On the evolution of the H I column density distribution in cosmological simulations

Alireza Rahmati,1⋆ Andreas H. Pawlik,2 Milan Raičević1 and Joop Schaye1

1Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands
2Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85748 Garching, Germany

Accepted 2013 January 8. Received 2012 December 3; in original form 2012 October 29

ABSTRACT
We use a set of cosmological simulations combined with radiative transfer calculations to investigate the distribution of neutral hydrogen in the post-reionization Universe. We assess the contributions from the metagalactic ionizing background, collisional ionization and diffuse recombination radiation to the total ionization rate at redshifts $z = 0–5$. We find that the densities above which hydrogen self-shielding becomes important are consistent with analytic calculations and previous work. However, because of diffuse recombination radiation, whose intensity peaks at the same density, the transition between highly ionized and self-shielded regions is smoother than what is usually assumed. We provide fitting functions to the simulated photoionization rate as a function of density and show that post-processing simulations with the fitted rates yield results that are in excellent agreement with the original radiative transfer calculations. The predicted neutral hydrogen column density distributions agree very well with the observations. In particular, the simulations reproduce the remarkable lack of evolution in the column density distribution of Lyman limit and weak damped Lyα systems below $z = 3$. The evolution of the low column density end is affected by the increasing importance of collisional ionization with decreasing redshift. On the other hand, the simulations predict the abundance of strong damped Lyα systems to broadly track the cosmic star formation rate density.

Key words: radiative transfer – methods: numerical – galaxies: evolution – galaxies: formation – galaxies: high-redshift – intergalactic medium.

1 INTRODUCTION
A substantial fraction of the interstellar medium (ISM) in galaxies consists of atomic hydrogen. This makes studying the distribution of neutral hydrogen (H i) and its evolution crucial for our understanding of various aspects of star formation. In the local universe, the H i content of galaxies is measured through 21 cm observations, but at higher redshifts this will not be possible until the advent of significantly more powerful telescopes such as the Square Kilometre Array.1 However, at $z \gtrsim 6$, i.e. after reionization, the neutral gas can already be probed through the absorption signatures imprinted by the intervening H i systems on the spectra of bright background sources, such as quasars (QSO).

The early observational constraints on the H i column density distribution function (H i CDDF), from quasar absorption spectroscopy at $z \lesssim 3$, were well described by a single power law in the range $N_{\text{HI}} \sim 10^{13}–10^{21} \text{ cm}^{-2}$ (Tytler 1987). Thanks to a significant increase in the number of observed quasars and improved observational techniques, more recent studies have extended these observations to both lower and higher H i column densities and to higher redshifts (e.g. Kim et al. 2002; Péroux et al. 2005; O’Meara et al. 2007, 2012; Noterdaeme et al. 2009, 2012; Prochaska, Worseck & O’Meara 2009; Prochaska & Wolfe 2009). These studies have revealed a much more complex shape which has been described using several different power-law functions (e.g. Prochaska, O’Meara & Worseck 2010; O’Meara et al. 2012). The shape of the H i CDDF is determined by both the distribution and ionization state of hydrogen. Consequently, determining the distribution function of H i column densities requires not only accurate modelling of the cosmological distribution of gas, but also radiative transfer (RT) of ionizing photons. As a starting point, the H i CDDF can be modelled by assuming a certain gas profile and exposing it to an ambient ionizing radiation field (e.g. Petitjean, Bergeron & Puget 1992; Zheng & Miralda-Escudé 2002). Although this approach captures the effect of self-shielding, it cannot be used to calculate the detailed shape and normalization of the H i CDDF which results from the cumulative effect of large numbers of objects.

*E-mail: rahmati@strw.leidenuniv.nl
1http://www.skatelescope.org/
with different profiles, total gas contents, temperatures and sizes. Moreover, the interaction between galaxies and the circum-galactic medium through accretion and various feedback mechanisms and its impact on the overall gas distribution are not easily captured by simplified models. Therefore, it is important to complement these models with cosmological simulations that model the evolution of the large-scale structure of the Universe and the formation of galaxies.

The complexity of the RT calculation depends on the H\textsc{i} column density. At low H\textsc{i} column densities (i.e. $N_{\text{HI}} \lesssim 10^{21}$ cm$^{-2}$; corresponding to the so-called Lyman $\alpha$ forest), hydrogen is highly ionized by the metagalactic ultraviolet background radiation (hereafter UVB) and largely transparent to the ionizing radiation. For these systems, the H\textsc{i} column densities can therefore be accurately computed in the optically thin limit. At higher H\textsc{i} column densities (i.e. $N_{\text{HI}} \gtrsim 10^{23}$ cm$^{-2}$; corresponding to the so-called Lyman Limit and Damped Lyman $\alpha$ systems), the gas becomes optically thick and self-shielded. As a result, the accurate computation of the H\textsc{i} column densities in these systems requires precise RT simulations. On the other hand, at the highest H\textsc{i} column densities where the gas is fully self-shielded and the recombination rate is high, non-local RT effects are not very important and the gas remains largely neutral. At these column densities, the hydrogen ionization rate may, however, be strongly affected by the local sources of ionization (Miralda-Escudé 2005; Schaye 2006; Rahmati et al., 2013). In addition, other processes like H$_2$ formation (Schaye 2001b; Krumholz et al. 2009b; Altay et al. 2011) or mechanical feedback from young stars and/or active galactic nuclei (AGNs) (Erikal et al. 2012) can also affect the highest H\textsc{i} column densities.

Despite the importance of RT effects, most of the previous theoretical works on the H\textsc{i} column density distribution did not attempt to model RT effects in detail (e.g. Katz et al. 1996; Gardner et al. 1997, 2001; Haehnelt, Steinmetz & Rauch 1998; Cen et al. 2003; Nagamine, Springel & Hernquist 2004; Nagamine et al. 2007). Only very recent works incorporated RT, primarily to account for the attenuation of the UVB (Razoumov et al. 2006; Pontzen et al. 2008; Altay et al. 2011; Fumagalli et al. 2011; McQuinn, Oh & Faucher-Giguère 2011) and found a sharp transition between optically thin and self-shielded gas that is expected from the exponential nature of extinction.

The aforementioned studies focused mainly on redshifts $z = 2$–3, for which observational constraints are strongest, without investigating the evolution of the H\textsc{i} distribution. They found that the H\textsc{i} CDDF in current cosmological simulations is in reasonable agreement with observations in a large range of H\textsc{i} column densities. Only at the highest H\textsc{i} column densities (i.e. $N_{\text{HI}} \gtrsim 10^{21}$ cm$^{-2}$) the agreement is poor. However, it is worth noting that the interpretation of these H\textsc{i} systems is complicated due to the complex physics of the ISM and ionization by local sources. Moreover, the observational uncertainties are also larger for these rare high $N_{\text{HI}}$ systems.

In this paper, we investigate the cosmological H\textsc{i} distribution and its evolution during the last $\gtrsim 12$ billion years (i.e. $z \lesssim 5$). For this purpose, we use a set of cosmological simulations which include star formation, feedback and metal-line cooling in the presence of the UVB. These simulations are based on the Overwhelmingly Large Simulations (OWLS) presented in Schaye et al. (2010). To obtain the H\textsc{i} CDDF, we post-processed the simulations with RT, accounting for both ionizing UVB radiation and ionizing recombination radiation (RR). In contrast to previous works, we account for the impact of RR explicitly, by propagating RR photons. Using these simulations we study the evolution of the H\textsc{i} CDDF in the range of redshifts $z = 0$–5 for column densities $N_{\text{HI}} \gtrsim 10^{16}$ cm$^{-2}$. We discuss how the individual contributions from the UVB, RR and collisional ionization to the total ionization rate shape the H\textsc{i} CDDF and assess their relative importance at different redshifts.

The structure of this paper is as follows. In Section 2 we describe the details of the hydrodynamical simulations and of the RT, including the treatment of the UVB and RR. In Section 3 we present the simulated H\textsc{i} CDDF and its evolution and compare it with observations. In the same section we also discuss the contributions of different ionizing processes to the total ionization rate and provide fitting functions for the total photoionization rate as a function of density which reproduce the RT results. Finally, we conclude in Section 4.

## 2 SIMULATION TECHNIQUES

### 2.1 Hydrodynamical simulations

We use density fields from a set of cosmological simulations performed using a modified version of the smoothed particle hydrodynamics code GADGET-3 (last described in Springel 2005). The subgrid physics is identical to that used in the reference simulation of the OWLS project (Schaye et al. 2010). Star formation is pressure dependent and reproduces the observed Kennicutt–Schmidt law (Schaye & Dalla Vecchia 2008). Chemical evolution is followed using the model of Wiersma et al. (2009b), which traces the abundance evolution of 11 elements by following stellar evolution assuming a Chabrier (2003) initial mass function. Moreover, a radiative heating and cooling implementation based on Wiersma et al. (2009a) calculates cooling rates element-by-element (i.e. using the above-mentioned 11 elements) in the presence of the uniform cosmic microwave background and the UVB model given by Haardt & Madau (2001). About 40 per cent of the available kinetic energy in Type II SNe is injected in winds with initial velocity of 600 km s$^{-1}$ and a mass loading parameter $n = 2$ (Dalla Vecchia & Schaye 2008). Our tests show that varying the implementation of the kinetic feedback only changes the H\textsc{i} CDDF in the highest column densities ($N_{\text{HI}} \gtrsim 10^{21}$ cm$^{-2}$). However, the differences caused by these variations are smaller than the evolution in the H\textsc{i} CDDF and observational uncertainties (see Altay et al., in preparation).

We adopt fiducial cosmological parameters consistent with the most recent Wilkinson Microwave Anisotropy Probe (WMAP) 7-year results: $\Omega_m = 0.272$, $\Omega_b = 0.0455$, $\Omega_{\Lambda} = 0.728$, $\sigma_8 = 0.81$, $n_s = 0.967$ and $h = 0.704$ (Komatsu et al. 2011). We also use cosmological simulations from the OWLS project which are performed with a cosmology consistent with WMAP 3-year values with $\Omega_m = 0.238$, $\Omega_b = 0.0418$, $\Omega_{\Lambda} = 0.762$, $\sigma_8 = 0.74$, $n_s = 0.951$ and $h = 0.73$. We use those simulations to avoid expensive resimulation with a WMAP 7-year cosmology. Instead, we correct for the difference in the cosmological parameters as explained in Appendix B.

Our simulations have box sizes in the range $L = 6.25$–100 comoving $h^{-1}$ Mpc and baryonic particle masses in the range $1.7 \times 10^9 h^{-1}$ M$_\odot$ to $7.8 \times 10^9 h^{-1}$ M$_\odot$. The suite of simulations allows us to study the dependence of our results on the box size and mass resolution. Characteristic parameters of the simulations are summarized in Table 1.

### 2.2 Radiative transfer with TRAPHIC

The RT is performed using TRAPHIC (Pawlik & Schaye 2008, 2011). TRAPHIC is an explicitly photon-conserving RT method designed
to transport radiation directly on the irregular distribution of SPH particles using its full dynamic range. Moreover, by tracing photon packets inside a discrete number of cones, the computational cost of the RT becomes independent of the number of radiation sources. TRAPHIC is therefore particularly well-suited for RT calculation in cosmological density fields with a large dynamical range in densities and large numbers of sources. In the following we briefly describe how TRAPHIC works. More details, as well as various RT tests, can be found in Pawlik & Schaye (2008, 2011).

