Simulating the cosmic distribution of neutral hydrogen and its connection with galaxies

Rahmati, A.

Citation
Rahmati, A. (2013, October 15). *Simulating the cosmic distribution of neutral hydrogen and its connection with galaxies*. Retrieved from https://hdl.handle.net/1887/21949

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/21949 holds various files of this Leiden University dissertation.

Author: Rahmati, Alireza
Title: Simulating the cosmic distribution of neutral hydrogen and its connection with galaxies
Issue Date: 2013-10-15

2

On the evolution of the HI column density distribution in cosmological simulations

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 yields 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.

Alireza Rahmati, Andreas H. Pawlik, Milan Raičevič, Joop Schaye

2.1   Introduction

A substantial fraction of the interstellar medium (ISM) in galaxies consists of atomic hydrogen. This makes studying the distribution of neutral hydrogen (HI) and its evolution crucial for our understanding of various aspects of star formation. In the local universe, the HI 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 Kilometer Array[1]. However, at $z \lesssim 6$, i.e., after reionization, the neutral gas can already be probed through the absorption signatures imprinted by the intervening HI systems on the spectra of bright background sources, such as quasars (QSOs).

The early observational constraints on the HI column density distribution function (HI CDDF hereafter), from quasar absorption spectroscopy at $z \lesssim 3$, were well described by a single power-law in the range $N_{\mathrm{HI}} \sim 10^{13} - 10^{21}\,\mathrm{cm}^{-2}$ (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 HI column densities and to higher redshifts (e.g., Kim et al., 2002; Péroux et al., 2005; O'Meara et al., 2007; Noterdaeme et al., 2009; Prochaska et al., 2009; Prochaska & Wolfe, 2009; O'Meara et al., 2013; Noterdaeme et al., 2012). These studies have revealed a much more complex shape which has been described using several different power-law functions (e.g., Prochaska et al., 2010; O'Meara et al., 2013).

The shape of the HI CDDF is determined by both the distribution and ionization state of hydrogen. Consequently, determining the distribution function of HI column densities requires not only accurate modeling of the cosmological distribution of gas, but also radiative transfer (RT) of ionizing photons. As a starting point, the HI CDDF can be modeled by assuming a certain gas profile and exposing it to an ambient ionizing radiation field (e.g., Petitjean et al., 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 HI CDDF which results from the cumulative effect of large numbers of objects 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 HI column density. At low HI column densities (i.e., $N_{\mathrm{HI}} \lesssim 10^{17}\,\mathrm{cm}^{-2}$, corresponding to the so-called Lyman-α forest), hydrogen is highly ionized by the metagalactic ultraviolet background radiation (hereafter UVB) and largely transparent to the ion-

---

[1]http://www.skatelescope.org/

izing radiation. For these systems, the HI column densities can therefore be accurately computed in the optically thin limit. At higher HI column densities (i.e., $N_{HI} \gtrsim 10^{17}\,\mathrm{cm}^{-2}$, corresponding to the so-called Lyman Limit and Damped Lyman-α systems), the gas becomes optically thick and self-shielded. As a result, the accurate computation of the HI column densities in these systems requires precise RT simulations. On the other hand, at the highest HI 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., 2009; Altay et al., 2011) or mechanical feedback from young stars and / or AGNs (Erkal et al., 2012), can also affect the highest HI column densities.

Despite the importance of RT effects, most of the previous theoretical works on the HI column density distribution did not attempt to model RT effects in detail (e.g., Katz et al., 1996; Gardner et al., 1997; Haehnelt et al., 1998; Gardner et al., 2001; Cen et al., 2003; Nagamine et al., 2004, 2007). Only very recent works incorporated RT, primarily to account for the attenuation of the UVB (Razoumov et al., 2006; Pontzen et al., 2008; Fumagalli et al., 2011; Altay et al., 2011; McQuinn et al., 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 HI distribution. They found that the HI CDDF in current cosmological simulations is in reasonable agreement with observations in a large range of HI column densities. Only at the highest HI column densities (i.e., $N_{HI} \gtrsim 10^{21}\,\mathrm{cm}^{-2}$) the agreement is poor. However, it is worth noting that the interpretation of these HI 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_{HI} systems.

In this chapter, we investigate the cosmological HI 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 HI 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 recombination radiation explicitly, by propagating RR photons. Using these simulations we study the evolution of the HI CDDF in the range of redshifts $z = 0 - 5$ for column densities $N_{HI} \gtrsim 10^{16}\,\mathrm{cm}^{-2}$. We discuss how the individual contributions from the UVB, RR and collisional ionization to the total ionization rate shape the HI CDDF and assess their relative importance at different redshifts.

The structure of this chapter is as follows. In §2.2 we describe the details of the hydrodynamical simulations and of the RT, including the treatment of the UVB and recombination radiation. In §2.3 we present the simulated HI 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 §2.4.

