), an alternate source of opacity must be included in the simulations. This could include, for example, absorption by neighbouring galaxies (not modelled here), small clumps of cold gas embedded in a hotter medium, and/or galactic-scale outflows. However, in order to satisfy the constraints imposed by the Lyα absorption, these phenomena cannot carry a very large column density of neutral hydrogen nor contribute to the kinematics apparent in the cold flows unless they replace the streams altogether. This would fail to supply the gas necessary to produce the observed high SFR plus a similar outflow rate. Unless these SFRs and outflows are sustained by a large amount of reservoir gas, the inflows and outflows must live together. This is a natural consequence of the fact that the inflows are along narrow dense streams, while the outflows are covering a large solid angle between the streams. Additional work in simulations is needed to explore stronger wind models and to understand the interaction between outflowing and inflowing gas (e.g. Razoumov 2009; Powell, Slyz & Devriendt 2010; van de Voort et al. 2011), before we can reconcile both the Lyα and metal absorption profiles with a combination of inflows and outflows.

7 SUMMARY AND CONCLUSIONS

We investigated the characteristics of cold gas in cold-stream fed galaxies in absorption using a sample of seven galaxies drawn at random for zoom-in cosmological hydro-AMR simulations with high resolution (35–70 pc) in the halo mass range 10^{10}–10^{12} M$_\odot$ between $z \sim 1.4$ and 4. We considered the contributions of the cold gas in the central galaxies and of the inflowing cold gas in the two components of the streams that feed the high-redshift massive galaxies from the cosmic web – the incoming small galaxies and the larger smooth component. Our analysis focused on the mass dependence and time evolution of optically thick and thin gas cross-section and covering factor, on the metallicity distribution in the streams and massive galaxies, and on the Lyα and metal line kinematics. The simulations were compared with available observations of ALSs in the foreground of quasars and with kinematics of the CGM in LBGs to highlight what observables are most useful to detect cold streams in absorption. The limited sample size and the fact that very strong outflows are not included in these simulations make our predictions rather specific to the study of cold streams, and caution is advised in generalizing our findings to the full population of ALSs. Our results and findings can be summarized as follows.

(i) After post-processing the simulations with an RT code that includes dust, collisional ionization and photoionization from the UVB and local stellar sources, cold streams appear highly ionized and not entirely self-shielded structures. Pockets of high column density neutral hydrogen, mostly associated with the central galaxy and instreaming galaxies, are embedded in a ionized medium with $N_{H\,I} \lesssim 10^{19}$ cm^{-2}. The densest regions are mostly affected by the UV radiation from local sources because the UVB is completely shielded in regions of $n_H > 0.1$ cm^{-3}. While the UV from stars is largely confined by hydrogen and dust absorption in the star-forming regions, the non-zero escape fraction is responsible for

the ionization of the gas in the immediate surroundings of discs ($10^{18} \lesssim N_{H\,I}/cm^{-2} \lesssim 10^{20}$). Since most of the mass is found at the highest column densities, UV from stars decreases the total neutral hydrogen mass in galaxies.

(ii) The smooth component of the cold streams dominates the cross-section of neutral hydrogen below $N_{H\,I} = 10^{18}$ cm^{-2} within $2R_{vir}$. At $z \leq 2$, less than half of the cross-section in the range $N_{H\,I} = 10^{19}$–10^{20} cm^{-2} is due to the streams. At these low redshifts, most of the cross-section for $N_{H\,I} > 10^{20}$ cm^{-2} is in galaxies, with the smooth component of the streams accounting only for a small fraction of the total budget. However, a rapid evolution is seen with redshift, with the gas clumps in the streams responsible for more than half of the DLA/SLLS covering factor at $z \sim 4$.

(iii) No more than 60 per cent of the DLA cross-section is associated with the central galaxies. Satellites and gas clumps in streams are important or even the dominant contributors to the neutral gas cross-section, especially in the most massive galaxies.

(iv) At all redshifts, the covering factor at $2R_{vir}$ is <40 per cent for $N_{H\,I} > 10^{15}$ cm^{-2}, <10 per cent for LLSs and <1 per cent for DLAs. The H I gas covering factor within $2R_{vir}$ appears to decrease slowly from $z \sim 4$ to $z \sim 1.5$, and the covering factor within a fixed aperture of 10 arcsec ($\sim$70–85 kpc) seems not evolving or slowly increasing with redshift. However, the small sample and the non-homogeneous sampling of the halo mass function at all the redshifts prevent us from deriving robust scaling relations. In these simulations, a certain degree of self-similarity is found in the column density probability distribution function for neutral hydrogen and the time evolution of the different covering factors appears to follow the universal expansion.

(v) Cold streams mostly appear as LLSs and contribute significantly to the column density distribution function in between $N_{H\,I} = 10^{17}$ and 10^{18} cm^{-2}. Massive discs, as described by these high-resolution simulations, shape the column density distribution function for very high column density DLAs. These models, especially if molecules are included, can reproduce the observed knee at $10^{21.5}$ cm^{-2} in the $f(N_{H\,I}, X)$. Conversely, there is a deficiency of systems between $10^{18.5}$ and 10^{20} cm^{-2}, likely due to the missing low-mass systems in our simulated sample and to finite numerical resolution which does not capture a clumpy medium. The simulated $f(N_{H\,I}, X)$ is very shallow below 10^{17} cm^{-2}, indicating that cold streams in proximity to massive galaxies are not the only sources of opacity for Lyman limit photons in the Universe; an additional contribution from the IGM is required. Finally, in agreement with observations, the simulated $f(N_{H\,I}, X)$ is not evolving in shape with redshift, due to a nearly self-similar distribution of neutral hydrogen in haloes.

(vi) Despite the small covering factor, comparing the observed and simulated incidence of ALSs, it appears that massive galaxies and cold streams are responsible for at least 30 per cent of the absorbers detected in the foreground of QSOs. In fact, due to a large enough number density of galaxies in the Universe, the cross-section required to match the observed incidence is only a factor of a few above the mean values in this sample. A modest extrapolation towards lower halo masses can fully reproduce the observed population. To the extent that cold streams are indeed the common mode of accreting gas in massive galaxies at high redshift, as emerges from the cosmological simulations, these streams have already been detected in absorption-line surveys, primarily as metal-poor LLSs.

(vii) In our simulations, most of the cross-section within $2R_{vir}$ is occupied by metal-poor gas ($Z \sim 10^{-2} Z_\odot$), which has been partly enriched by previous episodes of star formation. DLAs and SLLSs exhibit broad metallicity distributions, with a peak

at $Z \sim 10^{-1}\,Z_\odot$ and $Z \sim 10^{-1.5}\,Z_\odot$, respectively. Metallicity is a valuable tool to disentangle gas that is associated with galaxies or winds (with metallicity distribution skewed above $Z \sim 10^{-1.5}\,Z_\odot$) from LLSs in the cold streams (with the lowest level of enrichment).