The photon transport in TRAPHIC proceeds in two steps: the isotropic emission of photon packets with a characteristic frequency $\nu$ by source particles and their subsequent directed propagation on the irregular distribution of SPH particles. The spatial resolution of the RT is set by the number of neighbours for which we generally use the same number of SPH neighbours used for the underlying hydrodynamical simulations, i.e. $N_{\text{neigh}} = 48$.

After source particles emit photon packets isotropically to their neighbours, the photon packets travel along their propagation directions to other neighbouring SPH particles which are inside their transmission cones. Transmission cones are regular cones with an opening solid angle $4\pi/N_{\text{RC}}$ and are centred on the propagation direction. The parameter $N_{\text{RC}}$ sets the angular resolution of the RT, and we adopt $N_{\text{RC}} = 64$. We demonstrate convergence of our results with the angular resolution in Appendix C. Note that the transmission cones are defined locally at the transmitting particle, and hence the angular resolution of the RT is independent of the distance from the source.

It can happen that transmission cones do not contain any neighbouring SPH particles. In this case, additional particles (virtual particles, ViPs) are placed inside the transmission cones to accomplish the photon transport. The ViPs, which enable the particle-to-particle transport of photons along any direction independent of the spatially inhomogeneous distribution of the particles, do not affect the SPH simulation and are deleted after the photon packets have been transferred.

An important feature of the RT with TRAPHIC is the merging of photon packets which guarantees the independence of the computational cost from the number of sources. Different photon packets which are received by each SPH particle are binned based on their propagation directions in $N_{\text{RC}}$ reception cones. Then, photon packets with identical frequencies that fall in the same reception cone are merged into a single photon packet with a new direction set by the weighted sum of the directions of the original photon packets. Consequently, each SPH particle holds at most $N_{\text{RC}} \times N_{\nu}$ photon packets, where $N_{\nu}$ is the number of frequency bins. We set $N_{\text{RC}} = 8$ for which our tests yield converged results.

Photon packets are transported along their propagation direction until they reach the distance they are allowed to travel within the RT time-step by the finite speed of light, i.e. $c\Delta t$. Photon packets that cross the simulation box boundaries are assumed to be lost from the computational domain. We use a time-step $\Delta t = 1 \text{ Myr} \left(\frac{\text{L}_{100}}{\text{N}_{256}} \left(\frac{\text{M}_{\odot}}{\text{h}^{-1} \text{M}_{\odot}}\right) \left(\frac{\text{kpc}}{h}\right)^{2/3}\right)$, where $N_{\text{SPH}}$ is the number of SPH particles in each dimension. We verified that our results are insensitive to the exact value of the RT time-step: values that are smaller or larger by a factor of 2 produce essentially identical results. This is mostly because we evolve the ionization balance on smaller subcycling steps, and because we iterate for the equilibrium solution, as we discuss below. At the end of each time-step the ionization states of the particles are updated based on the number of absorbed ionizing photons.

The number of ionizing photons that are absorbed during the propagation of a photon packet from one particle to its neighbour is given by $\delta N_{\text{abs},\alpha} = \delta N_{\text{abs},\nu}[1 - \exp(-\tau(v))]$ where $\delta N_{\text{abs},\nu}$ and $\tau(v)$ are, respectively, the initial number of ionizing photons in the photon packet with frequency $\nu$ and the total optical depth of all the absorbing species. In this work we mainly consider hydrogen ionization, but in general the total optical depth is the sum $\tau(v) = \sum_{\alpha} \tau_{\alpha}(v)$ of the optical depth of each absorbing species (i.e. $\alpha \in \{ \text{H}, \text{He}, \text{He}^+ \}$). Assuming that neighbouring SPH particles have similar densities, we approximate the optical depth of each species using $\tau_{\alpha}(v) = \sigma_{\alpha}(v)n_{\nu}d_{\text{abs}}$, where $n_{\nu}$ is the number density of species, $d_{\text{abs}}$ is the absorption distance between the SPH particle and its neighbour and $\sigma_{\alpha}(v)$ is the absorption cross-section (Verner et al. 1996). Note that ViPs are deleted after each transmission, and hence the photons they absorb need to be distributed among their SPH neighbours. However, in order to decrease the amount of smoothing associated with this redistribution of photons, ViPs are assigned only five (instead of 48) SPH neighbours. We demonstrate convergence of our results with the number of ViP neighbours in Appendix C.

At the end of each RT time-step, every SPH particle has a total number of ionizing photons that have been absorbed by each species, $\Delta N_{\text{abs},\alpha}(v)$. This number is used in order to calculate the
photoionization rate of every species for that SPH particle. For instance, the hydrogen photoionization rate is given by

\[ \Gamma_{\text{H}}(\nu) = \frac{c n_{\text{e}}}{\nu} \sum_i \Delta N_{\text{abs},i}(\nu), \]

where \( N_{\text{H}} \) is the total number of hydrogen atoms inside the SPH particle and \( n_{\text{e}} \equiv n_{\text{e},i}/n_{\text{H}} \) is the hydrogen neutral fraction.

Once the photoionization rate is known, the evolution of the ionization state is calculated. For instance, the equation which governs the ionization state of hydrogen is

\[ \frac{d n_{\text{H}}}{dt} = \alpha_{\text{H}} n_{\text{e}} (1 - n_{\text{H}}) - \frac{1}{\tau_{\text{rec}}}, \]

where \( n_{\text{e}} \) is the free electron number density, \( \tau_{\text{rec}} \) is the recombination time-scale, and \( \alpha_{\text{H}} \) is the H\textsuperscript{+} ionization rate. The differential equations which govern the ionization balance (e.g. equation 2) are solved using a subcycling time-step, \( \Delta t = \min(\tau_{\text{rec}}, \Delta t) \) where \( \tau_{\text{rec}} \equiv \tau_{\text{ion}} + \tau_{\text{rec}} \), and \( f \) is a dimensionless factor which controls the integration accuracy (we set it to 10\textsuperscript{-3}). The subcycling scheme allows the RT time-step to be chosen independently of the photoionization and recombination time-scales without compromising the accuracy of the ionization state calculations.\(^2\)

We employ separate frequency bins to transport UVB and RR photons. Because the propagation directions of photons in different frequency bins are merged separately, this allows us to individually track the individual radiation components, i.e. UVB and RR, and to compute their contributions to the total photoionization rate. The implementation of the UVB and RR is described in Sections 2.3 and 2.4.

At the start of the RT, the hydrogen is assumed to be neutral. In addition, we use a common simplification (e.g. Faucher-Giguère et al. 2009; McQuinn & Switzer 2010; Altay et al. 2011) by assuming a hydrogen mass fraction of unity, i.e. we ignore helium (only for the RT). To calculate recombination and collisional ionizations rates, we set, in post-processing, the temperatures of star-forming gas particles with densities \( n_{\text{H}} > 0.1 \text{ cm}^{-3} \) to \( T_{\text{ISM}} = 10^4 \text{ K} \), which is typical of the observed warm-neutral phase of the ISM. This is needed because in our hydrodynamical simulations the star-forming gas particles follow a polytropic equation of state which defines their effective temperatures. These temperatures are only a measure of the imposed pressure and do not represent physical temperatures (see Schaye & Dalla Vecchia 2008). To speed up convergence, the hydrogen at low densities (i.e. \( n_{\text{H}} < 10^{-3} \text{ cm}^{-3} \)) or high temperatures (i.e. \( T > 10^5 \text{ K} \)) is assumed to be in ionization equilibrium with the UVB and collisional ionization (see Appendix A2). Typically, the neutral fraction of the box and the resulting H\textsuperscript{+} CDDF do not evolve after 2–3 light-crossing times (the light-crossing time for the extended box with \( L_{\text{box}} = 6.25 \text{ comoving h}^{-1} \text{ Mpc} \approx 7.5 \text{ Myr at } z = 3 \)).

### 2.3 Ionizing background radiation

Although our hydrodynamical simulations are performed using periodic boundary conditions, we use absorbing boundary conditions for the RT. This is necessary because our box size is much smaller than the mean-free-path of ionizing photons. We simulate the ionizing background radiation as plane-parallel radiation entering the simulation box from its sides. At the beginning of each RT step, we generate a large number of photon packets, \( N_{\text{bg}} \), on the nodes of a regular grid at each side of the simulation box and set their propagation directions perpendicular to the sides. The number of photon packets is chosen to obtain converged results. Furthermore, to avoid numerical artefacts close to the edges of the box, we use the periodicity of our simulations to extend the simulation box by the typical size of the region where we generate the background radiation (i.e. 2 per cent of the box size from each side). These extended regions are excluded from the analysis, thereby removing the artefacts without losing any information contained in the original simulation box.

The photon content of each packet is normalized such that in the absence of any absorption (i.e. assuming the optically thin limit), the total photon density of the box corresponds to the desired uniform hydrogen photoionization rate. If we assume that all the photons with frequencies higher than \( \nu_{\text{He}} \) are absorbed by helium, then the hydrogen photoionization rate can be written as

\[ \Gamma_{\text{UVB}} = \int_{\nu_{\text{He}}}^{\nu_{\text{bg}}} \int_{\lambda} \frac{J_{\nu}}{h \nu} \frac{\sigma_{\text{H},\nu}}{\nu} \, d \nu, \]

where \( J_{\nu} \) is the radiation intensity (in units erg cm\textsuperscript{-2} s\textsuperscript{-1} sr\textsuperscript{-1} Hz\textsuperscript{-1}), \( \nu_{\text{He}} \), and \( \nu_{\text{bg}} \) are, respectively, the frequency at the Lyman limit and the frequency at the He\textsuperscript{+} ionization edge, and \( \sigma_{\text{H},\nu} \) is the neutral hydrogen absorption cross-section for ionizing photons. In the last equation we have defined the grey absorption cross-section,

\[ \sigma_{\text{H}} = \int_{\nu_{\text{He}}}^{\nu_{\text{bg}}} \int_{\nu} \frac{J_{\nu}}{\nu} \frac{\sigma_{\text{H},\nu}}{\nu} \, d \nu. \]

The radiation intensity is related to the photon energy density, \( u_{\nu} \),

\[ J_{\nu} = \frac{u_{\nu} c}{4 \pi} = \frac{n_{\nu} \nu c}{4 \pi}, \]

where \( n_{\nu} \) is the number density of photons inside the box. Combining equations (3)–(5) yields

\[ \Gamma_{\text{UVB}} = n_{\nu} \frac{c}{4 \pi} \sigma_{\nu}, \]

where \( n_{\nu} \) is the number density of ionizing photons inside the box. The total number of ionizing photons in the box is therefore given by

\[ n_{\nu} L_{\text{box}} \approx 6 N_{\text{bg}} L_{\text{box}}, \]

where \( n_{\nu} \) is the number of ionizing photons carried by each photon packet. Now we can calculate the photon content of each packet that must be injected into the box during each step in order to achieve the desired H\textsuperscript{+} photoionization rate:

\[ n_{\nu} = \frac{\Gamma_{\text{UVB}} L_{\text{box}}^2}{6 \sigma_{\nu} N_{\text{bg}}} \]