2.2   Simulation techniques

2.2.1   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. (2009a), which traces the abundance evolution of eleven elements by following stellar evolution assuming a Chabrier (2003) initial mass function. Moreover, a radiative heating and cooling implementation based on Wiersma et al. (2009b) 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 kms^{-1} and a mass loading parameter $\eta = 2$ (Dalla Vecchia & Schaye, 2008). Our tests show that varying the implementation of the kinetic feedback only changes the HI CDDF in the highest column densities ($N_{HI} \gtrsim 10^{21}$ cm^{-2}). However, the differences caused by these variations are smaller than the evolution in the HI CDDF and observational uncertainties (see Altay et al. in prep.).

We adopt fiducial cosmological parameters consistent with the most recent 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 $1.7 \times 10^5\ h^{-1}$M$_\odot - 8.7 \times 10^7\ h^{-1}M_\odot$. 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 2.1.

2.2.2 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 ν 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 neighbors for which we generally use the same number of SPH neighbors used for the underlying hydrodynamical simulations, i.e., $N_{ngb} = 48$.

After source particles emit photon packets isotropically to their neighbors, the photon packets travel along their propagation directions to other neighboring SPH particles which are inside their transmission cones. Transmission cones are regular cones with opening solid angle $4\pi/N_{TC}$ and are centered on the propagation direction. The parameter N_{TC} sets the angular resolution of the RT, and we adopt $N_{TC} = 64$. 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 neighboring 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_{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 $N_{RC} \times N_\nu$ photon packets, where N_ν is the number of

Table 2.1: List of cosmological simulations used in this work. All the simulations use model ingredients identical to the reference simulation of Schaye et al. (2010). From left to right the columns show: simulation identifier; comoving box size; number of dark matter particles (there are equally many baryonic particles); initial baryonic particle mass; dark matter particle mass; comoving (Plummer-equivalent) gravitational softening; maximum physical softening; final redshift; cosmology. The last column shows whether the simulation was post-processed with RT. In simulations without RT, the Hɪ distribution is obtained by using a fit to the photoionization rates as a function of density measured from simulations with RT.

Simulation	L $(h^{-1}\mathrm{Mpc})$	N	m_{b} $(h^{-1}\mathrm{M_\odot})$	m_{dm} $(h^{-1}\mathrm{M_\odot})$	ϵ_{com} $(h^{-1}\mathrm{kpc})$	ϵ_{prop} $(h^{-1}\mathrm{kpc})$	z_{end}	Cosmology	RT
L006N256	6.25	256^3	1.7×10^5	7.9×10^5	0.98	0.25	2	WMAP7	✓
L006N128	6.25	128^3	1.4×10^6	6.3×10^6	1.95	0.50	0	WMAP7	✓
L012N256	12.50	256^3	1.4×10^6	6.3×10^6	1.95	0.50	2	WMAP7	✓
L025N512	25.00	512^3	1.4×10^6	6.3×10^6	1.95	0.50	2	WMAP7	✗
L006N128-W3	6.25	128^3	1.4×10^6	6.3×10^6	1.95	0.50	2	WMAP3	✓
L025N512-W3	25.00	512^3	1.4×10^6	6.3×10^6	1.95	0.50	2	WMAP3	✗
L025N128-W3	25.00	128^3	8.7×10^7	4.1×10^8	7.81	2.00	0	WMAP3	✓
L050N256-W3	50.00	256^3	8.7×10^7	4.1×10^8	7.81	2.00	0	WMAP3	✓
L050N512-W3	50.00	512^3	1.1×10^7	5.1×10^7	3.91	1.00	0	WMAP3	✓
L100N512-W3	100.00	512^3	8.7×10^7	4.1×10^8	7.81	2.00	0	WMAP3	✗

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 $\Delta t = 1$ Myr $\left(\frac{L_{\text{box}}}{6.25\ h^{-1}\text{Mpc}}\right)\left(\frac{4}{1+z}\right)\left(\frac{128}{N_{\text{SPH}}}\right)$, where N_{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 two 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 neighbor is given by $\delta\mathcal{N}_{\text{abs},\nu} = \delta\mathcal{N}_{\text{in},\nu}[1 - \exp(-\tau(\nu))]$ where $\delta\mathcal{N}_{\text{in},\nu}$ and $\tau(\nu)$ are, respectively, the initial number of ionizing photons in the photon packet with frequency ν 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 $\tau(\nu) = \sum_\alpha \tau_\alpha(\nu)$ of the optical depth of each absorbing species (i.e., $\alpha \in \{\text{HI}, \text{HeI}, \text{HeII}\}$). Assuming that neighboring SPH particles have similar densities, we approximate the optical depth of each species using $\tau_\alpha(\nu) = \sigma_\alpha(\nu) n_\alpha d_{\text{abs}}$, where n_α is the number density of species, d_{abs} is the absorption distance between the SPH particle and its neighbor and $\sigma_\alpha(\nu)$ 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 neighbors. However, in order to decrease the amount of smoothing associated with this redistribution of photons, ViPs are assigned only 5 (instead of 48) SPH neighbors. We demonstrate convergence of our results with the number of ViP neighbors 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, $\Delta\mathcal{N}_{\text{abs},\alpha}(\nu)$. 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:

$$\Gamma_{\text{HI}} = \frac{\sum_\nu \Delta\mathcal{N}_{\text{abs,HI}}(\nu)}{\eta_{\text{HI}} \mathcal{N}_{\text{H}} \Delta t}, \tag{2.1}$$

where $\mathcal{N}_{\text{H}}$ is the total number of hydrogen atoms inside the SPH particle and $\eta_{\text{HI}} \equiv n_{\text{HI}}/n_{\text{H}}$ is the hydrogen neutral fraction.