(viii) Because of the partial covering factor of neutral gas, composites of Lyα absorption lines are not black at the line centre, even within 40 kpc from the central galaxy. The typical linewidth outside the central galaxy and its immediate neighbourhood is $\sim$200–300 km s^{-1} FWHM reflecting the bulk inflow velocity in cold streams. Our current simulations reproduce the Lyα equivalent width distribution observed in the CGM of LBGs. Conversely, the predicted optical depth of metal lines is suppressed by more than 5 orders of magnitude and the predicted equivalent width of Si $\text{II}\,\lambda1260$ is systematically lower than the observed values. An increase in the metallicity of the streams is not enough to match the observations. Another mechanism is required, such as massive outflows beyond the existing winds in the current simulations.

From the systematic analysis of our simulations and a careful comparison with observations, we conclude that the theoretical predictions of cold gas accretion along filaments of the cosmic web are consistent with present-day observations. This conclusion is not driven by a null claim that the cold streams are undetectable in absorption, which would have been rather unsatisfactory. In fact, our current simulations, with all the caveats previously discussed, indicate that cold streams contribute significantly to the incidence and kinematics of absorbers observed at redshift $z \sim 2$–3.

Progress can now be made in two directions. While this work offers a first quantitative analysis of detectability of cold streams in absorption, simulations representative of a larger spectrum of masses will facilitate robust statistical comparisons with observations. At the same time, improved modelling of outflows in simulations will allow a more detailed exploration of the interaction between inflowing and outflowing gas. This is important to better assess the cross-section distribution of neutral hydrogen, a key quantity for the visibility of cold streams in absorption.

Observationally, it should be possible to identify candidates of cold streams by searching the increasing spectroscopic data sets of quasars and galaxy pairs. Despite the difficulties of measuring metallicity in LLSs, one can investigate their kinematics in detail to disentangle between metal-rich gas that is outflowing at velocities >600 km s^{-1} from gas that is inflowing at velocities $\lesssim 200$ km s^{-1} with metallicity $\lesssim 0.01\,Z_\odot$.

Observed metal-poor LLSs are suggestive of the cold mode of inflow (e.g. Prochter et al. 2010; Ribaudo et al. 2011b) and a confirmation of this phenomenon may arise if a nearby galaxy is observed to be fed by cold inflows. However, at low redshift the cold mode is expected to be limited to galaxies with stellar mass below $10^{11}\,M_\odot$ (Dekel & Birnboim 2006), and to accrete from a wide angle at a lower column density. This is significantly different from the cold, narrow, dense streams that are predicted at high redshift and only direct evidence for cold inflow at high redshift would be a confirmation of these theoretical predictions. In future, models of cold gas accretion can be also tested statistically against observations, provided reliable estimates of observables such as the column density distribution, metallicities and kinematics from simulations.

Following these paths, it may be possible to uncover the modes of gas accretion in high-redshift galaxies, and their interplay with outflows.

REFERENCES

Abel T., Wandelt B. D., 2002, MNRAS, 330, L53
Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45
Agertz O., Teyssier R., Moore B., 2009, MNRAS, 397, L64
Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, ApJ, 737, L37
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Bahcall J. N., Peebles P. J. E., 1969, ApJ, 156, L7
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Birnboim Y., Dekel A., 2011, MNRAS, 415, 2566
Cantalupo S., 2010, MNRAS, 403, L16
Cantalupo S., Porciani C., 2011, MNRAS, 411, 1678
Cantalupo S., Porciani C., Lilly S. J., Miniati F., 2005, ApJ, 628, 61
Cen R., 2010, preprint (arXiv:1010.5014)
Ceverino D., Klypin A., 2009, ApJ, 695, 292
Ceverino D., Dekel A., Bournaud F., 2010, MNRAS, 404, 2151
Christensen L., Wisotzki L., Roth M. M., Sánchez S. F., Kelz A., Jahnke K., 2007, A&A, 468, 587
Daddi E. et al., 2010, ApJ, 713, 686
Dall'Aglio A., Wisotzki L., Worseck G., 2009, preprint (arXiv:0906.1484)
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A., Birnboim Y., 2008, MNRAS, 383, 119
Dekel A., Sari R., Ceverino D., 2009a, ApJ, 703, 785
Dekel A. et al., 2009b, Nat, 457, 451
Dijkstra M., Loeb A., 2009, MNRAS, 400, 1109
Draine B. T., 2003, ARA&A, 41, 241
Dwek E., 1998, ApJ, 501, 643
Fardal M. A., Katz N., Gardner J. P., Hernquist L., Weinberg D. H., Davé R., 2001, ApJ, 562, 605
Faucher-Giguère C.-A., Kereš D., 2011, MNRAS, 412, L118
Faucher-Giguere C., Keres D., Ma C., 2011, preprint (arXiv:1103.0001)
Faucher-Giguère C.-A., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 682, L9
Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, ApJ, 725, 633
Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761
Fumagalli M., Krumholz M. R., Hunt L. K., 2010a, ApJ, 722, 919
Fumagalli M., O'Meara J. M., Prochaska J. X., Kanekar N., 2010b, MNRAS, 408, 362
Furlanetto S. R., Schaye J., Springel V., Hernquist L., 2005, ApJ, 622, 7
Giavalisco M. et al., 2011, preprint (arXiv:1106.1205)
Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
Gnat O., Sternberg A., 2007, ApJS, 168, 213
Gnedin N. Y., 2010, ApJ, 721, L79
Gnedin N. Y., Abel T., 2001, New Astron., 6, 437
Goerdt T., Dekel A., Sternberg A., Ceverino D., Teyssier R., Primack J. R., 2010, MNRAS, 407, 613
Haardt F., Madau P., 1996, ApJ, 461, 20
Haiman Z., Spaans M., Quataert E., 2000, ApJ, 537, L5