We use the redshift-dependent UVB spectrum of Haardt & Madau (2001) to calculate \( \Gamma_{\text{UVB}} \) and \( \sigma_{\nu} \). The Haardt & Madau (2001) UVB model successfully reproduces the relative strengths of the observed metal absorption lines in the intergalactic medium (Aguirre...
et al. 2008) and has been used to calculate heating/cooling in our cosmological simulations.\footnote{Note that during the hydrodynamical simulations, photoheating from the UVB is applied to all gas particles. This ignores the self-shielding of hydrogen atoms against the UVB that occurs at densities $n_H \gtrsim 10^{-3}$ to $10^{-2}$ cm$^{-3}$. This inconsistency, which could affect both collisional ionization rates and the small-scale structure of the absorbers, has been found to have no significant impact on the simulated H$\,\text{i}$ CDDF (Pontzen et al. 2008; McQuinn & Switzer 2010; Altay et al. 2011).}

To reduce the computational cost, we treat the multi-frequency problem in the grey approximation. In other words, we transport the UVB radiation using a single frequency bin, inside which photons are absorbed using the grey cross-section $\sigma_{\text{eq}}$, defined in equation (4). Note that the grey approximation ignores the spectral hardening of the radiation field that would occur in multi-frequency simulations. In Appendix D we show the result of repeating our simulations using multiple frequency bins, and also explicitly accounting for the absorption of photons by helium. These results clearly show the expected spectral hardening. The impact of spectral hardening on the hydrogen neutral fractions and the H$\,\text{i}$ CDDF is small. However, we note that spectral hardening can change the temperature of the gas in self-shielded regions and that this effect is not captured in our simulations.

Hydrogen photoionization rates and average absorption cross-sections for UVB radiation at different shifts are listed in Table 2 for our fiducial UVB model based on Haardt & Madau (2001) together with Haardt & Madau (2012) and Faucher-Giguère et al. (2009). The photoionization rate peaks at $z \approx 2$–3 in all three UVB models and the equivalent effective photon energy\footnote{We defined the equivalent effective photon energy, $E_{\text{eq}}$, which corresponds to the absorption cross-section, $\sigma_{\text{eq}}$, as: $E_{\text{eq}} \equiv 13.6 \, \text{eV} (n_{	ext{H}}/\sigma_0)^{-1/3}$ where $\sigma_0 = 6.3 \times 10^{-18}$ cm$^2$.} of the background radiation changes only weakly with redshift, compared to the total photoionization rate.

| Redshift | UVB          | $\Gamma_{\text{UVB}}$ (s$^{-1}$) | $\sigma_{\text{eq}}$ (cm$^2$) | $E_{\text{eq}}$ (eV) | $n_{\text{H, SSB}}$ (cm$^{-3}$) |
|----------|--------------|----------------------------------|-------------------------------|---------------------|----------------------------------|
| $z = 0$  | HM01         | $8.34 \times 10^{-14}$           | $3.27 \times 10^{-18}$        | 16.9                | $1.1 \times 10^{-3}$             |
|          | HM12         | $2.27 \times 10^{-14}$           | $2.68 \times 10^{-18}$        | 18.1                | $5.1 \times 10^{-4}$             |
|          | FG09         | $3.99 \times 10^{-14}$           | $2.59 \times 10^{-18}$        | 18.3                | $7.7 \times 10^{-4}$             |
| $z = 1$  | HM01         | $7.39 \times 10^{-13}$           | $2.76 \times 10^{-18}$        | 17.9                | $5.1 \times 10^{-3}$             |
|          | HM12         | $3.42 \times 10^{-13}$           | $2.62 \times 10^{-18}$        | 18.2                | $3.3 \times 10^{-3}$             |
|          | FG09         | $3.03 \times 10^{-13}$           | $2.37 \times 10^{-18}$        | 18.8                | $3.1 \times 10^{-3}$             |
| $z = 2$  | HM01         | $1.50 \times 10^{-12}$           | $2.55 \times 10^{-18}$        | 18.3                | $8.7 \times 10^{-3}$             |
|          | HM12         | $8.98 \times 10^{-13}$           | $2.61 \times 10^{-18}$        | 18.2                | $6.1 \times 10^{-3}$             |
|          | FG09         | $6.00 \times 10^{-13}$           | $2.27 \times 10^{-18}$        | 19.1                | $5.1 \times 10^{-3}$             |
| $z = 3$  | HM01         | $1.16 \times 10^{-12}$           | $2.49 \times 10^{-18}$        | 18.5                | $7.4 \times 10^{-3}$             |
|          | HM12         | $8.74 \times 10^{-13}$           | $2.61 \times 10^{-18}$        | 18.2                | $6.0 \times 10^{-3}$             |
|          | FG09         | $5.53 \times 10^{-13}$           | $2.15 \times 10^{-18}$        | 19.5                | $5.0 \times 10^{-3}$             |
| $z = 4$  | HM01         | $7.92 \times 10^{-13}$           | $2.45 \times 10^{-18}$        | 18.6                | $5.8 \times 10^{-3}$             |
|          | HM12         | $6.14 \times 10^{-13}$           | $2.60 \times 10^{-18}$        | 18.3                | $4.7 \times 10^{-3}$             |
|          | FG09         | $4.31 \times 10^{-13}$           | $2.02 \times 10^{-18}$        | 19.9                | $4.4 \times 10^{-3}$             |
| $z = 5$  | HM01         | $5.43 \times 10^{-13}$           | $2.45 \times 10^{-18}$        | 18.6                | $4.5 \times 10^{-3}$             |
|          | HM12         | $4.57 \times 10^{-13}$           | $2.58 \times 10^{-18}$        | 18.3                | $3.9 \times 10^{-3}$             |
|          | FG09         | $3.52 \times 10^{-13}$           | $1.94 \times 10^{-18}$        | 20.1                | $4.0 \times 10^{-3}$             |

### 2.4 Recombination radiation

Photons produced by the recombination of positive ions and electrons can also ionize the gas. If the recombining gas is optically thin, RR can escape and its ionizing effects can be ignored (i.e. the so-called Case A). However, for regions in which the gas is optically thick, the proper approximation is to assume the ionizing RR is absorbed on the spot. In this case, the effective recombination rate can be approximated by excluding the transitions that produce ionizing photons (e.g. Osterbrock & Ferland 2006). This scenario is usually called Case B. A possible way to take into account the effect of RR is to use Case A recombination at low densities and Case B recombination at high densities (e.g. Altay et al. 2011; McQuinn et al. 2011), but this will be inaccurate in the transition regime.

In this work we explicitly treat the ionizing photons emitted by recombining hydrogen atoms and follow their propagation through the simulation box. This is facilitated by the fact that the computational cost of RT with \textsc{trapsh} is independent of the number of sources. This is particularly important noting that every SPH particle is potentially a source. The photon production rates of SPH particles depend on their recombination rates and the radiation is emitted isotropically once at the beginning of every RT time-step (see Raicevic et al., in preparation for full details).

We do not take into account the redshifting of the recombination photons by peculiar velocities of the emitters, or the Hubble flow. Instead, we assume that all recombination photons are monochromatic with energy 13.6 eV. In reality, recombination photons cannot travel to large cosmological distances without being redshifted to frequencies below the Lyman edge. Therefore, neglecting the cosmological redshifting of RR will result in overestimation of its photoionization rate on large scales. However, because of the small size of our simulation box, the total photoionization rate that is produced by RR on these scales remains negligible compared to the UVB photoionization rate. Consequently, the neglect of RR redshifting is not expected to affect our results.
orders of magnitudes lower than the typical hydrogen absorption cross-section for ionizing photons (Weingartner & Draine 2001). In other words, the absorption of ionizing photons by dust particles is not significant compared to the absorption by the neutral hydrogen. Consequently, as also found in cosmological simulations with ionizing radiation (Gnedin, Kravtsov & Chen 2008), dust absorption does not noticeably alter the overall distribution of ionizing photons and hydrogen neutral fractions.

The observed cut-off in the abundance of very high \(N_{\text{HI}}\) systems may be related to the conversion of atomic hydrogen into H\(_2\) (e.g. Schaye 2001b; Krumholz et al. 2009b; Prochaska & Wolfe 2009; Altay et al. 2011). Following Altay et al. (2011) and Duffy et al. (2012), we adopt an observationally driven scaling relation between gas pressure and hydrogen molecular fraction (Blitz & Rosolowsky 2006) in post-processing, which reduces the amount of observable H\(_2\) at high densities. This scaling relation is based on observations of low-redshift galaxies and may not cover the low metallicities relevant for higher redshifts. This could be an issue, since the H\(_2\)--H\(_2\) relation is known to be sensitive to the dust content and hence to the metallicity (e.g. Schaye 2001b, 2004; Krumholz, McKee & Tumlinson 2009a).

3 RESULTS

In this section we report our findings based on various RT simulations which include UVB ionizing radiation and diffuse RR from ionized gas. As we demonstrate in Section 3.3, the dependence of the photoionization rate on density obtained from our RT simulations shows a generic trend for different resolutions and box sizes. Therefore, we can use the results of RT calculations obtained from smaller boxes (e.g. \(L_{50N128}\) or \(L_{60N256}\)) which are computationally cheaper, to calculate the neutral hydrogen distribution in larger boxes. The last column of Table 1 indicates for which simulations this was done.

In the following, we will first present the predicted H\(_1\) CDDF and compare it with observations. Next we discuss other aspects of our RT results and the effects of ionization by the UVB, RR and collisional ionization on the resulting H\(_1\) distributions at different redshifts.

3.1 Comparison with observations

In Fig. 1 we compare the simulation results with a compilation of observed H\(_1\) CDDFs, after converting both to the WMAP 7-year cosmology. The data points with error bars show results from high-redshift (\(z = 1.7\)–5.5) QSO absorption line studies and the orange filled circles show the fitting function reported by Zwaan et al. (2005) based on 21 cm observations of nearby galaxies. The latter observations only probe column densities \(N_{\text{HI}} > 10^{19}\) cm\(^{-2}\).

We note that the OWLS simulations have already been shown to agree with observations by Altay et al. (2011), but only for \(z = 3\) and based on a different RT method (see Appendix C3 for a comparison). Overall, our RT results are also in good agreement with the observations. At high column densities (i.e. \(N_{\text{HI}} > 10^{17}\) cm\(^{-2}\)) the observations probe \(0 < z < 5.5\) and are consistent with each other. This implies weak or no evolution with redshift. The simulation is consistent with this remarkable observational result, predicting only weak evolution for \(10^{17}\) cm\(^{-2} \ll N_{\text{HI}} < 10^{21}\) cm\(^{-2}\) (i.e. Lyman limit systems, LLCs, and weak DLAs) especially at \(z \lesssim 3\).