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

$$\frac{d\eta_{\text{HI}}}{dt} = \alpha_{\text{HII}} n_e (1 - \eta_{\text{HI}}) - \eta_{\text{HI}}(\Gamma_{\text{HI}} + \Gamma_{e,\text{H}} n_e), \tag{2.2}$$

where n_e is the free electron number density, $\Gamma_{e,\mathrm{H}}$ is the collisional ionization rate and α_{HII} is the HII recombination rate. The differential equations which govern the ionization balance (e.g., equation 2.2) are solved using a subcycling time step, $\delta t = \min(f\tau_{\mathrm{eq}}, \Delta t)$ where $\tau_{\mathrm{eq}} \equiv \tau_{\mathrm{ion}}\tau_{\mathrm{rec}}/(\tau_{\mathrm{ion}} + \tau_{\mathrm{rec}})$, and f is a dimensionless factor which controls the integration accuracy (we set it to 10^{-3}), $\tau_{\mathrm{rec}} \equiv 1/\sum_i n_e\alpha_i$ and $\tau_{\mathrm{ion}} \equiv 1/\sum_i(\Gamma_i + n_e\Gamma_{e,i})$. 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 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 the recombination radiation is described in § 2.2.3 and § 2.2.4 below.

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 $n_{\mathrm{H}} > 0.1\,\mathrm{cm}^{-3}$ to $T_{\mathrm{ISM}} = 10^4$ 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., $n_{\mathrm{H}} < 10^{-3}\,\mathrm{cm}^{-3}$) or high temperatures (i.e., $T > 10^5$ K) is assumed to be in ionization equilibrium with the UVB and the collisional ionization rate (see Appendix A2). Typically, the neutral fraction of the box and the resulting H I CDDF do not evolve after 2-3 light-crossing times (the light-crossing time for the extended box with $L_{\mathrm{box}} = 6.25$ comoving h^{-1}Mpc is ≈ 7.5 Myr at z = 3).

2.2.3 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 radi-

---

[2]Other considerations prevent the use of arbitrarily large RT time steps. The RT assumes that species fractions and hence opacities do not evolve within a RT time step. This approximation becomes increasingly inaccurate with increasing RT time steps. Note that in this work, we iterate for ionization equilibrium which help to render our results robust against changes in the RT time step, as our convergence studies confirm.

Table 2.2: Hydrogen photoionization rate, absorption cross-section, equivalent gray approximation frequency and the self-shielding density threshold (i.e., based on equation 2.13) for three UVB models: Haardt & Madau (2001) (HM01; used in this work), Haardt & Madau (2012) (HM12) and Faucher-Giguère et al. (2009) (FG09) at different redshifts. For the calculation of the photoionization rate and absorption cross-sections, only photons with energies below 54.4 eV are taken into account, effectively assuming that more energetic photons are absorbed by He.

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

ation entering the simulation box from its sides. At the beginning of each RT step, we generate a large number of photon packets, N_{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 artifacts 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% of the box size from each side). These extended regions are excluded from the analysis, thereby removing the artifacts 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 ν_{HeII} are absorbed by helium, then the hydrogen photoionization rate can be written as:

$$
\begin{aligned}
\Gamma_{UVB} &= \int_{\nu_{HI}}^{\nu_{HeII}} 4\pi \, \frac{J_\nu}{h\nu} \, \sigma_{HI,\circ} \, d\nu \\
&\equiv \frac{4\pi \, \bar{\sigma}_{\nu_{HI}}}{h} \int_{\nu_{HI}}^{\nu_{HeII}} \frac{J_\nu}{\nu} \, d\nu,
\end{aligned}
\tag{2.3}
$$

where J_ν is the radiation intensity (in units $\mathrm{erg\ cm^{-2}\ s^{-1}\ sr^{-1}\ Hz^{-1}}$), ν_{HI} and ν_{HeII} are respectively the frequency at the Lyman-limit and the frequency at the HeII ionization edge, and $\sigma_{HI,\circ}$ is the neutral hydrogen absorption cross-section for ionizing photons. In the last equation we have defined the gray absorption cross-section,

$$
\bar{\sigma}_{\nu_{HI}} \equiv \frac{\int_{\nu_{HI}}^{\nu_{HeII}} J_\nu/\nu \, \sigma_{HI,\circ} \, d\nu}{\int_{\nu_{HI}}^{\nu_{HeII}} J_\nu/\nu \, d\nu}.
\tag{2.4}
$$