Hennawi J. F., Prochaska J. X., 2007, ApJ, 655, 735
Hennawi J. F. et al., 2006, ApJ, 651, 61
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Kereš D., Hernquist L., 2009, ApJ, 700, L1
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kimm T., Slyz A., Devriendt J., Pichon C., 2011, MNRAS, 413, L51
Klypin A., Trujillo-Gomez S., Primack J., 2010, preprint (arXiv:1002.3660)
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kohler K., Gnedin N. Y., 2007, ApJ, 655, 685
Komatsu E. et al., 2009, ApJS, 180, 330
Kravtsov A. V., 2003, ApJ, 590, L1
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kroupa P., 2001, MNRAS, 322, 231
Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36
Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 693, 216
Laursen P., Sommer-Larsen J., Andersen A. C., 2009, ApJ, 704, 1640
Leitherer C. et al., 1999, ApJS, 123, 3
Lejeune T., Schaerer D., 2001, A&A, 366, 538
Li A., Draine B. T., 2001, ApJ, 554, 778
McKee C. F., Krumholz M. R., 2010, ApJ, 709, 308
McQuinn M., Oh S. P., Faucher-Giguere C.-A., 2011, preprint (arXiv:1101.1964)
Maller A. H., Prochaska J. X., Somerville R. S., Primack J. R., 2001, MNRAS, 326, 1475
Maller A. H., Prochaska J. X., Somerville R. S., Primack J. R., 2003, MNRAS, 343, 268
Matsuda Y. et al., 2010, MNRAS, L171
Miller G. E., Scalo J. M., 1979, ApJS, 41, 513
Monier E. M., Turnshek D. A., Rao S., 2009, MNRAS, 397, 943
Nagamine K., Wolfe A. M., Hernquist L., Springel V., 2007, ApJ, 660, 945
Nagamine K., Choi J.-H., Yajima H., 2010, ApJ, 725, L219
Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087
O'Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Petkova M., Springel V., 2009, MNRAS, 396, 1383
Pontzen A. et al., 2008, MNRAS, 390, 1349
Powell L. C., Slyz A., Devriendt J., 2011, MNRAS, 414, 3671
Prochaska J. X., Hennawi J. F., 2009, ApJ, 690, 1558
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Prochaska J. X., Gawiser E., Wolfe A. M., Castro S., Djorgovski S. G., 2003, ApJ, 595, L9
Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123
Prochaska J. X., Worseck G., O'Meara J. M., 2009, ApJ, 705, L113
Prochaska J. X., O'Meara J. M., Worseck G., 2010, ApJ, 718, 392
Prochter G. E., Prochaska J. X., O'Meara J. M., Burles S., Bernstein R. A., 2010, ApJ, 708, 1221
Rafelski M., Wolfe A. M., Chen H.-W., 2011, ApJ, 736, 48
Razoumov A. O., 2009, ApJ, 707, 738
Razoumov A. O., Norman M. L., Prochaska J. X., Wolfe A. M., 2006, ApJ, 645, 55
Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, ApJ, 712, 1070
Ribaudo J., Lehner N., Howk J. C., 2011a, ApJ, 735, L1
Ribaudo J., Lehner N., Howk J. C., Werk J. K., Tripp T. M., Prochaska J. X., Meiring J. D., Tumlinson J., 2011b, preprint (arXiv:1105.5381)
Schaye J., 2006, ApJ, 643, 59
Shen S., Wadsley J., Stinson G., 2010, MNRAS, 407, 1581
Simcoe R. A., Sargent W. L. W., Rauch M., Becker G., 2006, ApJ, 637, 648
Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289
Stewart K. R., Kaufmann T., Bullock J. S., Barton E. J., Maller A. H., Diemand J., Wadsley J., 2011, ApJ, 735, L1
Tacconi L. J. et al., 2010, Nat, 463, 781
Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411

van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, MNRAS, 414, 2458
Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427
Wise J. H., Abel T., 2011, MNRAS, 414, 3458
Wood K., Haffner L. M., Reynolds R. J., Mathis J. S., Madsen G., 2005, ApJ, 633, 295
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181
Zaghloul M. R., 2007, MNRAS, 375, 1043
Zheng Z., Miralda Escudé J., 2002, ApJ, 568, L71
Zwaan M. A., Prochaska J. X., 2006, ApJ, 643, 675

APPENDIX A: A DETAILED COMPARISON OF THE RT CALCULATIONS

In this Appendix, we provide a more detailed comparison of the results from the different RT calculations. At first, we discuss the effects that the UVB and local sources have on the column density distribution. Then, we quantify the hydrogen density threshold above which the gas is self-shielded and we provide a numerical approximation useful to correct the CIE approximation in the absence of the UVB. Finally, we examine the impact that photoionization has on the neutral mass of galaxies and satellites.

A1 Effects on the hydrogen column density distribution

In Fig. A1, we highlight the effect that different ionization sources have on the H I column density distribution by showing the number of AMR cells in bins of $N_{H I}$ from an arbitrary projection of MW3 at $z = 2.3$. The different RT models have been normalized with respect to the STAR model (black horizontal solid line). The CIE model (blue dash–dotted line) offers an upper limit to the neutral fraction at all the column densities. The CIE approximation holds only in highly optically thick regions ($N_{H I} \gtrsim 10^{20}$ cm^{-2}), while at lower column density overpredicts by a large amount the neutral fraction in comparison to the UVB case (red dashed line). Photoionization cannot be neglected, particularly given the fact that streams are not entirely self-shielded, as evident from Fig. 3.
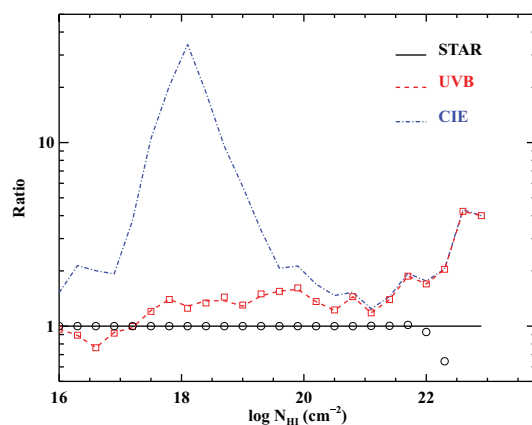
Figure A1. Comparison of the neutral hydrogen column density in the STAR model (black line), the CIE model (blue dash–dotted line) and the UVB model, with and without dust (red dashed line and open squares). The STAR model corrected for molecular hydrogen is shown with open circles. For each RT calculation, we plot the number of projected cells in a given interval of $N_{H I}$, in comparison to the STAR model. The CIE approximation holds only in highly optically thick regions. At high column densities, neutral hydrogen is further ionized by local sources and depleted by molecules.

Comparing the UVB with the STAR model, it appears that radiation from local sources affects primarily the high column densities ($N_{H_I} \gtrsim 10^{21}$ cm^{-2}) which are completely self-shielded from the UVB. Since most of the ionizing photons are absorbed locally, the effect of local sources is modest below $N_{H_I} \sim 10^{20}$ cm^{-2} and almost negligible below $N_{H_I} \sim 10^{18}$ cm^{-2}. Note that quantities shown in Fig. A1 are normalized and not proportional to the area covered by gas at a given column density. In fact, radiation from local sources that leaks from the central disc can affect the H I column density distribution in the interval 10^{18}–10^{20} cm^{-2}, when weighted by area. Dust produces little or no effect on the UVB model (compare the red dashed line with open squares).

Although not included in the our RT calculation, in Fig. A1 we explore how molecules affect the shape of the hydrogen column density distribution. When atomic hydrogen is shielded from UV radiation, molecules can form, lowering the effective H I column density. To investigate the importance of this effect, we compute the molecular fraction following the McKee & Krumholz (2010) formalism. This analytic model reproduces to the first order the molecular gas fraction in nearby spirals (Krumholz, McKee & Tumlinson 2009) and metal-poor dwarfs (Fumagalli, Krumholz & Hunt 2010a) and approximates the results from numerical simula-

tions (Krumholz & Gnedin 2011). In each AMR cell, we estimate the molecular gas fraction as a function of the total neutral gas column density $\Sigma_{gas} = 2m_p N_{H_2} + m_p N_{H_I}$ and metallicity. Since this model is designed to describe the molecular fraction in individual clouds, we correct for smearing due to the resolution by assuming $\Sigma_{cloud} = c\Sigma_{gas}$ with $c = 5$ for all the AMR cells with sizes above 200 pc and $c = 1$ elsewhere. Although not a dominant correction, the inclusion of molecules lowers the H I column densities above $N_{H_I} \sim 10^{22}$ cm^{-2} (open circles). For this reason, the values of N_{H_I} used in our analysis should be regarded as upper limits at the highest density. This column density threshold is in good agreement with the knee in the observed DLA column density distribution (Prochaska et al. 2005; Noterdaeme et al. 2009) that, indeed, is interpreted as the transition between atomic and molecular gas (e.g. Zwaan & Prochaska 2006).