The simulation predicts some variation with redshift for strong DLAs (\(N_{\text{HI}} > 10^{21}\) cm\(^{-2}\)). The abundance of strong DLAs in the
On the evolution of the H\textsubscript{i} CDDF

Figure 1. CDDF of neutral gas at different redshifts in the presence of the UVB and diffuse recombination radiation for L50N512-W3. A column density dependent amplitude correction has been applied to make the results consistent with WMAP 7-year cosmological parameters. The observational data points represent a compilation of various quasar absorption line observations at high redshifts (i.e. $z = [1.7, 5.5]$) taken from Péroux et al. (2005) with $z = [1.8, 3.5]$, O'Meara et al. (2007) with $z = [1.7, 4.5]$, Noterdaeme et al. (2009) with $z = [2.2, 5.5]$ and Prochaska & Wolfe (2009) with $z = [2.2, 5.5]$. The coloured data points in the top-left corner of the left panel are taken from Kim et al. (2002) with $z = [2.9, 3.5]$ and $z = [1.7, 2.4]$ for the yellow crosses and orange diamonds, respectively. The orange filled circles show the best fit based on the low-redshift 21 cm observations of Zwaan et al. (2005). The high column density end of the H\textsubscript{i} distribution is magnified in the right panel and for clarity only the simulated H\textsubscript{i} CDDF of redshifts $z = 1$, $3$ and $5$ are shown. The top-section of each panel shows the ratio between the H\textsubscript{i} CDDFs at different redshifts and the H\textsubscript{i} CDDF at $z = 3$. The simulation results are in reasonably good agreement with the observations and, like the observations, show only a weak evolution for Lyman limit and weak DLAs below $z = 3$.

At low column densities (i.e. $N_{\text{H} i} \lesssim 10^{17}$ cm$^{-2}$) the simulation results agree very well with the observations. This is apparent from the agreement between the simulated $f(N_{\text{H} i}, z)$ at $z = 3$ and $z = 4$, and the observed values for redshifts $2.9 < z < 3.5$ (Kim et al. 2002) which are shown by the yellow crosses in the left panel of Fig. 1. The simulated $f(N_{\text{H} i}, z)$ at lower and higher redshifts deviate from those at $z \approx 3$ showing the abundance of those systems decreases with decreasing redshift and remains nearly constant at $z \lesssim 2$. This is consistent with the Ly\textalpha forest observations at lower redshifts (Kim et al. 2002; Janknecht et al. 2006; Lehner et al. 2007; Prochaska & Wolfe 2009; Ribaudo, Lehner & Howk 2011), as illustrated with the orange diamonds which correspond to $z \approx 2$ observations, in the top-left corner of the left panel in Fig. 1.

The evolution of the H\textsubscript{i} CDDF with redshift results from a combination of the expanding Universe and the growing intensity of the UVB radiation down to redshifts $z \approx 2–3$. At low redshifts (i.e. $z \approx 0$) the intensity of the UVB radiation has dropped by more than one order of magnitude leading to higher hydrogen neutral fractions and higher H\textsubscript{i} column densities. However, as we show in Section 3.5, at lower redshifts an increasing fraction of low-density gas is shock-heated to temperatures sufficiently high to become collisionally ionized and this compensates for the weaker UVB radiation at low redshifts.

The simulated H\textsubscript{i} CDDFs at all redshifts are consistent with each other and the observations. However, as illustrated in the right panel of Fig. 1, there is a $\approx 0.2$ dex difference between the simulation results and the observations of LLS and DLAs at all redshifts. We found that the normalization of the H\textsubscript{i} CDDF in those regimes is sensitive to the adopted cosmological parameters (see Appendix B). Notably, the cosmology consistent with the WMAP 7-year results that are shown here produces a better match to the observations than a cosmology based on the WMAP 3-year results with smaller values for $\Omega_0$ and $\sigma_8$. This suggests that a higher value of $\sigma_8$ may...
explain the small discrepancy between the simulation results and the observations.

3.2 The shape of the $\text{H} \text{i} \text{ CDDF}$

The shape of the $\text{H} \text{i} \text{ CDDF}$ is determined by the distribution of hydrogen and by the different ionizing processes that set the hydrogen neutral fractions of the absorbers. One can assume that overdense hydrogen resides in self-gravitating systems that are in local hydrostatic equilibrium. Then, the typical scales of the systems can be calculated as a function of the gas density based on a Jeans scaling argument (Schaye 2001a). Assuming that absorbers have universal baryon fractions (i.e. $f_b = \frac{\Omega_b}{\Omega_m}$) and typical temperatures of $T_i \equiv (T/10^4 \text{K}) \sim 1$ (i.e. collisional ionization is unimportant), one can calculate the total hydrogen column density (Schaye 2001a):

$$N_{\text{HI}} \sim 1.6 \times 10^{21} \text{cm}^{-2} n_{\text{H}}^{1/2} T_i^{1/2} \left(\frac{f_b}{0.17}\right)^{1/2}. \quad (10)$$

Assuming that the gas is highly ionized and in ionization equilibrium with the ambient ionizing radiation field with the photoionization rate, $\Gamma_{-12} = \Gamma/10^{-12} \text{s}^{-1}$, one gets (Schaye 2001a)

$$N_{\text{HI}} \sim 2.3 \times 10^{13} \text{cm}^{-2} \left(\frac{n_{\text{H}}}{10^{-3} \text{cm}^{-3}}\right)^{3/2} \times T_i^{-0.26} \Gamma_{-12}^{-1} \left(\frac{f_b}{0.17}\right)^{1/2}. \quad (11)$$

At high densities where the gas is nearly neutral, equation (10) provides a relation between $N_{\text{HI}}$ and $n_{\text{H}}$. Equation (11), on the other hand, gives the relation for optically thin, highly ionized gas. The latter is derived assuming that the UVB photoionization is the dominant source of ionization, which is a good assumption at high redshifts and explains the relation between density and column density in Lyman forest simulations (e.g. Davé et al. 2010; Altay et al. 2011; McQuinn et al. 2011; Tepper-García et al. 2012). However, as we will show in the following sections, photoionization domination breaks down at lower redshifts where collisional ionization plays a significant role.

The column density at which hydrogen starts to be self-shielded against the UVB radiation follows from setting $\tau_{\text{HI}} = 1$:

$$N_{\text{HI,SSh}} \sim 4 \times 10^{17} \text{cm}^{-2} \left(\frac{\sigma_{\text{HI}}}{2.49 \times 10^{-18} \text{cm}^2}\right)^{-1}. \quad (12)$$

which can be used together with equation (11) to find the typical densities at which the self-shielding begins (e.g. Furlanetto et al. 2005):

$$n_{\text{HI,SSh}} \sim 6.73 \times 10^{-3} \text{cm}^{-3} \left(\frac{\sigma_{\text{HI}}}{2.49 \times 10^{-18} \text{cm}^2}\right)^{-2/3} \times T_i^{-0.17} \Gamma_{-12}^{-2/3} \left(\frac{f_b}{0.17}\right)^{-1/3}. \quad (13)$$

These relations are compared with the $n_{\text{H}_{\text{SSh}}}$-weighted total hydrogen number density as a function of $N_{\text{HI}}$ in the L06N256 simulation at $z = 3$ in the left panel of Fig. 2. The solid curve shows the median and the red (blue) shaded area represents the central 70 per cent (90 per cent) percentile. The diagonal grey solid line, which converges with the simulation results at low column densities, shows equation (11) and the steeper grey dotted line which converges with the simulation results at high column densities is based on equation (10). The agreement between the expected slopes of the $n_{\text{H}_{\text{SSh}}}$ relation and the simulations at low and high column densities confirms our initial assumption that hydrogen resides in self-gravitating systems which are close to local hydrostatic equilibrium.

As expected from equation (13), at low densities the gas is optically thin and follows the Jeans scaling relation of the highly ionized gas. At $n_{\text{H}_{\text{SSh}}} \sim 0.01 \text{cm}^{-3}$, however, the relation between density and column density starts to deviate from equation (11) and approaches that of a nearly neutral gas. Consequently, for densities above the self-shielding threshold the $\text{H} \text{i}$ column density increases rapidly over a narrow range of densities, leading to a flattening in the $n_{\text{H}_{\text{SSh}}} - N_{\text{HI}}$ relation and in the resulting $\text{H} \text{i} \text{ CDDF}$ at $N_{\text{HI}} > 10^{18} \text{cm}^{-2}$ (see Fig. 2). The results from the RT simulation deviate from the magenta dotted lines, which are obtained assuming optically thin gas, at $N_{\text{HI}} > 4 \times 10^{17} \text{cm}^{-2}$. As the dotted line in the right panel of Fig. 2 shows, in absence of self-shielding, the slope of $f(N_{\text{HI}}, z) \propto N_{\text{HI}}^{-1}$ is constant all the way up to DLAs at $f_{\text{Ly}\alpha} \approx -1.6$. However, because of self-shielding, the $\text{H} \text{i} \text{ CDDF}$ flattens to $f_{\text{LSS}} \approx -1.1$ at $10^{18} \text{cm}^{-2} \lesssim N_{\text{HI}} \lesssim 10^{20} \text{cm}^{-2}$ in the RT simulation (solid curve). These predicted slopes are in excellent agreement with the latest observational constraints of $f_{\text{Ly}\alpha} \lesssim -1.6$ for $10^{17} \text{cm}^{-2} < N_{\text{HI}} < 10^{19} \text{cm}^{-2}$ for $f_{\text{LSS}} \approx -1$ in the LLS regime (O’Meara et al. 2012). We also note that $f_{\text{Ly}\alpha} \lesssim -1.6$ is predicted to be almost the same for all redshifts, which agrees well with observations (Janknecht et al. 2006; Lehner et al. 2007; Ribaudo et al. 2011).

At densities $n_{\text{H}_{\text{SSh}}} > 0.1 \text{cm}^{-3}$ the gas is nearly neutral and the Jeans scaling in equation (10) controls the $n_{\text{H}_{\text{SSh}}} - N_{\text{HI}}$ relation. Consequently, the rate at which $N_{\text{HI}}$ responds to changes in $n_{\text{H}_{\text{SSh}}}$ slows down, causing a steepening in the resulting $f(N_{\text{HI}}, z)$ in the DLAs range (i.e. $N_{\text{HI}} > 10^{19} \text{cm}^{-2}$). However, as the thick solid curve in the right panel of Fig. 2 illustrates, the slope of $f(N_{\text{HI}}, z)$ remains constant for $N_{\text{HI}} = 10^{21} - 10^{22} \text{cm}^{-2}$. This is in contrast with observed trends indicating a sharp cut-off at $N_{\text{HI}} > 3 \times 10^{21} \text{cm}^{-2}$ (Prochaska et al. 2010; O’Meara et al. 2012; but see Noterdaeme et al. 2012). At those column densities a large fraction of hydrogen is expected to form $\text{H}_2$ molecules and be absent from $\text{H} \text{i}$ observations (Schaye 2001b; Krumholz et al. 2009b; Altay et al. 2011). As the thin solid curve in the right panel of Fig. 2 shows, accounting for $\text{H}_2$ using the empirical relation between $\text{H}_2$ fraction and pressure, based on $z = 0$ observations (Blitz & Rosolowsky 2006), does reproduce a sharp cut-off. If the observed relation does not cut off (Noterdaeme et al. 2012), then this may imply that $\text{H}_2$ fractions are lower at $z = 3$ than at $z = 0$. We also note that the ionizing effect of local sources (Rahmati et al., 2013), increasing the efficiency of stellar feedback, e.g. by using a top-heavy IMF, and AGN feedback can also affect these high $\text{H} \text{i}$ column densities (Altay et al., in preparation).