The radiation intensity is related to the photon energy density, u_ν,

$$
J_\nu = \frac{u_\nu \, c}{4\pi} = \frac{n_\nu \, h\nu \, c}{4\pi},
\tag{2.5}
$$

where n_ν is the number density of photons inside the box. Combining Equations 2.3-2.5 yields

$$
\Gamma_{UVB} = n_{\nu_{HI}} \, c \, \bar{\sigma}_{\nu_{HI}},
\tag{2.6}
$$

where $n_{\nu_{HI}}$ is the number density of ionizing photons inside the box. The total number of ionizing photons in the box is therefore given by

$$
n_{\nu_{HI}} L_{box}^3 = n_\gamma \, 6 \, N_{bg} \frac{L_{box}}{c \, \Delta t},
\tag{2.7}
$$

where n_γ 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 HI photoionization rate:

$$n_\gamma = \frac{\Gamma_{\rm UVB}\, L_{\rm box}^2\, \Delta t}{6\, \bar{\sigma}_{\nu_{\rm HI}}\, N_{\rm bg}},$$ (2.8)

We use the redshift-dependent UVB spectrum of Haardt & Madau (2001) to calculate $\Gamma_{\rm UVB}$ and $\bar{\sigma}_{\nu_{\rm HI}}$. 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 [3]. We note however that using more recent models for the UVB is not expected to change our main results. On can show that varying the UVB photoionization rate by a factor of 3, only changes the HI CDDF by less than 0.2 dex for LLSs (e.g., Altay et al., 2011). As shown in Table 2.2, the differences between photoionization rates in different UVB models are smaller that a factor of 3, particularly at $z > 1$, where the photoionization by the UVB is not subdominant (see §2.3.5). The variations in the adopted UVB model is even less important for systems with higher HI column densities (i.e., DLAs) which remain highly neutral for reasonable UVB models (e.g., Haardt & Madau, 2012; Faucher-Giguère et al., 2009).

To reduce the computational cost, we treat the multi-frequency problem in the gray approximation. In other words, we transport the UVB radiation using a single frequency bin, inside which photons are absorbed using the gray cross-section $\bar{\sigma}_{\nu_{\rm HI}}$ defined in equation 2.4. Note that the gray approximation ignores the spectral hardening of the radiation field that would occur in multifrequency 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 HI 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 redshifts are listed in Table 2.2 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 those models and the equivalent effective photon energy[4] of the background radiation changes only weakly with redshift, compared to the total photoionization rate.

---

[3]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 $n_{\rm H} \gtrsim 10^{-3} - 10^{-2}\, {\rm cm}^3$. 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 HI CDDF (Pontzen et al., 2008; McQuinn & Switzer, 2010; Altay et al., 2011).

[4]We defined the equivalent effective photon energy, $\mathcal{E}_{\rm eq}$, which corresponds to the absorption cross section, $\bar{\sigma}_{\nu_{\rm HI}}$, as: $\mathcal{E}_{\rm eq} \equiv 13.6\ {\rm eV}\ (\bar{\sigma}_{\nu_{\rm HI}}\, /\sigma_0)^{-1/3}$ where $\sigma_0 = 6.3 \times 10^{-18}\, {\rm cm}^2$.

2.2.4 Recombination radiation

Photons produced by the recombination of positive ions and electrons can also ionize the gas. If the recombining gas is optically thin, recombination radiation 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 recombination radiation 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 recombination radiation 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 TRAPHIC 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 prep. 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.

2.2.5 The HI column density distribution function

In order to compare the simulation results with observations, we compute the CDDF of neutral hydrogen, $f(N_{\mathrm{HI}}, z)$, a quantity that is somewhat straightforward to measure in QSO absorption line studies and is defined as the number of absorbers per unit column density, per unit absorption length, dX:

$$f(N_{\mathrm{HI}}, z) \equiv \frac{d^2 n}{dN_{\mathrm{HI}} dX} \equiv \frac{d^2 n}{dN_{\mathrm{HI}} dz} \frac{H(z)}{H_0} \frac{1}{(1+z)^2}. \tag{2.9}$$

We project the H I content of the simulation box along each axis onto a grid with 5000^2 or 10000^2 pixels (for the $128^3 - 256^3$ and 512^3 simulations, respectively)[5]. This is done using the actual kernels of SPH particles and for each of the three axes. The projection may merge distinct systems along the line of sight. However, for the small box sizes and high column densities with $N_{HI} > 10^{17}\,cm^{-2}$, which are the focus of this work, the chance of overlap between multiple absorbing systems in projection is negligible. Based on our numerical experiments, we expect that this projection effect starts to appear only at $N_{HI} < 10^{16}\,cm^{-2}$ if one uses a single slice for the projection of the entire *L50N512-W3* simulation box at $z = 3$. To make sure our results are insensitive to this effect, we use, depending on redshift, 25 or 50 slices for projecting the *L50N512-W3* simulation.