A2 Self-shielding and numerical approximations

Fig. A2 shows the neutral fraction x_{H_I} as a function of total gas volume density n_H, for the CIE model (top-left panel), the UVB model (top-right panel) and the STAR model (bottom-left panel), again for MW3 at $z = 2.3$. In the presence of the UVB, hydrogen
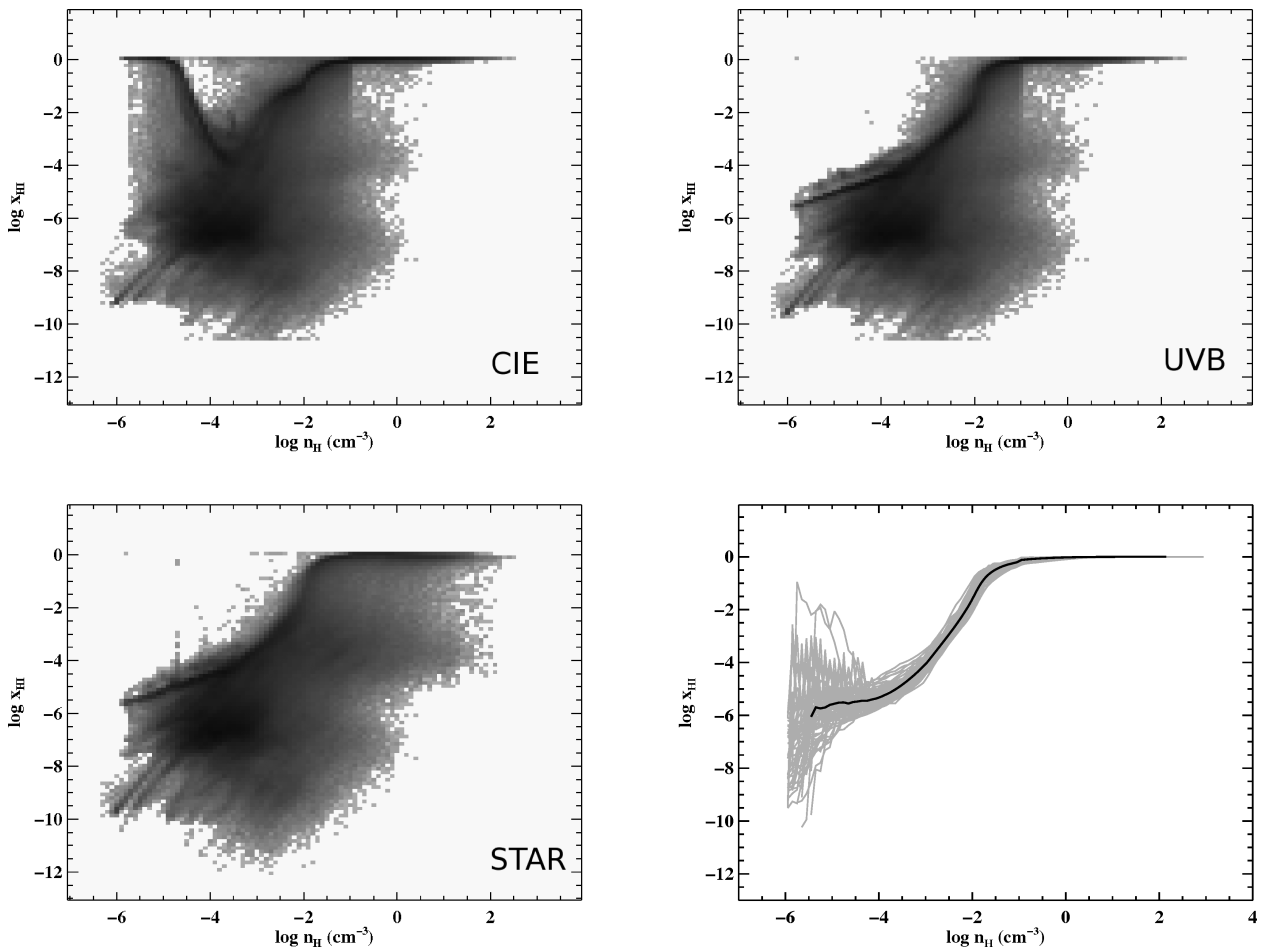
Figure A2. Neutral fraction as a function of the hydrogen volume density for the CIE model (top left), the UVB model (top right) and the STAR model (bottom left) in MW3 at $z \sim 2.3$. The colour scale is proportional to the number of AMR cells in each bin, with the most populated region in dark grey. The arithmetic mean in bin of $\log n_H$ for the UVB model in individual galaxies is shown in the bottom-right panel (grey lines). The central values common to all the galaxies are also superimposed (black solid line). Although with significant scatter, this distribution can be used as a crude approximation to the UVB model. RT with local sources is required to capture the intrinsic scatter in the neutral fraction above the $n_H \sim 0.01$–0.1 cm^{-3}, where gas becomes shielded from the UVB.

is highly ionized below $n_H = 0.01 \, \mathrm{cm}^{-3}$, whereas it is fully neutral above $n_H = 0.1 \, \mathrm{cm}^{-3}$. A similar volume density for self-shielding has been consistently found with different numerical techniques (e.g. Razoumov et al. 2006; Faucher-Giguère et al. 2010), justifying the use of a fixed threshold above which the UVB can be suppressed while running the simulations.

Additional photoheating, currently neglected in our RT calculations, will produce a smoother transition between ionized and shielded regions due the enhanced collisional ionization in partially shielded cells. Also, both the UVB and CIE models do not account for the presence of ionizing sources that increase the dispersion of $x_{H\,I}$ (bottom-left panel in Fig. A2). A significant fraction of the gas above $n_H = 0.01$–$0.1 \, \mathrm{cm}^{-3}$ remains fully neutral, but local fluctuations in the ionization state of up to 5 order of magnitude can be found. Indeed, a locus of points with $n_H > 0.01 \, \mathrm{cm}^{-3}$ and $x_{H\,I} < 0.01$ is visible in Fig. A2. Since this gas mostly resides in the central disc and in satellites where the majority of the star formation is present, the inclusion of ionizing photons seems to be an essential ingredient in models of galaxy formation. Ionizing photons from stars are absorbed mostly in the inner part of the galaxy by dust and gas and there is only a minor contribution of the ionization below $n_H = 0.01 \, \mathrm{cm}^{-3}$, in qualitative agreement with current constraints on the escape fraction of ionizing radiation. However, the degree of precision required in comparing models and observations requires the inclusion of this RT effects on the neutral gas, commonly neglected or treated with some approximations.