To the first order, one can mimic the effect of RT by assuming gas with $n_{\text{H}_{\text{SSh}}} < n_{\text{LSS}}$ to be optically thin (i.e. Case A recombination) and gas with $n_{\text{H}_{\text{SSh}}} > n_{\text{H}_{\text{LSS}}}$ to be fully neutral. Simulations with three different self-shielding density thresholds are shown in Fig. 2. The dot–dashed, dot–dot–dashed and long dashed curves correspond to $n_{\text{H}_{\text{SSh}}} = 10^{-4}, 10^{-3}$ and $10^{-2} \text{cm}^{-3}$, respectively. Although all of these simulations predict the flattening of $f(N_{\text{HI}}, z)$, they produce a transition between optically thin and neutral gas that is too steep.

7 One should note that the above-mentioned Jeans argument provides an order of magnitude calculation due to its simplifying assumptions (e.g. uniform density, universal baryon fraction, etc.). Although we may expect the predicted scaling relations to be correct, the very close agreement of the normalization with the simulations at low densities is coincidental. As the steeper grey dotted line which is based on equation (10) shows, the simulated $N_{\text{HI}}$ for a given $n_{\text{H}_{\text{SSh}}}$ is $\approx 0.5$ dex higher than implied by the Jeans scaling for the nearly neutral case (i.e. steep, grey solid line).
As explained in more detail (equation 10; diagonal solid line) and for neutral gas (equation 11; steeper dotted line). A second solid line with the same slope expected from equation (11) but a different normalization is illustrated by the second solid line which is identical to the dotted line but shifted by 0.5 dex to higher \( n_{\text{H}} \). The pink and blue shaded areas in the right panel indicate the 70 per cent and 99 per cent scatter, respectively, while the solid curves show the median for the RT result. All the other curves are also medians. This shows that the \( n_{\text{H}} \)-\( N_{\text{H}} \) relationship can be explained by the Jeans scaling and that the flattening in the CDDF is due to self-shielding. Right: \( \text{H}_1 \) CDDF in the presence of the UVB and diffuse RR for simulation L06N256. Simulations shown with different curves are identical to the left panel. In addition, the effect of \( \text{H}_2 \) formation is shown by the green solid curve which deviates from the brown solid curve at \( N_{\text{H}} \gtrsim 3 \times 10^{21} \text{ cm}^{-2} \). Finally the red dashed curve, which is indistinguishable from the brown solid curve, shows the result of assuming the median of the photoionization rate profile of the RT results to calculate the neutral fractions (see Section 3.3 and Appendix A1). The top-section in the right panel shows the ratio between different \( \text{H}_1 \) CDDFs and the one resulting from the RT simulations.

In contrast, the RT results show a transition between highly ionized and highly neutral gas that is more gradual, as observed.

### 3.3 Photoionization rate as a function of density

Fig. 3 illustrates the RT results for neutral fractions and photoionization rates as a function of density in the presence of UVB radiation and diffuse RR for the L06N128, L06N256 and L12N256 simulations at \( z = 3 \). For comparison, the results for the optically thin limit are shown by the green dotted curves. The sharp transition between highly ionized and neutral gas and its deviation from the optically thin case are evident in the left panel. This transition can also be seen in the photoionization rate (right panel) which drops at \( n_{\text{H}} \gtrsim 0.01 \text{ cm}^{-3} \), consistent with equation (13) and previous studies (Tajiri & Umemura 1998; Razoumov et al. 2006; Faucher-Giguère et al. 2010; Nagamine, Choi & Yajima 2010; Altay et al. 2011; Fumagalli et al. 2011).

The medians and the scatter around them are insensitive to the resolution of the underlying simulation and to the box size. This suggests that one can use the photoionization rate profile obtained from the RT simulations for calculating the hydrogen neutral fractions in other simulations for which no RT has been performed.

Moreover, as we show in Section 3.5, the total photoionization rate as a function of the hydrogen number density has the same shape at different redshifts. This shape can be characterized by three features: (i) a knee at densities around the self-shielding density threshold, (ii) a relatively steep fall-off at densities higher than the self-shielding threshold and (iii) a flattening in the fall-off after the photoionization rate has dropped by \( \sim 2 \text{ dex} \) from its maximum value which is caused by the RR photoionization. These features are captured by the following fitting formula:

\[
\frac{\Gamma_{\text{Phot}}}{\Gamma_{\text{UVB}}} = 0.98 \left[ 1 + \left( \frac{n_{\text{H}}}{n_{\text{H},\text{SSb}}} \right) \right]^{-2.28} + 0.02 \left[ 1 + \left( \frac{n_{\text{H}}}{n_{\text{H},\text{SSb}}} \right) \right]^{-0.84},
\]

where \( \Gamma_{\text{UVB}} \) is the background photoionization rate and \( \Gamma_{\text{Phot}} \) is the total photoionization rate. Moreover, the self-shielding density threshold, \( n_{\text{H},\text{SSb}} \), is given by equation (13) and is thus a function of \( \Gamma_{\text{UVB}} \) and \( \sigma_{98} \), which vary with redshift. As explained in more detail in Appendix A1, the numerical parameters representing the shape of the profile are chosen to provide a redshift-independent best fit to our RT results. In addition, the parametrization is based on the main RT related quantities, namely the intensity of UVB radiation and its spectral shape. It can therefore be used for UVB models similar to the Haardt & Madau (2001) model we used in this work (e.g. Faucher-Giguère et al. 2009; Haardt & Madau 2012). For a given UVB model, one only needs to know \( \Gamma_{\text{UVB}} \) and \( \sigma_{98} \) in order

---

**Figure 2.** Left: \( n_{\text{H}} \)-weighted total hydrogen number density as a function of \( N_{\text{H}} \). The brown solid curve shows the RT result and the purple dotted curve shows the optically thin limit. Blue dot-dashed, dot-dot-dot-dashed and long dashed curves assume models with self-shielding density thresholds of \( n_{\text{H},\text{SSb}} = 10^{-4}, 10^{-2} \) and \( 10^{-3} \text{ cm}^{-3} \), respectively. All of the above mentioned curves show the median of the \( n_{\text{H}} \)-weighted total hydrogen number density at a given \( N_{\text{H}} \). The grey thin lines show the expected Jeans scaling relations for optically thin gas (equation 10; diagonal solid line) and for neutral gas (equation 11; steeper dotted line). A second solid line with the same slope expected from equation (11) but a different normalization is illustrated by the second solid line which is identical to the dotted line but shifted by 0.5 dex to higher \( n_{\text{H}} \). The pink and blue shaded areas in the right panel indicate the 70 per cent and 99 per cent scatter, respectively, while the solid curves show the median for the RT result. All the other curves are also medians. This shows that the \( n_{\text{H}} \)-\( N_{\text{H}} \) relationship can be explained by the Jeans scaling and that the flattening in the CDDF is due to self-shielding. Right: \( \text{H}_1 \) CDDF in the presence of the UVB and diffuse RR for simulation L06N256. Simulations shown with different curves are identical to the left panel. In addition, the effect of \( \text{H}_2 \) formation is shown by the green solid curve which deviates from the brown solid curve at \( N_{\text{H}} \gtrsim 3 \times 10^{21} \text{ cm}^{-2} \). Finally the red dashed curve, which is indistinguishable from the brown solid curve, shows the result of assuming the median of the photoionization rate profile of the RT results to calculate the neutral fractions (see Section 3.3 and Appendix A1). The top-section in the right panel shows the ratio between different \( \text{H}_1 \) CDDFs and the one resulting from the RT simulations.
to determine the corresponding $n_{\text{HI,SSS}}$ from (13) (see also Table 2). Then, after using equation (14) to calculate the photoionization rate as a function of density, the equilibrium hydrogen neutral fraction for different densities, temperatures and redshifts can be readily calculated as explained in Appendix A1.

We note that the parameters used in equation (14) are only accurate for photoionization-dominated cases. As we show in Section 3.5, at $z \sim 0$ the collisional ionization rate is greater than the total photoionization rate around the self-shielding density threshold. Consequently, equation (13) does not provide an accurate estimate of the self-shielding density threshold at low redshifts. In Appendix A1 we therefore report the parameters that best reproduce our RT results at $z = 0$. Our tests show that simulations that use equation (14) reproduce the $f(N_{\text{HI}}, z)$ accurately to within 10 per cent for $z \gtrsim 1$ where photoionization is dominant (see Appendix A1).

Although using the relation between the median photoionization rate and the gas density is a computationally efficient way of calculating equilibrium neutral fractions in big simulations, it comes at the expense of the information encoded in the scatter around the median photoionization rate at a given density. However, our experiments show that the error in $f(N_{\text{HI}}, z)$ that results from neglecting the scatter in the photoionization rate profile is negligible for $N_{\text{HI}} \gtrsim 10^{18} \text{ cm}^{-3}$ and less than $\lesssim 0.1$ dex at lower column densities (see Appendix A1).

### 3.4 The roles of diffuse recombination radiation and collisional ionization at $z = 3$

To study the interplay between different ionizing processes and their effects on the distribution of H$_{\text{i}}$, we compare their ionization rates at different densities. We start the analysis by presenting the results at $z = 3$ and extend it to other redshifts in Section 3.5.

The total photoionization rate profiles shown in the right panel of Fig. 3 are almost flat at low densities and decrease with increasing density, starting at densities $n_{\text{H}} \sim 10^{-4} \text{ cm}^{-3}$. Just below $n_{\text{H}} = 10^{-2} \text{ cm}^{-3}$ self-shielding causes a sharp drop, but the fall-off becomes shallower for $n_{\text{H}} > 10^{-2} \text{ cm}^{-3}$ and the photoionization rate starts to increase at $n_{\text{H}} > 10^{-3} \text{ cm}^{-3}$. As shown in Fig. 4, the shallower fall-off in the total photoionization rate with increasing density is caused by RR. The increase in the photoionization rate with density at the highest densities on the other hand is an artefact of the imposed temperature for ISM particles (i.e. $T = 10^4 \text{ K}$) which produces a rising collisional ionization rate with increasing density. As the comparison between the UVB and RR photoionization profiles shows (see Fig. 4), RR only starts to dominate the total photoionization rate at $n_{\text{H}} > 10^{-2} \text{ cm}^{-3}$, where
the UVB photoionization rate has dropped by more than one order of magnitude and the gas is no longer highly ionized. RR reduces the total \( H_1 \) content of high-density gas by \( \approx 20 \) per cent. Although ionization rates remain non-negligible at higher densities, they cannot keep the hydrogen highly ionized. For instance at \( n_H \sim 1 \text{ cm}^{-3} \), a photoionization rate of \( \Gamma \sim 10^{-14} \text{ s}^{-1} \) can only ionize the gas by \( \lesssim 20 \) per cent.