To produce a converged $f(N_{HI}, z)$ from simulations, one needs to use cosmological boxes that are large enough to capture the relevant range of overdensities. This is particularly demanding at very high H I column densities: for instance, van de Voort et al. (2012) showed that most of the gas with $N_{HI} > 10^{21}\,cm^{-2}$ resides in galaxies with halo masses $\gtrsim 10^{11}\,M_\odot$ which are relatively rare. As we show in Appendix B, the box size required to produce a converged H I CDDF up to $N_{HI} \sim 10^{22}\,cm^{-2}$ is $L \gtrsim 50$ comoving h^{-1}Mpc. Simulating RT in such a large volume is expensive. However, as we show in §2.3.4, at a given redshift the photoionization rates are fit very well by a function of the hydrogen number density. This relation is conserved with respect to both box size[6] and resolution and can therefore be applied to our highest resolution simulation (i.e., *L50N512-W3*), allowing us to keep the numerical cost tractable. Finally, since repeating the high-resolution simulations is expensive, we apply a redshift-independent correction which accounts for the difference between the WMAP year 3 parameters used for *L50N512-W3* simulation and the WMAP year 7 values. This is done by multiplying all the H I CDDFs produced based on the WMAP year 3 cosmology by the ratio between the H I CDDFs for *L25N512* and *L25N512-W3* at $z = 3$.

2.2.6   Dust and molecular hydrogen

Dust and star formation are highly correlated and infrared observations indicate that the prevalence of dusty galaxies follows the average star formation history of the Universe (e.g., Rahmati & van der Werf, 2011). Nevertheless, dust extinction is a physical processes that is not treated in our simulations. Assuming a constant dust-to-gas ratio, the typical dust absorption cross-section per atom is orders of magnitudes lower than the typical hydrogen absorption cross-section

---

[5]Using 5000^2 cells, the corresponding cell size is similar to the minimum smoothing length, and ~ 100 times smaller than the mean smoothing length, of SPH particles at $z = 3$ in the *L06N128* simulation.

[6]One should note that the box size can indirectly change the resulting photoionization rate profile. For instance, self-shielding can be affected by collisional ionization, which become stronger at lower redshifts and whose importance depends on the abundance of massive objects, which is more sensitive to the box size.
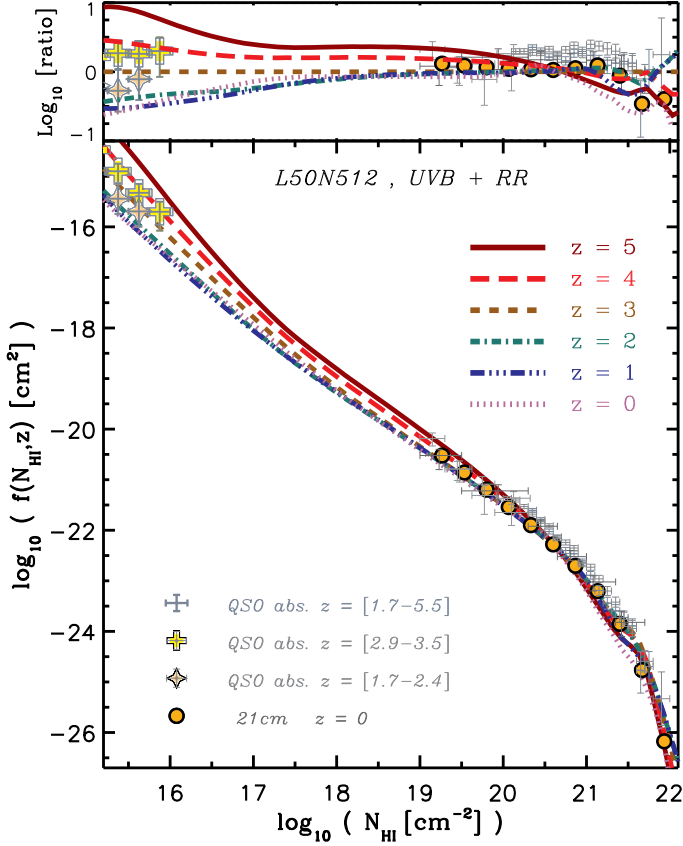
Figure 2.1: CDDF of neutral gas at different redshifts in the presence of the UVB and diffuse recombination radiation for *L50N512-W3*. A column density dependent amplitude correction has been applied to make the results consistent with WMAP year 7 cosmological parameters. The observational data points 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 colored 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 I distribution is magnified in the *right panel* and for clarity only the simulated H I CDDF of redshifts $z = 1$, 3 & 5 are shown. The *top-section* of each panel shows the ratio between the H I CDDFs at different redshifts and the H I CDDF at $z = 3$. The simulation results are in reasonably good agreement with the observations and, like the observations, show only a remarkably weak evolution for Lyman Limit and weak damped Lyα systems below $z = 3$.
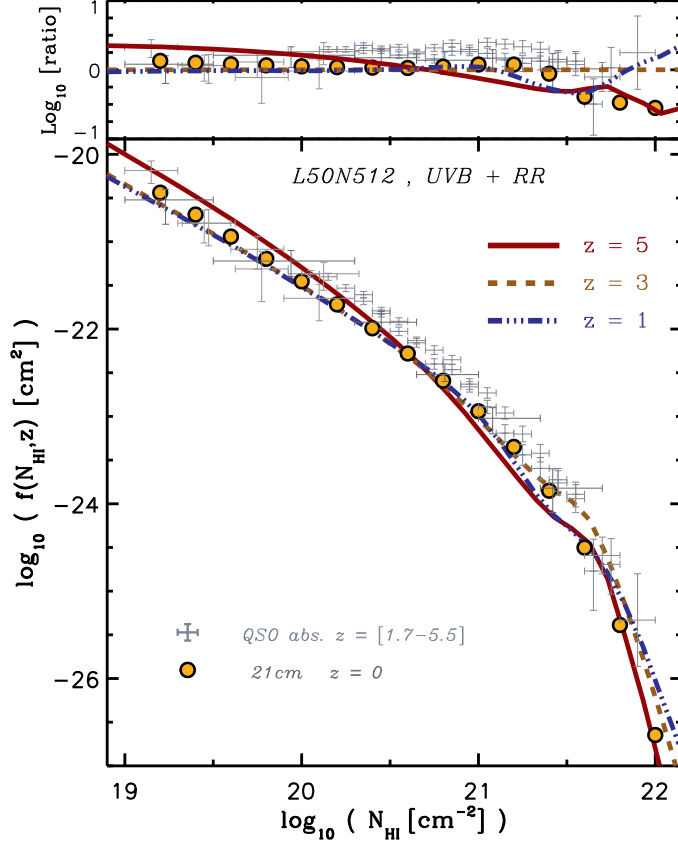
Figure 2.2: The high column density end of the HI distribution shown in Figure 2.1 is magnified here. For clarity, only the simulated HI CDDF of redshifts $z = 1$, 3 & 5 are shown. The *top-section* shows the ratio between the HI CDDFs at different redshifts and the HI CDDF at $z = 3$. The simulation results are in reasonably good agreement with the observations and, like the observations, show only a weak evolution for Lyman Limit and weak damped Lyα systems below $z = 3$.