Although below $n_H = 0.01 \, \mathrm{cm}^{-3}$ the scatter in $x_{H\,I}$ at a given volume density is large, most of the AMR cells accumulate in a well-defined locus of points. Due to a self-similarity in the probability distribution function of the hydrogen column density, there is little scatter for different galaxies with both mass and redshift, as seen from the bottom-right panel of Fig. A2, where we plot the arithmetic mean of $x_{H\,I}$ for the UVB model (grey lines for individual galaxies). This invariance not only corroborates the assumption of a fixed threshold above which the UVB is suppressed, but also allows one to assume a lower limit on the hydrogen volume density in systems that host neutral gas, such as DLAs. Further, we can derive a crude approximation to the UVB model. In the CIE model, fully neutral or marginally ionized gas is found even below $n_H = 0.01 \, \mathrm{cm}^{-3}$, with an error up to several order of magnitudes in comparison to the UVB model. Using the fact that the mean $x_{H\,I}$ as a function of n_H is sensitive to the intensity of the UVB only, we derive a relation $x_{H\,I} = x_{H\,I}(n_H)$ that improves the CIE model in presence of a UVB. We obtain such relation by combining all the arithmetic means computed in individual galaxies with a geometric mean (see Fig. A2 and Table B3).

A3 Neutral hydrogen mass

For each RT model, we measure the total H I mass enclosed in the virial radius of the main haloes and satellites (see Table B2). A comparison is provided in Fig. A3. In the top panel, we show the relation between the total and the neutral hydrogen mass in the STAR model. In massive haloes above $10^9 \, \mathrm{M_\odot}$, less than $\sim 1/5$ of the hydrogen mass is neutral. A robust linear regression in logarithmic space for the main haloes only gives $\log M_{H\,I} = 0.03 + 0.93 \log M_H$. Less massive satellites cover a wider range in H I mass, from almost neutral to highly ionized. The CIE and UVB models predict nearly identical masses above $10^9 \, \mathrm{M_\odot}$ ($\log M_{H\,I,CIE} = 0.02 + 1.00 \log M_{H\,I,UVB}$ for the main haloes), while in the satellites the CIE model predicts larger neutral hydrogen masses. H I masses from the STAR model are systematically lower

than those in the CIE or the UVB model above $10^9 \, \mathrm{M_\odot}$, up to 50–60 per cent above $10^{10} \, \mathrm{M_\odot}$ ($\log M_{H\,I,STAR} = 0.41 + 0.93 \log M_{H\,I,CIE}$ and $\log M_{H\,I,STAR} = 0.45 + 0.93 \log M_{H\,I,UVB}$ for the main haloes). Satellites have a larger scatter.

The explanation for this behaviour naturally emerges from the previous considerations. In the main haloes, most of the mass is in high-density regions where the gas is self-shielded from the UVB. Photoionization from local sources and, to the second order, electron collisions regulates the amount of mass in neutral hydrogen. For satellites, the observed scatter is due to the different SFRs in these systems and on the different distances from the central discs. Also, at lower masses, systems are more vulnerable to ionization from the UVB. This behaviour is qualitatively in agreement with the simulations by Gnedin (2010) who finds that the Lyman–Werner radiation from the UVB is negligible compared to the local radiation field and that the ionizing radiation from the UVB plays a significant role only up to the edge of the H I discs. This result has implications for understanding the processes that shape the massive end of the H I mass function and stresses further the necessity to account for local photoionization as an important feedback mechanism (e.g. Cantalupo 2010).

A4 Uncertainties and future improvements

Although our model significantly improves upon simpler RT calculations, few approximations contribute the final error budget. For example, these simulations do not include the presence of an active galactic nucleus and additional ionization from this harder spectrum is neglected in this work. Similarly, we include only photoionization at 912 Å due to stars and we do not consider harder spectral energy distributions (SEDs) that could enhance the ionization via heating. Similarly, He recombination emission is neglected. Finally, we ignore the possibility that these galaxies could lie in the proximity of other ionizing sources, such as QSOs (e.g. Cantalupo et al. 2005). We note that these effects would reduce further the neutral gas fraction and enhance the effects of ionizing radiation on the neutral volume and column density.

In the previous sections, we have clearly shown that the inclusion of both the UVB and local sources shapes the distribution of gas volume and column density at the low and high densities. However, the exact amplitude of these effects is proportional to the intensity of the input radiation field which is subject to uncertainties, as discussed in Section 3.1. Heating from the UVB and local sources is already included in these simulations, but during the post-processing we do not update the gas temperature (see a discussion in Cantalupo & Porciani 2011). Moreover, a sensible determination of the temperature has obvious implications for collision ionization and emission properties (Faucher-Giguère et al. 2010).