The shape of the photoionization rate profile produced by diffuse RR can be understood by noting that the production rate of RR increases with the density of ionized gas. At number densities \( n_H < 10^{-2} \text{ cm}^{-3} \), where the gas is highly ionized, the photoionization rate due to recombination photons is proportional to the density (i.e., \( \Gamma_{RR} \propto n_H \)). At higher densities on the other hand, the gas becomes neutral. As a result, the density of ionized gas decreases with increasing density and the production rate of recombination photons decreases. Therefore, there is a peak in the photoionization rate due to RR around the self-shielding density. At very low densities, the superposition of recombination photons which have escaped from higher densities becomes dominant and the net photoionization rate of recombination photons flattens. Note that our simulations may underestimate this asymptotic rate because our simulation volumes are small compared to the mean free path for ionizing radiation (which is \( \sim 100 \text{ Mpc} \) at \( z \sim 3 \)). On the other hand, the neglect of cosmological redshifting for RR will result in overestimation of its photoionization rate on large scales. Recombination photons also leak from lower densities to self-shielded regions, smoothing the transition between highly ionized and highly neutral gas. At high densities, in the absence of the UVB ionizing photons, RR and collisional ionization can boost each other by providing more free electrons and ions.

In Fig. 5 we compare hydrogen neutral fraction and photoionization rate profiles for different assumptions about RR. The hydrogen neutral fraction profile based on a precise RT calculation of RR is close to the Case A result at low densities \( (n_H \lesssim 10^{-3} \text{ cm}^{-3}) \) but converges to the Case B result at high densities \( (n_H \gtrsim 10^{-1} \text{ cm}^{-3}) \). This suggests that the neutral fraction profile, though not the ionization rate, can be modelled by switching from Case A to Case B recombination at \( n_H \sim n_{H,SSh} \) (e.g. Altay et al. 2011; McQuinn et al. 2011).

### 3.5 Evolution

The general trends in the profile of the photoionization rates with density and their influences on the distribution of \( H_1 \) are not very sensitive to redshift. However, as shown in Table 2, the intensity and hardness of the UVB radiation change with redshift which, in turn, changes the self-shielding density. Moreover, as the Universe expands, the average density of absorbers decreases and their distributions evolve. The larger structures that form at lower redshifts drastically change the temperature structure of the gas at low and intermediate densities where collisional ionization becomes the dominant process. In the top-left panel of Fig. 6, the evolution of the hydrogen neutral fraction is illustrated for the \( L50N512-W3 \) simulation. As discussed in Section 3.3 and Appendix A1, since the photoionization rate profiles are converged with box size and resolution, we apply the profiles derived from an RT simulation of a smaller box, or a subset of the big box at lower redshifts,\(^8\) to calculate the neutral fractions in this big box. Fig. 6 shows that the neutral fraction profiles are similar in shape at high redshifts but that at \( z \lesssim 1 \) the profiles are largely different, particularly at low hydrogen number densities, due to the evolving collisional ionization rates.

The evolution of the collisional ionization rate profiles is shown in the bottom-left panel of Fig. 6. At \( z \gtrsim 2 \) and for \( n_H < 10^{-2} \text{ cm}^{-3} \), the collisional ionization rate is not high enough to compete with the UVB photoionization rate. At lower redshifts and for number

---

\(^8\) Strong collisional ionization at low redshifts can change the self-shielding. Therefore, for our RT simulation at low redshifts (i.e. \( z \lesssim 1 \)) we used a representative sub-volume of the \( L50N512-W3 \) simulation sufficiently large for the collisional ionization rates to be converged.
Figure 6. Evolution of the hydrogen neutral fraction profile (top-left) and various ionization rates as a function of density. Top-right, bottom-left and bottom-right panels show, respectively, the total photoionization rates, collisional ionization rates and RR photoionization rates. All the H I fractions and collisional ionization rates which are sensitive to both collisional ionization and photoionization are taken from the L50N512-W3 simulation. Photoionization rates at $z \geq 2$ are based on the L06N128 simulation. At lower redshifts (i.e. $z = 0$ and 1), where the box size becomes important because of the collisional ionization and its effect on changing the self-shielding, we used a representative sub-volume of the L50N512-W3 simulation to calculate the photoionization rate profile with density. While the overall shape of the UVB photoionization rate profile is similar at different redshifts, the collisional ionization becomes increasingly stronger at lower redshifts and strongly reduces the hydrogen neutral fractions at densities $n_H \lesssim 10^{-3} \text{ cm}^{-3}$.

At lower redshifts, this causes a steeper high-density fall-off in the collisional ionization rate as shown in the bottom-left panel of Fig. 6. The differences between the total photoionization rates at different redshifts shown in the top-right panel of Fig. 6, are caused by the evolution of the UVB intensity and its hardness, which affects the self-shielding density thresholds (see equation 13). On the other hand, as we showed in the previous section, the peak of the photoionization rate produced by RR tracks the self-shielding density. As a result the peak of the RR photoionization rate also changes with redshift as illustrated in the bottom-right panel of Fig. 6.

The filled circles in Fig. 7 indicate the UVB photoionization rate versus the number density at which the RR photoionization rate peaks. The self-shielding density expected from the Jeans scaling argument (equation 13) is also shown (green dotted line). The peaks in the RR photoionization rate in RT simulations follow this expected scaling for $z \geq 1$. However, the $z = 0$ result deviates from this trend since collisional ionization affects the self-shielding density threshold, a factor that is not captured by equation (13).

As a result of the RR photoionization rate peaking around the self-shielding density threshold, the transition between highly ionized and nearly neutral gas becomes more extended at all redshifts. To illustrate the smoothness of this transition, the densities at which...
The CDDF, \( f(N_{\text{HI}}, z) \), predicted by our RT simulations is in excellent agreement with observational constraints at all redshifts \( z = 0 \)–5 and reproduces the slopes of the observed \( f(N_{\text{HI}}, z) \) function for a wide range of HI column densities. At low HI column densities, the CDDF is a steep function which decreases with increasing \( N_{\text{HI}} \) before it flattens at \( N_{\text{HI}} \gtrsim 10^{13} \text{ cm}^{-2} \) due to self-shielding. At \( N_{\text{HI}} \gtrsim 10^{21} \text{ cm}^{-2} \), on the other hand, \( f(N_{\text{HI}}, z) \) is determined mainly by the intrinsic distribution of total hydrogen and the \( \text{H}_2 \) fraction.

We showed that the \( N_{\text{HI}}-n_{\text{HI}} \) relationship can be explained by a simple Jeans scaling. This argument assumes HI absorbers to be self-gravitating systems close to local hydrostatic equilibrium (Schaye 2001a) and to be either neutral or in photoionization equilibrium in the presence of an ionizing radiation field. However, at \( z \approx 0 \) the analytic treatment underestimates the self-shielding density threshold due to its neglect of collisional ionization.

The high HI column density end of the predicted \( f(N_{\text{HI}}, z) \) evolves only weakly from \( z = 5 \) to \( z = 0 \), consistent with observations. In the Lyman limit range of the distribution function, the slope of \( f(N_{\text{HI}}, z) \) remains the same at all redshifts. However, at \( z > 3 \) the number of absorbers increases with redshift as the Universe becomes denser while the UVB intensity remains similar. At lower redshifts, on the other hand, the combination of a decreasing UVB intensity and the expansion of the Universe results in a non-evolving \( f(N_{\text{HI}}, z) \). In contrast, the number of absorbers with lower HI column densities (i.e. the Ly\( \alpha \) forest) decreases significantly from \( z \sim 3 \). We showed that this results in part from the stronger collisional ionization at redshifts \( z \lesssim 1 \), which compensates for the lower intensity of the UVB. The increasing importance of collisional ionization is due to the rise in the fraction of hot gas due to shock-heating associated with the formation of structure.

The inclusion of diffuse RR smooths the transition between optically thin and thick gas. Consequently, the transition to highly neutral gas is not as sharp as what has been assumed in some previous works (e.g. Nagamine et al. 2010; Yajima et al. 2012; Goerdt et al. 2012). For instance, the difference in the gas density at which hydrogen is highly ionized (i.e. \( n_{\text{HII}}/n_{\text{H}} \approx 0.01 \)) and the density at which gas is highly neutral (i.e. \( n_{\text{HII}}/n_{\text{H}} \gtrsim 0.5 \)) is more than one order of magnitude (see Fig. 7). As a result, assuming a sharp self-shielding density threshold at the density for which the optical depth of ionizing photons is \( \sim 1 \) overestimates the resulting neutral hydrogen mass by a factor of a few.

Our simulations adopted some commonly used approximations (e.g. neglecting helium RT effects, using a grey approximation in order to mimic the UVB spectra, neglecting absorption by dust and local sources of ionizing radiation). Our tests show that most of those approximations have negligible effects on our results. But there are some assumptions which require further investigation. For instance, the presence of young stars in high-density regions could change the HI CDDF, especially at high HI column densities through feedback and emission of ionizing photons. Indeed, we will show in Rahmati et al. (2013) that for very high column densities the ionizing radiation from young stars can reduce the \( f(N_{\text{HI}}, z) \) by 0.5–1 dex.

**ACKNOWLEDGEMENTS**

We thank the anonymous referee for a helpful report. We thank Garbriel Altay for providing us with his simulation results and a compilation of the observed HI CDDF. We also would like to thank Kristian Finlator, J. Xavier Prochaska, Tom Theuns and all the members of the OWLS team for valuable discussions and Marcel...
REFERENCES