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 et al., 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_{HI} systems may be related to the conversion of atomic hydrogen into H_2 (e.g., Schaye, 2001b; Krumholz et al., 2009; 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 HI 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 HI-H_2 relation is known to be sensitive to the dust content and hence to the metallicity (e.g., Schaye, 2001b, 2004; Krumholz et al., 2009a).

2.3   Results

In this section we report our findings based on various RT simulations which include UVB ionizing radiation and diffuse recombination radiation from ionized gas. As we demonstrate in §2.3.3, 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., *L06N128* or *L06N256*) which are computationally cheaper, to calculate the neutral hydrogen distribution in larger boxes. The last column of Table 2.1 indicates for which simulations this was done.

In the following, we will first present the predicted HI CDDF and compare it with observations. Next we discuss other aspects of our RT results and the effects of ionization by the UVB, recombination radiation and collisional ionization on the resulting HI distributions at different redshifts.

2.3.1   Comparison with observations

In Figure 2.1 & 2.2 we compare the simulation results with a compilation of observed HI CDDFs, after converting both to the WMAP year 7 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 $N_{HI} \gtrsim 10^{19}$ 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., $N_{HI} > 10^{17}$ cm^{-2}) the observations probing $0 < z < 5.5$ 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 ($\lesssim 50\%$) for 10^{17} cm$^{-2} < N_{HI} < 10^{21}$ cm^{-2} (i.e., Lyman limit systems, LLSs, and weak Damped Ly-α systems, DLAs) especially at $z \lesssim 3$.

The simulation predicts some variation with redshift for strong DLAs ($N_{HI} \gtrsim 10^{21}$ cm^{-2}). The abundance of strong DLAs in the simulations follows a similar redshift-dependent trend as the average star formation density in our simulations which peaks at $z \approx 2 - 3$ (Schaye et al., 2010). This result is consistent with the DLA evolution found by Cen (2012) in two zoomed simulations of a cluster and a void. One should, however, note that at very high column densities (e.g., $N_{HI} \gtrsim 10^{21.5}$ cm^{-2}) both observations and simulations are limited by small number statistics and the simulation results are more sensitive to the adopted feedback scheme (Altay et al. in prep.). Moreover, as we will show in Rahmati et al. (2013), including local stellar ionizing radiation can decrease the H I CDDF by up to ≈ 1 dex for $N_{HI} \gtrsim 10^{21}$ cm^{-2}, especially at redshifts $z \approx 2 - 3$ for which the average star formation activity of the Universe is near its peak.