We try to quantify the impact that an increase in temperature has on our results by performing a simple test with MW3 at $z \sim 2.3$, using the STAR model. After having increased by 50 per cent the temperature in photoionized cells with $x_{H\,I} < 0.8$, we recompute the neutral fraction for CIE ($x_{H\,I,cie}$). Then, we derive the projected H I distribution function using an updated neutral fraction $x_{H\,I,new} = \min(\sqrt{x_{H\,I,old} x_{H\,I,cie}}; x_{H\,I,old}; x_{H\,I,cie})$, defined in this way to favour high ionization. Results are summarized in Fig. A4. In the bottom panel, we show the column density distribution function for the original STAR model (black solid line) and for the model at higher temperature (blue dash line). In the top panel we display the ratio of the two distributions. Additional heating produces higher ionization in the interval $\log N_{H\,I} = 18.3$–$19.8 \, \mathrm{cm}^{-2}$, yielding a
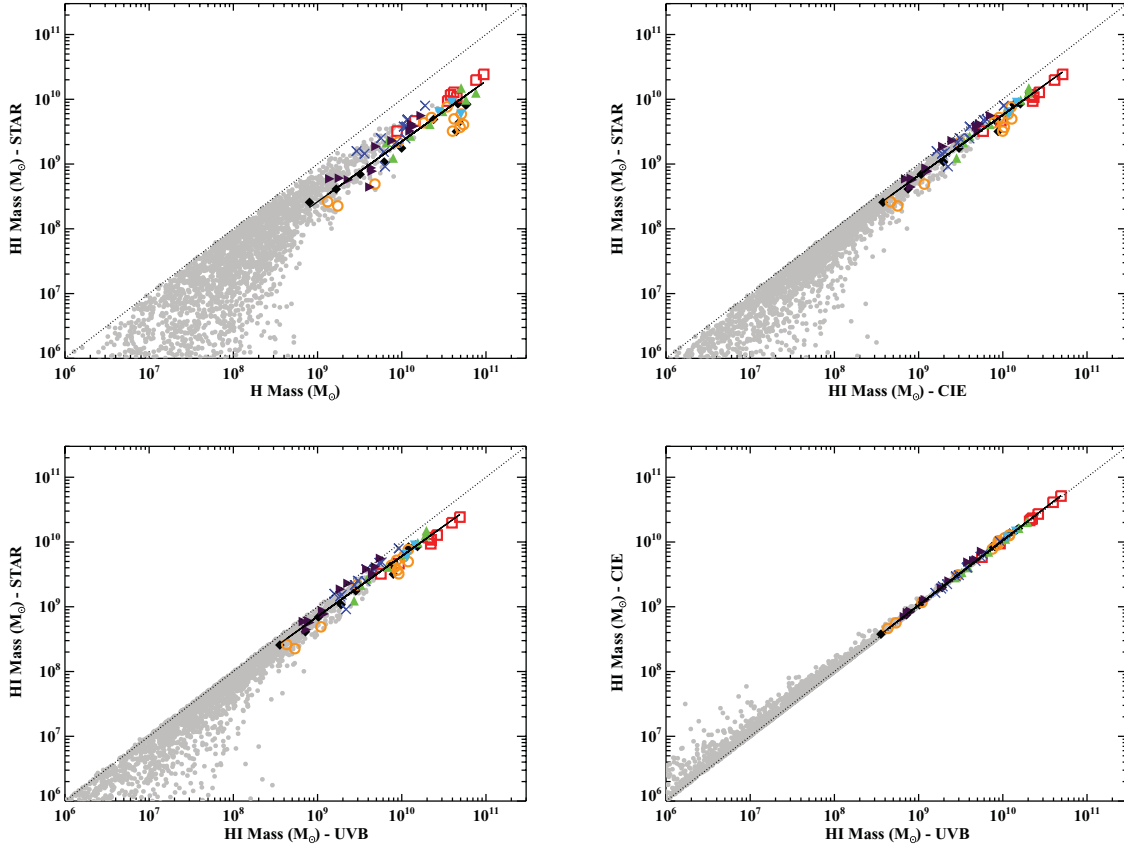
Figure A3. Comparison of the neutral and total hydrogen mass enclosed within $R_{\rm vir}$ for the three RT models. In each panel, main haloes are colour coded as in Fig. 1. Satellites are displayed with grey circles. Dotted lines indicate a 1:1 relation, while dashed lines show a robust linear regressions for the main haloes only. In massive galaxies, only 20–25 per cent of the total hydrogen mass is neutral, with local sources and electron collisions being responsible for most of the ionization. The UVB has no effects on the neutral mass due to nearly complete self-shielding at the highest hydrogen volume densities.
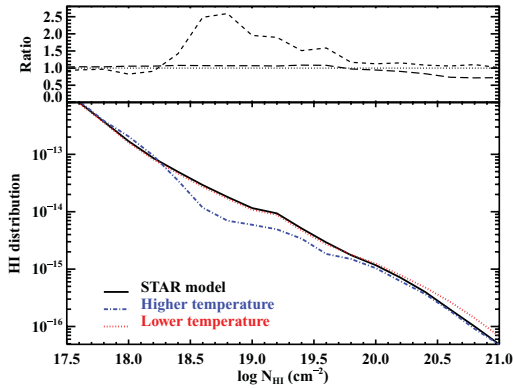
Figure A4. Bottom panel: column density distribution function for the STAR model (black solid line), for a model with increased gas temperature (blue dashed line) and a model at lower temperature (red dotted line). Top panel: the ratio of the distribution for increased (decreased) temperature and the STAR model is displayed with a dashed (long dashed) line. Identity between the models is marked with a dotted line. Additional heating produces higher ionization in the interval $\log N_{\rm H\,I} = 18.3$–$19.8\,{\rm cm}^{-2}$, but the covering factor of optically thick gas is nearly unperturbed, owing to a rising $f(N_{\rm H\,I}, X)$ towards lower column densities. Lowering the temperature, little differences (less than a factor of 2) are visible since the UV radiation is the dominant source of ionization in unshielded regions, while shielded regions are already cold and neutral.

flatter $f(N_{\rm H\,I}, X)$. The cross-section of SLLSs is reduced by a factor of ~ 2, increasing even further the discrepancy with observations. The DLA cross-section is decreased by no more than 20 per cent and the effect on the cumulative covering factor of LLSs is small, owing to a rising $f(N_{\rm H\,I}, X)$ towards lower column densities. This simple test suggests that a moderate increase in temperature does not have a major impact on the result presented in this paper. Conversely, if cooling is more effective than what assumed in these simulations, hydrogen is less ionized by electron collisions. In Fig. A4, we show the hydrogen distribution (red dotted line) for a model in which the temperature is reduced to 50 per cent of its value in all the cells with $n_{\rm H} > 0.05\,{\rm cm}^{-2}$, i.e. partially or totally shielded. Gas below this density is highly photoionized and the ionization state is insensitive to the gas temperature. In this case, the ionization fraction is assumed to be $x_{\rm H\,I,new} = x_{\rm H\,I,old}$ for $n_{\rm H} < 0.05\,{\rm cm}^{-2}$ and $x_{\rm H\,I,new} = x_{\rm H\,I,cie}$ for $n_{\rm H} \geq 0.05\,{\rm cm}^{-2}$. Since gas above $n_{\rm H} \sim 0.05\,{\rm cm}^{-2}$ is already at low temperature, little difference (less than a factor of 2) is visible in the final H I distribution.

APPENDIX B: TABLES

Table B1. Summary of the baryonic and dark matter properties of the seven galaxies used in this study.