Aguirre A., Dow-Hygelund C., Schaye J., Theuns T., 2008, ApJ, 689, 851
Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, ApJ, 737, L37
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933
Cen R., 2012, ApJ, 748, 121
Cen R., Ostriker J. P., Prochaska J. X., Wolfe A. M., 2003, ApJ, 598, 741
Chabrier G., 2003, PASP, 115, 763
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Davé R., Oppenheimer B. D., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 408, 2051
Duffy A. R., Kay S. T., Battye R. A., Booth C. M., Dalla Vecchia C., Schaye J., 2012, MNRAS, 420, 2799
Erb D., Gnedin N. Y., Kravtsov A. V., 2012, ApJ, 761, 54
Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416
Faucher-Giguère C.-A., Keres D., Dijkstra M., Springel V., Hernquist L., Zaldarriaga M., 2011, ApJ, 725, 633
Friedrich M. M., Mellema G., Iliev I. T., Shapiro P. R., 2012, MNRAS, 421, 2232
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D.,Primack J. R., 2011, MNRAS, 418, 1796
Furlanetto S. R., Schaye J., Springel V., Hernquist L., 2005, ApJ, 622, 7
Gardner J. P., Katz N., Weinberg D. H., 1997, ApJ, 484, 31
Gardner J. P., Katz N., Weinberg D. H., 2001, ApJ, 559, 131
Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, ApJ, 672, 765
Goerdt T., Dekel A., Sternberg A., Gnat O., Ceverino D., 2012, preprint (arXiv:1205.2021)
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays
Haardt F., Madau P., 2012, ApJ, 746, 125
Haehnelt M. G., Steinmetz M., Rauch M., 1998, ApJ, 495, 647
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Janknecht E., Reimers D., Lopez S., Tytler D., 2006, A&A, 458, 427
Katz N., Weinberg D. H., Hernquist L., Miralda-Escude J., 1996, ApJ, 457, L57
Kim T.-S., Carswell R. F., Cristiani S., D’Odorico S., Gilangono E., 2002, MNRAS, 335, 555
Komatsu E. et al., 2011, ApJS, 192, 18
Krumholz M. R., McKee C. F., Tumlinson J., 2009a, ApJ, 693, 216
Krumholz M. R., Ellison S. L., Prochaska J. X., Tumlinson J., 2009b, ApJ, 701, L12
Lehner N., Savage B. D., Richter P., Sembach K. R., Tripp T. M., Wakker B. P., 2007, ApJ, 658, 680
McQuinn M., Switzer E. R., 2010, MNRAS, 408, 1945
McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, ApJ, 743, 82
Miralda-Escudé J., 2005, ApJ, 620, L91
Nagamine K., Springel V., Hernquist L., 2004, MNRAS, 348, 421
Nagamine K., Wolfe A. M., Hernquist L., Springel V., 2007, ApJ, 660, 945
Nagamine K., Choi J.-H., Yajima H., 2010, ApJ, 725, L219
Noteberda P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087
Noteberda P. et al., 2012, preprint (arXiv:1210.1213)
O’Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666
O’Meara J. M., Prochaska J. X., Worseck G., Chen H.-W., Madau P., 2012, preprint (arXiv:1204.3093)
Osterbrock D. E., Ferland G. J., 2006, in Osterbrock D.E., Ferland G.J., eds, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, 2nd. edn. University Science Books, Sausalito, CA
Pawlik A. H., Schaye J., 2008, MNRAS, 389, 651
Pawlik A. H., Schaye J., 2011, MNRAS, 412, 1493
Péroux C., Dessauges-Zavadsky M., D’Odorico S., Sun Kim T., McMahon R. G., 2005, MNRAS, 363, 479
Petitjean P., Bergeron J., Puget J. L., 1992, A&A, 265, 375
Ponzen et al., 2008, MNRAS, 390, 1349
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
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
Rahmati A., van der Werf P. F., 2011, MNRAS, 418, 176
Rahmati A., Schaye J., Pawlik A. H., Raicevic M., 2013, arXiv:1301.1978
Razoumov A. O., Norman M. L., Prochaska J. X., Wolfe A. M., 2006, ApJ, 645, 55
Ribaudo J., Lehner N., Howk J. C., 2011, ApJ, 736, 42
Schaye J., 2001a, ApJ, 559, 507
Schaye J., 2001b, ApJ, 562, L95
Schaye J., 2004, ApJ, 699, 667
Schaye J., 2006, ApJ, 643, 59
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J. et al., 2010, MNRAS, 402, 1493
Springel V., 2005, MNRAS, 364, 1105
Tajiri Y., Umemura M., 1998, ApJ, 502, 59
Tepper-Garcia T., Richter P., Schaye J., Booth C. M., Dalla Vecchia C., Theuns T., 2012, MNRAS, 425, 1640
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Tytler D., 1987, ApJ, 321, 49
van de Voort F., Schaye J., Altay G., Theuns T., 2012, MNRAS, 421, 2809
Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, ApJ, 465, 487
Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296
Wiersma R. P. C., Schaye J., 2009a, MNRAS, 393, 99
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574
Yajima H., Choi J.-H., Nagamine K., 2012, MNRAS, 427, 2889
Zhang Z., Miralda-Escude J., 2002, ApJ, 568, L71
Zwaan M. A., van der Hulst J. M., Briggs F. H., Verheijen M. A. W., 2005, MNRAS, 364, 1467

APPENDIX A: PHOTIONIZATION RATE AS A FUNCTION OF DENSITY

A1 Replacing the RT simulations with a fitting function

In Section 3.3 we demonstrated that the median of the simulated relation between the total photoionization rate, \( \Gamma_{\text{phot}} \), and density is converged with respect to resolution and box size. We used this result and provided fits to the median of this relation. We have exploited these fits to compute the neutral hydrogen fraction in cosmological simulations under the assumption of ionization equilibrium (see Appendix A2), without performing the computationally demanding RT. In this section, we discuss the accuracy of these fits.
Figure A1. Left: the ratio between the H I CDDF calculated using the RT based $\Gamma_{\text{phot}}-n_H$ relationship and the actual RT results for the L06N128 simulation in the presence of the UVB and diffuse RR, at $z = 3$. The orange solid line shows the result of using the median photoionization rate–density profile predicted by the RT simulation. The blue dashed curve shows the result of including the scatter around the median in the calculations. Right: H I CDDFs calculated using the $\Gamma_{\text{phot}}$–density fitting function (i.e. equation 14) are compared to the H I CDDFs for which the actual $\Gamma_{\text{phot}}$–density relation from the RT simulations are used. Blue and green curves are for $z = 0$ and $z = 2$, respectively, and the red curve is for $z = 4$. The difference between the RT result and the result of using the fitting function at $z = 0$ is due to the importance of collisional ionization at $z = 0$. To capture this effect and to reproduce the RT results at $z = 0$, we advise using the best-fitting parameters shown in Table A2. All the CDDFs are for the L50N512-W3 simulation and in the presence of the UVB and diffuse RR.

Table A1. The best-fitting parameters for equation (A1) at different redshifts based on RT results in the L06N128 simulation.

| Redshift | $\log n_0 (\text{cm}^{-3})$ | $\alpha_1$ | $\alpha_2$ | $\beta$ | $1-f$ |
|----------|--------------------------|------------|------------|--------|------|
| $z = 1-5$ | $\log [n_{\text{H}, \text{SSB}}]$ | $-2.28$ | $-0.84$ | $1.64$ | $0.98$ |
| $z = 0$    | $-2.94$ | $-3.98$ | $-1.09$ | $1.29$ | $0.99$ |
| $z = 1$    | $-2.29$ | $-2.94$ | $-0.90$ | $1.21$ | $0.97$ |
| $z = 2$    | $-2.06$ | $-2.22$ | $-1.09$ | $1.75$ | $0.97$ |
| $z = 3$    | $-2.13$ | $-1.99$ | $-0.88$ | $1.72$ | $0.96$ |
| $z = 4$    | $-2.23$ | $-2.05$ | $-0.75$ | $1.93$ | $0.98$ |
| $z = 5$    | $-2.35$ | $-2.63$ | $-0.57$ | $1.77$ | $0.99$ |

The left panel of Fig. A1 shows that using the median photoionization rates produces an H I CDDF in very good agreement with the H I CDDF obtained from the corresponding RT simulation (orange solid curve) at $N_{\text{HI}} \gtrsim 10^{18}$ cm$^{-2}$. However, there is a small systematic difference at lower column densities. One may think that this small difference is caused by the loss of information contained in the scatter in the photoionization rates at fixed density. We tested this hypothesis by including a log-normal random scatter around the median photoionization rate consistent with the scatter exhibited by the RT result. However, after accounting for the random scatter, the $f(N_{\text{HI}}, z)$ is slightly over-produced compared to the full RT result at nearly all H I column densities.

We exploit the insensitivity of the shape of the $\Gamma_{\text{phot}}$–density relation to the redshift, and propose the following fit to the photoionization rate, $\Gamma_{\text{phot}}$:

$$\frac{\Gamma_{\text{phot}}}{\Gamma_{\text{UVB}}} = (1-f) \left[ 1 + \left( \frac{n_H}{n_0} \right)^{\alpha_1} \right] + f \left[ 1 + \left( \frac{n_H}{n_0} \right)^{\alpha_2} \right],$$  \hspace{1cm} (A1)

Figure A2. Comparisons between the total photoionization rates as a function of density in the L06N128 simulation. Photoionization rates based on the RT simulations and best-fitting functions at $z = 4$ and $z = 0$ are shown in the left and right panels, respectively. In each panel, the RT result is shown with the orange solid curve. The best fit to the RT result at a given redshift (equation A1 and Table A1) is shown with the blue dashed curve and the best fit to the RT results at $z = 1-5$ (equation 14) is shown with the purple dotted curve. As shown in the right panel, because of the impact of collisional ionization on self-shielding, the low redshift photoionization curve (the blue dashed curve) deviates from the best fit to the results at higher redshifts (the purple dotted curve). To resolve this issue and to capture the impact of collisional ionization, we advise using the best-fitting parameters shown in Table A2 for $z = 0$. 

Downloaded from https://academic.oup.com/mnras/article-abstract/430/3/2427/1751889 by guest on 28 July 2018
Table A2. The best-fitting parameters for equation (A1) at $z = 0$ based on RT results for the $L50NS12$ simulation. To capture the impact of collisional ionization on the self-shielding, one needs to use large cosmological simulations. The simulation with a box size of $50 \, h^{-1} \, \text{Mpc}$ results in converged collisional ionizations.

| Redshift | $\log (n_0) \, (\text{cm}^{-3})$ | $\alpha_1$ | $\alpha_2$ | $\beta$ | $1 - f$ |
|----------|---------------------------------|-----------|-----------|-------|-------|
| $z = 0$  | -2.56                           | -1.86     | -0.51     | 2.83  | 0.99  |

where $\Gamma_{\text{UVB}}$ is the photoionization rate due to the ionizing background, and $n_0$, $\alpha_1$, $\alpha_2$ and $\beta$ are parameters of the fit. The best-fitting values of these parameters are listed in Table A1 and the photoionization rate–density relations they produce are compared with the RT simulations at redshifts $z = 0$ and $z = 4$ in Fig. A2. At all redshifts, the best-fitting value of $n_0$ is almost identical to the self-shielding density threshold, $n_{\text{H,SSS}}$, defined in equation (13), and the characteristic slopes of the photoionization rate–density relation are similar. This suggests that one can find a single set of best-fitting values to reproduce the RT results at $z \gtrsim 1$. The corresponding best-fitting parameter values are (see also equation 14) $\alpha_1 = -2.28 \pm 0.31$, $\alpha_2 = -0.84 \pm 0.11$, $n_0 = (1.003 \pm 0.005) \times n_{\text{H,SSS}}$, $\beta = 1.64 \pm 0.19$ and $f = 0.02 \pm 0.0089$.

In the right panel of Fig. A1, the ratio between the H II CDDF calculated using the fitting function presented in equation (14) and the RT based H I CDDF (i.e. calculated using the median of the photoionization rate–density relation in the RT simulations) is shown for the $L50NS12$-$W2$ and at $z = 0$, 2 and 4. This illustrates that the fitting function reproduces the RT results accurately, except at $z = 0$. As explained in Section 3, this is expected since at low redshifts collisional ionization affects the self-shielding and the resulting photoionization rate–density profile. However, a separate fit can be obtained using converged RT results at $z = 0$. The parameters that define such a fit are shown in Table A2.