At low column densities (i.e., $N_{HI} \lesssim 10^{17}$ cm^{-2}) the simulation results agree very well with the observations. This is apparent from the agreement between the simulated $f(N_{HI}, z)$ 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 Figure 2.1. The simulated $f(N_{HI}, z)$ 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-α forest observations at lower redshifts (Kim et al., 2002; Janknecht et al., 2006; Lehner et al., 2007; Prochaska & Wolfe, 2009; Ribaudo et al., 2011), as illustrated with the orange diamonds which correspond to $z \approx 2$ observations, in the top-left corner of Figure 2.1.

The evolution of the H 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 I column densities. However, as we show in §2.3.5, 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 I CDDFs at all redshifts are consistent with each other and the observations. However, as illustrated in Figure 2.2, there is a ≈ 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 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 is shown here, pro-
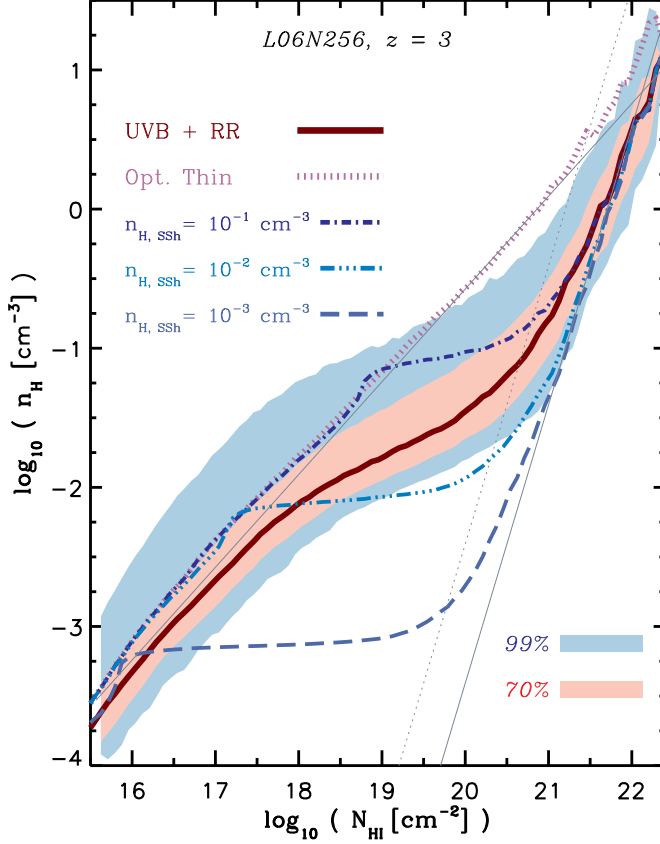
Figure 2.3: n_{HI}-weighted total hydrogen number density as a function of N_{HI}. The brown solid curve shows the RT results and the purple dotted curve shows the optically thin limit. Blue dot-dashed, dot-dot-dot-dashed and long dashed curves assume models with self-shielding density thresholds of $n_{H,SSh} = 10^{-1}$, 10^{-2} and 10^{-3} cm^{-3}, respectively. All of the above mentioned curves show the median of the n_{HI}-weighted total hydrogen number density at a given N_{HI}. The gray thin lines show the expected Jeans scaling relations for optically thin gas (equation 2.10; diagonal solid line) and for neutral gas (equation 2.11; steeper dotted line). A second solid line with the same slope expected from equation (2.11) 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_{HI}. The pink and blue shaded areas in the right panel indicate the 70% and 99% scatter respectively, while the solid curves shows the median for the RT result. All the other curves are also medians. This shows that the $n_H - N_{HI}$ relationship can be explained by the Jeans scaling and that the flattening in the CDDF is due to self-shielding.
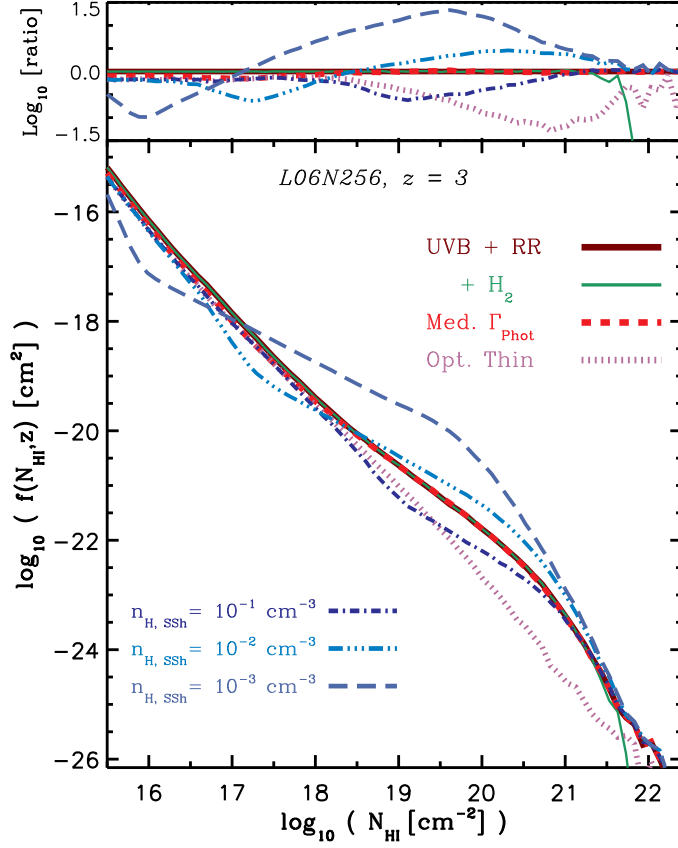
Figure 2.4: HI CDDF in the presence of the UVB and diffuse recombination radiation for simulation *L06N256*. Simulations shown with different curves are identical to those in Figure 2.3. In addition, the effect of H_2 formation is shown by the green solid curve which deviates from the brown solid curve at $N_{HI} \gtrsim 3 \times 10^{21}\,\mathrm{cm}^{-2}$. 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 §2.3.3 and Appendix A1). The *top-section* in the right panel shows the ratio between different HI CDDFs and the one resulting from the RT simulations.