Redshift	MW1	MW2	MW3	MW4	MW5	MW8	MW9	MW1	MW2	MW3	MW4	MW5	MW8	MW9
	Virial mass (10^{11} M$_\odot$)							Virial radius (kpc)						
4.00	0.11	0.91	0.29	1.26	6.09	0.62	0.19	15.00	30.00	20.50	33.50	56.50	26.25	18.25
3.55	0.26	1.51	0.37	1.73	9.58	0.73	0.33	21.75	39.00	24.50	41.00	72.25	30.50	23.50
3.17	0.53	3.14	1.06	2.73	11.70	1.13	0.48	29.75	57.00	38.00	51.75	84.25	38.50	29.00
2.85	1.13	4.19	1.62	4.23	–	1.52	1.22	41.75	64.50	47.25	64.50	–	46.00	42.75
2.57	1.68	5.04	2.66	5.81	–	1.80	1.31	51.50	74.25	60.50	77.25	–	52.25	47.00
2.33	3.34	5.92	3.55	9.21	–	2.31	1.29	68.75	83.25	70.50	96.50	–	60.75	50.00
2.13	6.49	7.64	5.73	12.01	–	2.54	1.32	91.25	97.25	88.00	112.25	–	66.75	53.50
1.94	8.83	9.12	8.06	15.17	–	2.79	1.75	106.75	108.75	104.00	129.00	–	73.00	62.50
1.78	10.20	–	9.27	–	–	2.85	2.87	118.50	–	114.75	–	–	77.75	77.50
1.63	11.47	–	11.81	–	–	2.85	3.00	129.50	–	131.25	–	–	81.75	82.75
1.50	13.03	–	12.59	–	–	2.86	3.39	141.75	–	140.25	–	–	85.50	90.50
1.38	13.69	–	13.52	–	–	5.10	4.31	150.50	–	150.00	–	–	108.00	102.25
	Star formation rate (M$_\odot$ yr^{-1})							Stellar mass (10^{10} M$_\odot$)						
4.00	0.94	37.79	3.74	21.19	49.49	4.23	2.02	0.03	1.62	0.20	1.22	6.25	0.63	0.13
3.55	4.61	58.25	2.70	28.66	140.61	5.70	2.88	0.17	3.12	0.26	1.69	10.66	0.80	0.28
3.17	4.42	148.23	14.44	15.77	76.14	4.48	3.12	0.31	7.52	0.86	2.79	13.32	1.31	0.41
2.85	14.76	157.48	5.54	38.11	–	40.65	6.98	0.87	11.16	1.14	3.91	–	1.70	1.27
2.57	9.65	104.36	25.95	48.22	–	4.73	8.28	1.14	14.24	1.76	4.94	–	1.85	1.48
2.33	35.71	99.77	30.49	44.21	–	7.16	7.14	2.25	17.51	2.63	8.69	–	2.21	1.61
2.13	52.47	143.23	46.77	73.75	–	5.03	2.45	5.54	22.26	4.34	12.07	–	2.42	1.70
1.94	62.03	187.86	71.46	77.22	–	13.63	2.38	8.19	27.48	7.14	15.38	–	2.96	1.94
1.78	108.52	–	84.00	–	–	6.26	5.44	10.57	–	8.90	–	–	3.04	3.19
1.63	20.71	–	50.19	–	–	3.42	17.68	11.43	–	12.10	–	–	3.09	3.46
1.50	25.65	–	39.10	–	–	3.61	4.25	12.80	–	13.15	–	–	3.20	3.98
1.38	23.19	–	45.64	–	–	5.31	6.77	14.19	–	14.38	–	–	6.42	5.49
	Dark matter mass (10^{11} M$_\odot$)							Hydrogen mass (10^{10} M$_\odot$)						
4.00	0.10	0.66	0.25	1.07	5.19	0.53	0.16	0.080	0.879	0.132	0.663	2.761	0.291	0.140
3.55	0.23	1.06	0.33	1.49	8.13	0.61	0.29	0.167	1.431	0.174	0.791	3.921	0.361	0.182
3.17	0.46	2.04	0.93	2.32	9.86	0.95	0.41	0.320	3.517	0.480	1.265	5.041	0.566	0.232
2.85	0.98	2.66	1.42	3.62	–	1.29	1.05	0.619	4.138	0.906	2.156	–	0.635	0.494
2.57	1.46	3.24	2.30	5.01	–	1.55	1.11	0.996	3.803	1.823	3.057	–	0.614	0.447
2.33	2.88	3.72	3.07	7.84	–	2.01	1.09	2.390	4.502	2.249	5.065	–	0.862	0.419
2.13	5.54	4.65	4.95	10.23	–	2.20	1.11	3.990	7.622	3.453	5.778	–	1.058	0.437
1.94	7.54	5.43	6.92	12.88	–	2.39	1.48	4.648	9.399	4.181	7.590	–	1.109	0.767
1.78	8.70	–	7.98	–	–	2.43	2.42	4.449	–	4.009	–	–	1.136	1.312
1.63	9.85	–	10.11	–	–	2.42	2.53	4.791	–	4.926	–	–	1.162	1.255
1.50	11.17	–	10.77	–	–	2.42	2.86	5.763	–	5.097	–	–	1.177	1.382
1.38	11.68	–	11.54	–	–	4.27	3.59	5.920	–	5.413	–	–	1.895	1.701

Table B2. Summary of the neutral gas properties of the seven galaxies used in this study.

Redshift	MW1	MW2	MW3	MW4	MW5	MW8	MW9	MW1	MW2	MW3	MW4	MW5	MW8	MW9
	H I mass (10^{10} M$_\odot$)–UVB							H I Mass (10^{10} M$_\odot$)–STAR						
4.00	0.036	0.565	0.043	0.326	1.124	0.158	0.068	0.026	0.322	0.026	0.215	0.637	0.158	0.059
3.55	0.072	0.926	0.054	0.271	1.423	0.186	0.079	0.041	0.464	0.023	0.124	0.891	0.145	0.060
3.17	0.103	2.210	0.110	0.389	1.147	0.302	0.081	0.069	0.940	0.049	0.270	0.589	0.252	0.057
2.85	0.190	2.631	0.295	0.676	–	0.218	0.187	0.108	1.281	0.217	0.410	–	0.091	0.187
2.57	0.285	2.085	0.751	1.071	–	0.194	0.120	0.176	1.169	0.434	0.651	–	0.153	0.078
2.33	0.805	2.230	0.869	1.963	–	0.277	0.074	0.490	1.081	0.522	1.494	–	0.216	0.044
2.12	1.454	3.950	1.195	1.548	–	0.379	0.115	0.901	1.983	0.763	0.977	–	0.379	0.087
1.94	1.537	4.902	1.193	1.902	–	0.372	0.230	0.850	2.427	0.498	1.252	–	0.248	0.230
1.78	0.794	–	0.922	–	–	0.476	0.488	0.318	–	0.323	–	–	0.379	0.415
1.63	0.745	–	0.885	–	–	0.544	0.452	0.441	–	0.369	–	–	0.478	0.324
1.50	1.207	–	1.048	–	–	0.579	0.385	0.810	–	0.591	–	–	0.487	0.382
1.38	1.202	–	0.818	–	–	0.922	0.560	0.839	–	0.407	–	–	0.798	0.559

Table B3. Average DLA cross-sections within R_{vir} for all the haloes at all redshifts.

Halo Mass (M$_\odot$)	Total cross-section (kpc^2)		Central galaxy only (kpc^2)	
	Mean	Standard deviation	Mean	Standard deviation
3.6×10^{10}	46	24	30	21
8.6×10^{10}	138	93	82	58
1.6×10^{11}	209	170	104	86
2.6×10^{11}	312	225	182	145
3.9×10^{11}	723	396	251	184
6.8×10^{11}	623	498	215	252
1.1×10^{12}	522	252	216	163

Table B4. DLAs and LLSs covering factor from the STAR model. The top table lists values for the streams and galaxies, while the bottom one lists values for the streams alone.