A2 The equilibrium hydrogen neutral fraction

In this section we explain how to derive the neutral fraction in ionization equilibrium. Equating the total number of ionizations per unit time per unit volume with the total number of recombinations per unit time per unit volume, we obtain

$$n_{\text{HI}}, \Gamma_{\text{TOT}} = a_A \, n_a \, n_{\text{HI}},$$

(A2)

where $n_{\text{HI}}$, $n_a$ and $n_{\text{HI}}$ are the number densities of neutral hydrogen atoms, free electrons and protons, respectively. $\Gamma_{\text{TOT}}$ is the total ionization rate per neutral hydrogen atom and $a_A$ is the Case A recombination rate\(^{10}\) for which we use the fitting function given by Hui & Gnedin (1997):

$$a_A = 1.269 \times 10^{-13} \frac{\lambda^{1.503}}{(1 + (\lambda / 0.522)^{0.47})^{1.927}} \text{ cm}^3 \text{ s}^{-1},$$

(A3)

where $\lambda = 315614 / T$.

Defining the hydrogen neutral fraction as the ratio between the number densities of neutral hydrogen and total hydrogen, $\eta = n_{\text{HI}} / n_a$, and ignoring helium (which is an excellent approximation, see Appendix D2), we can rewrite equation (A3) as

$$\eta \, \Gamma_{\text{TOT}} = a_A \, (1 - \eta)^2 \, n_a.$$

(A4)

Furthermore, we can assume that the total ionization rate, $\Gamma_{\text{TOT}}$, consists of two components: the total photoionization rate, $\Gamma_{\text{Phot}}$, and the collisional ionization rate, $\Gamma_{\text{Col}}$:

$$\Gamma_{\text{TOT}} = \Gamma_{\text{Phot}} + \Gamma_{\text{Col}},$$

(A5)

where $\Gamma_{\text{Col}} = \Lambda_{\text{T}} \, (1 - \eta) \, n_{\text{HI}}$. The photoionization rate can be expressed as a function of density using equation (14). For $\Lambda_{\text{T}}$, which depends only on temperature, we use a relation given in Theuns et al. (1998):

$$\Lambda_{\text{T}} = 1.17 \times 10^{-10} \frac{T^{1/2} \exp(-157809 / T)}{1 + \sqrt{T} / 10^5} \text{ cm}^3 \text{ s}^{-1}.$$  

(A6)

We can now rearrange equation (A4) as a quadratic equation:

$$A \, \eta^2 - B \, \eta + C = 0,$$

(A7)

with $A = a_A + \Lambda_{T}$, $B = 2a_A + \frac{T_{\text{phot}}}{\Lambda_{T}} + \Lambda_{T}$ and $C = a_A$ which gives

$$\eta = \frac{B - \sqrt{B^2 - 4AC}}{2A}.$$  

(A8)

Using the last equation one can calculate the equilibrium hydrogen neutral fraction for a given $n_a$ and temperature.

APPENDIX B: THE EFFECTS OF BOX SIZE, COSMOLOGICAL PARAMETERS AND RESOLUTION ON THE H I CDDF

The size of the simulation box may limit the abundance and the density of the densest systems captured by the simulation. In other words, very massive structures, which may be associated with the highest H I column densities, cannot be formed in a small cosmological box. Indeed, as shown in the top panels of Fig. B1, one needs to use cosmological boxes larger than $\gtrsim 25$ comoving $h^{-1}$ Mpc in order to achieve convergence in the H I distribution (see also Altay et al. 2011). On the other hand, the bottom-right panel of Fig. B1 shows that changing the resolution of the cosmological simulations also affects $f(N_{\text{HI}}, z)$, although the effect is small.

The adopted cosmological parameters also affect the gas distribution and hence the H I CDDF. For instance, one expects that the number of absorbers at a given density varies with the density parameter $\Omega_m$, and the root mean square amplitude of density fluctuations $\sigma_8$. The bottom-left panel of Fig. B1 shows the ratio of column densities in simulations assuming WMAP 7-year and 3-year parameters. The ratio is only weakly dependent on the box size of the simulation and its resolution. This motivates us to use this ratio to convert the H I CDDF between the two cosmologies for all box sizes and resolutions (at any given redshift). While this is an approximate way of correcting for the difference in the cosmological parameters, it does not affect the main conclusions presented in this work [e.g. the lack of evolution of $f(N_{\text{HI}}, z)$].

APPENDIX C: RT CONVERGENCE TESTS

C1 Angular resolution

The left panel of Fig. C1 shows the dependence of photoionization rates on the adopted angular resolution, i.e. the opening angle...
Figure B1. The relative changes in the H\textsc{i} CDDF using different resolutions, box sizes and cosmologies in the presence of the UVB and diffuse RR. The top-left panel shows the effect of box size on $f(N_{\text{HI}}, z)$ for a fixed resolution at $z = 3$, where the orange solid (blue dashed) curve shows the difference between using a box size of $L = 25 (50)$ comoving $h^{-1}$Mpc and a box size of $L = 100$ comoving $h^{-1}$Mpc. The top-right panel shows the same effect but for smaller box sizes: the orange solid (blue dashed) curve shows the difference between using a box size of $L = 6 (12)$ comoving $h^{-1}$Mpc and a box size of $L = 25$ comoving $h^{-1}$Mpc. The bottom-left shows the effect of using a cosmology consistent with WMAP 3-year results instead of using a cosmology based on the WMAP 7-year constraints. The orange solid and blue dashed curves show this effect for simulations with box sizes of $L = 6$ and 25 comoving $h^{-1}$Mpc, respectively. The bottom-right panel shows the effect of resolution.

Figure C1. The UVB photoionization rate is converged for our adopted angular resolution, i.e. $N_{\text{TC}} = 64$, as shown in the left panel and our adopted number of ViP neighbours, i.e. $N_{\text{GB}} = 5$, as shown in the right panel. Photoionization rate profiles are shown for the $L06N128$ simulation in the presence of the UVB radiation where the Case A recombination is adopted. The curves show the medians and the shaded areas around them indicate the 15–85 percentiles.
of the transmission cones $4\pi/N_{TC}$. The photoionization rates are converged for $N_{TC} = 64$ (our fiducial value) or higher.

C2 The number of ViP neighbours

The right panel of Fig. C1 shows the dependence of the photoionization rates on the number of SPH neighbours of ViPs. As discussed in Section 2.2, ViPs distribute the ionizing photons they absorb among their NGBVip nearest SPH neighbours. The larger the number of neighbours, the larger the volume over which photons are distributed, and the more extended is the transition between highly ionized and self-shielded gas. The photoionization rates converge for $\lesssim 5$ ViP neighbours (our fiducial value is 5).

C3 Direct comparison with another RT method

Altay et al. (2011) used cosmological simulations from the reference model of the OWLS project (Schaye et al. 2010), i.e. a simulation run with the same hydro code as we used in this work, to investigate the effect of the UVB on the H\textsc{i} CDDF at $z = 3$. However, they employed a ray-tracing method very different from the RT method we use here. Furthermore, they did not explicitly treat the transfer of RR. In Fig. C2, we compare one of our UVB photoionization rate profiles\textsuperscript{11} with the photoionization rate found by Altay et al. (2011) in a similar simulation. The overall agreement is very good, but the comparison also reveals important differences.

Altay et al. (2011) calculate the average optical depth around every SPH particle within a distance of 100 proper kpc, assuming the UVB is unattenuated at larger distances. Then, they use this optical depth to calculate the attenuation of the UVB photoionization rate for every particle. This procedure may underestimate the small but non-negligible absorption of UVB ionizing photons on large scales. Indeed, by tracing the self-consistent propagation of photons inside the simulation box, we have found that the UVB photoionization rate decreases gradually with increasing density up to the density of self-shielding. However, we note that the small differences between our UVB photoionization rates and those calculated by Altay et al. (2011) at densities below the self-shielding become slightly smaller by increasing the angular resolution in our RT calculations (see the left panel of Fig. C1).

APPENDIX D: APPROXIMATED PROCESSES

D1 Multi-frequency effects

As discussed in Section 2.3, in our RT simulations we have treated the multi-frequency nature of the UVB radiation in the grey approximation (see equation 4). This approach does not capture the spectral hardening which is a consequence of variation of the absorption cross-sections with frequencies. We tested the impact of spectral hardening on the H\textsc{i} fractions by repeating the $L06N128$ simulation at $z = 3$ with the UVB using three frequency bins. We used energy intervals [13.6–16.6], [16.6–24.6] and [24.6–54.4] eV and assumed that photons with higher frequencies are absorbed by He. The result is illustrated in the top section of the left panel in Fig. D1 by plotting the ratio between the resulting hydrogen neutral fraction, $n_\text{H}^\text{\alpha}$, and the same quantity in the original simulation that uses the grey approximation. This comparison shows that the simulation that uses multi-frequency predicts hydrogen neutral fractions <10 per cent lower at low densities (i.e. $n_\text{H}^\text{\alpha} \lesssim 10^{-4}$ cm$^{-3}$). This does not change the resulting $f(N_{\text{Hi}}, z)$ noticeably at the column densities of interest here.

The spectral hardening captured in the simulation with three frequency bins is illustrated in the right panel of Fig. D1. This figure shows the fractional contribution of different frequencies to the total UVB photoionization rate as a function of density. The red solid curve shows the contribution of the bin with the lowest frequency and drops at the self-shielding density threshold. On the other hand, the fractional contribution of the hardest frequency bin increases at higher densities, as shown with the blue dashed curve. Despite the differences in the fractional contributions to the total UVB photoionization rate, the absolute photoionization rates drop rapidly at densities higher than the self-shielding threshold for all frequency bins.

D2 Helium treatment

A simplifying assumption frequently used in RT simulations which aim to calculate the distribution of neutral hydrogen is to ignore helium in the ionization processes (e.g. Faucher-Giguère et al. 2009; McQuinn & Switzer 2010; Altay et al. 2011). We adopted the same

---

\textsuperscript{11} Note that in our simulations the UVB photoionization rate is converged with the box size and the resolution as shown in Section 3.3.
assumption in our RT calculations which implies that we implicitly assumed the ionization state of neutral helium and its interaction with free electrons to be similar to the trends followed by neutral hydrogen. This has been shown to be a good assumption (Osterbrock & Ferland 2006; McQuinn & Switzer 2010; Friedrich et al. 2012). Nevertheless, we tested the validity of our approximate helium treatment by repeating the L06N128 simulation at $z = 3$ with the UVB using four frequency bins and an explicit He treatment. The first three frequency bins are identical to the bins used in the previous section (i.e. [13.6–16.6], [16.6–24.6] and [24.6–54.4] eV) and the last bin is chosen to cover higher frequencies which are capable of He II ionization. We adopted a helium mass fraction of 25 per cent and a Case A recombination rate. The ratio between the resulting hydrogen neutral fraction and the same quantity when a single frequency is used and helium is not treated explicitly is illustrated in the bottom-left panel of Fig. D1. The hydrogen neutral fractions are very close in the two simulations. However, the simulation with multi-frequency and explicit He treatment results in hydrogen neutral fractions that are ~10 per cent higher at low densities (i.e. $n_H \lesssim 10^{-4}$ cm$^{-3}$). This difference is barely noticeable in the comparison between the two H I CDDFs (not shown).

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