Redshift	MW1	MW2	MW3	MW4	MW5	MW8	MW9	MW1	MW2	MW3	MW4	MW5	MW8	MW9
	DLA covering factor in per cent at $2R_{vir}$ (streams and galaxies)							LLS covering factor in per cent at $2R_{vir}$ (streams and galaxies)						
4.00	1.78	3.47	1.09	2.68	1.42	1.57	1.55	10.80	12.54	6.70	12.19	12.43	11.20	11.72
3.55	1.03	3.65	0.79	1.35	1.22	0.85	1.02	7.92	9.02	6.24	8.56	11.11	9.35	8.27
3.17	0.94	1.95	0.60	1.76	1.08	0.88	0.76	8.03	5.08	2.35	13.12	8.68	4.71	6.41
2.85	0.79	2.65	1.42	1.86	–	0.28	0.62	4.88	5.81	10.41	11.03	–	1.69	5.36
2.57	1.12	2.84	0.88	1.60	–	0.31	0.29	10.25	7.13	6.52	10.30	–	4.08	2.98
2.33	1.37	2.22	1.19	1.54	–	0.31	0.49	9.87	6.89	8.92	9.11	–	3.80	4.72
2.12	1.07	2.23	0.75	1.00	–	0.59	0.82	6.64	6.79	5.74	7.03	–	4.03	5.06
1.94	0.62	2.07	0.38	0.72	–	0.28	0.73	4.91	6.85	3.19	4.85	–	2.03	5.09
1.78	0.23	–	0.17	–	–	0.25	0.66	1.98	–	1.83	–	–	1.63	4.20
1.63	0.36	–	0.21	–	–	0.51	0.30	4.25	–	2.84	–	–	2.46	2.59
1.50	0.35	–	0.27	–	–	0.51	0.35	2.97	–	2.82	–	–	2.87	5.04
1.38	0.26	–	0.32	–	–	0.48	0.44	1.84	–	2.40	–	–	1.66	4.21
	DLA covering factor in per cent at $2R_{vir}$ (streams only)							LLS covering factor in per cent at $2R_{vir}$ (streams only)						
4.00	1.47	2.94	0.82	2.41	0.77	0.95	1.29	10.55	12.05	6.30	11.63	11.82	10.60	11.42
3.55	0.70	3.02	0.52	0.98	0.68	0.34	0.74	7.46	8.28	5.78	7.95	10.41	8.75	7.83
3.17	0.62	1.48	0.36	1.28	0.39	0.33	0.57	7.51	4.41	1.87	12.45	7.85	3.96	6.09
2.85	0.41	2.24	1.00	1.27	–	0.05	0.37	4.17	5.23	9.67	10.24	–	1.13	4.85
2.57	0.69	2.40	0.32	0.91	–	0.03	0.15	9.73	6.58	5.51	9.34	–	3.40	2.49
2.33	0.70	1.89	0.54	0.95	–	0.08	0.15	8.99	6.42	7.88	8.29	–	3.13	4.04
2.12	0.50	1.83	0.32	0.50	–	0.15	0.26	5.66	6.31	5.10	6.26	–	3.31	4.21
1.94	0.22	1.75	0.16	0.44	–	0.03	0.37	4.23	6.41	2.60	4.38	–	1.56	4.50
1.78	0.07	–	0.04	–	–	0.03	0.34	1.62	–	1.44	–	–	1.02	3.42
1.63	0.08	–	0.07	–	–	0.24	0.05	3.64	–	2.47	–	–	1.98	2.04
1.50	0.07	–	0.07	–	–	0.15	0.12	2.30	–	2.31	–	–	2.22	4.30
1.38	0.02	–	0.14	–	–	0.28	0.24	1.34	–	1.97	–	–	1.19	3.64

Table B5. $f(N_{H_I}, X)$ from the STAR model within R_{vir}, $2R_{vir}$ and in the streams alone.

log N_{H_I} (cm^{-2})	R_{vir}	$2R_{vir}$	Streams
13.5	2.22×10^{-15}	8.18×10^{-15}	8.30×10^{-15}
14.5	2.61×10^{-16}	1.73×10^{-15}	1.74×10^{-15}
15.5	2.16×10^{-17}	1.35×10^{-16}	1.36×10^{-16}
16.5	1.94×10^{-18}	5.98×10^{-18}	6.06×10^{-18}
17.5	9.10×10^{-20}	3.03×10^{-19}	3.00×10^{-19}
18.5	2.91×10^{-21}	9.72×10^{-21}	9.52×10^{-21}
19.5	2.64×10^{-22}	7.70×10^{-22}	7.58×10^{-22}
20.5	2.65×10^{-23}	6.43×10^{-23}	5.46×10^{-23}
21.5	1.38×10^{-24}	2.55×10^{-24}	1.17×10^{-24}
22.5	7.80×10^{-27}	1.26×10^{-26}	1.78×10^{-28}
23.5	9.92×10^{-30}	1.24×10^{-29}	–
24.5	4.08×10^{-32}	4.08×10^{-32}	–

Table B6. Equivalent widths for stacked Lyα and Si IIλ1260 absorption profiles at different impact parameters.

Radius (kpc)	W(Lyα) $z \sim 2.3$ (Å)	W(Si II) $z \sim 2.3$ (Å)	W(Lyα) $z \sim 3.2$ (Å)	W(Si II) $z \sim 3.2$ (Å)	W(Si II) enriched (Å)
0–10	7.753	0.295	6.454	0.200	0.329
10–41	2.053	0.050	1.300	0.023	0.082
41–82	0.875	0.010	0.578	0.003	0.023
82–125	0.679	0.006	0.390	0.001	0.014
125–200	0.465	0.003	–	–	0.006

Table B7. Geometric mean of the x_{H_1} versus n_H relation from the UVB model.

$\log n_H$ (cm^{-3})	$\log x_{H_1}$	$\log n_H$ (cm^{-3})	$\log x_{H_1}$	$\log n_H$ (cm^{-3})	$\log x_{H_1}$	$\log n_H$ (cm^{-3})	$\log x_{H_1}$
−4.750	−5.512E+00	−2.950	−3.959E+00	−1.150	−2.522E−01	0.650	−5.050E−03
−4.650	−5.554E+00	−2.850	−3.739E+00	−1.050	−2.097E−01	0.750	−4.217E−03
−4.550	−5.496E+00	−2.750	−3.506E+00	−0.950	−1.192E−01	0.850	−3.391E−03
−4.450	−5.475E+00	−2.650	−3.279E+00	−0.850	−1.012E−01	0.950	−2.770E−03
−4.350	−5.445E+00	−2.550	−3.054E+00	−0.750	−8.607E−02	1.050	−2.307E−03
−4.250	−5.448E+00	−2.450	−2.820E+00	−0.650	−7.321E−02	1.150	−1.850E−03
−4.150	−5.401E+00	−2.350	−2.579E+00	−0.550	−6.224E−02	1.250	−1.530E−03
−4.050	−5.360E+00	−2.250	−2.323E+00	−0.450	−5.388E−02	1.350	−1.256E−03
−3.950	−5.293E+00	−2.150	−2.050E+00	−0.350	−4.524E−02	1.450	−9.829E−04
−3.850	−5.212E+00	−2.050	−1.764E+00	−0.250	−3.801E−02	1.550	−6.561E−04
−3.750	−5.128E+00	−1.950	−1.432E+00	−0.150	−3.123E−02	1.650	−4.945E−04
−3.650	−5.020E+00	−1.850	−1.098E+00	−0.050	−2.520E−02	1.750	−3.221E−04
−3.550	−4.906E+00	−1.750	−8.472E−01	0.050	−2.037E−02	1.850	−2.372E−04
−3.450	−4.775E+00	−1.650	−6.688E−01	0.150	−1.579E−02	1.950	−1.611E−04
−3.350	−4.631E+00	−1.550	−5.363E−01	0.250	−1.247E−02	2.050	−1.826E−04
−3.250	−4.479E+00	−1.450	−4.328E−01	0.350	−1.015E−02	2.150	−8.635E−05
−3.150	−4.321E+00	−1.350	−3.546E−01	0.450	−7.837E−03		
−3.050	−4.155E+00	−1.250	−2.978E−01	0.550	−6.192E−03		

This paper has been typeset from a TEX/LATEX file prepared by the author.