duces a better match to the observations than a cosmology based on the WMAP 3 year results with smaller values for Ω_b and σ_8. This suggests that a higher value of σ_8 may explain the small discrepancy between the simulation results and the observations.

2.3.2   The shape of the HI CDDF

The shape of the HI 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 over-dense 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, 2001). Assuming that absorbers have universal baryon fractions (i.e., $f_g = \frac{\Omega_b}{\Omega_m}$) and typical temperatures of $T_4 \equiv (T/10^4 \text{ K}) \sim 1$ (i.e., collisional ionization is unimportant), one can calculate the total hydrogen column density (Schaye, 2001):

$$N_H \sim 1.6 \times 10^{21} \text{ cm}^{-2} \, n_H^{1/2} \, T_4^{1/2} \left(\frac{f_g}{0.17} \right)^{1/2}. \tag{2.10}$$

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, 2001):

$$N_{HI} \sim 2.3 \times 10^{13} \text{ cm}^{-2} \left(\frac{n_H}{10^{-5} \text{ cm}^{-3}} \right)^{3/2}$$

$$\times \, T_4^{-0.26} \, \Gamma_{-12}^{-1} \left(\frac{f_g}{0.17} \right)^{1/2}. \tag{2.11}$$

At high densities where the gas is nearly neutral, equation (2.10) provides a relation between N_{HI} and n_H. Equation (2.11) 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 Lyα 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 $\tau_{HI} = 1$:

$$N_{HI,SSh} \sim 4 \times 10^{17} \text{ cm}^{-2} \left(\frac{\bar{\sigma}_{\nu_{HI}}}{2.49 \times 10^{-18} \text{ cm}^2} \right)^{-1} \tag{2.12}$$

which can be used together with equation (2.11) to find the typical densities at

which the self-shielding begins (e.g., Furlanetto et al., 2005):

$$n_{\mathrm{H,SSh}} \sim 6.73 \times 10^{-3} \, \mathrm{cm}^{-3} \left(\frac{\bar{\sigma}_{\nu_{\mathrm{HI}}}}{2.49 \times 10^{-18} \, \mathrm{cm}^2} \right)^{-2/3}$$

$$\times \, T_4^{0.17} \, \Gamma_{-12}^{2/3} \left(\frac{f_{\mathrm{g}}}{0.17} \right)^{-1/3}. \tag{2.13}$$

These relations are compared with the n_{HI}-weighted total hydrogen number density as a function of N_{HI} in the *L06N256* simulation at $z = 3$ in Figure 2.3. The solid curve shows the median and the red (blue) shaded area represents the central 70% (90%) percentile. The diagonal gray solid line which converges with the simulation results at low column densities, shows equation (2.11) and the steeper gray dotted line which converges with the simulation results at high column densities is based on equation (2.10). The agreement between the expected slopes of the $n_{\mathrm{H}} - N_{\mathrm{HI}}$ 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[7].

As expected from equation (2.13), at low densities the gas is optically thin and follows the Jeans scaling relation of the highly ionized gas. At $n_{\mathrm{H}} \gtrsim 0.01 \, \mathrm{cm}^{-3}$ however, the relation between density and column density starts to deviate from equation (2.11) and approaches that of a nearly neutral gas. Consequently, for densities above the self-shielding threshold the HI column density increases rapidly over a narrow range of densities, leading to a flattening in the $n_{\mathrm{H}} - N_{\mathrm{HI}}$ relation and in the resulting HI CDDF at $N_{\mathrm{HI}} \gtrsim 10^{18} \, \mathrm{cm}^{-2}$ (see Figure 2.4). The results from the RT simulation deviate from the magenta dotted lines, which are obtained assuming optically thin gas, at $N_{\mathrm{HI}} \gtrsim 4 \times 10^{17} \, \mathrm{cm}^{-2}$. As the dotted line in Figure 2.4 shows, in the absence of self-shielding, the slope of $f(N_{\mathrm{HI}}, z) \propto N_{\mathrm{HI}}^{\beta}$ is constant all the way up to DLAs at $\beta_{\mathrm{Ly}\alpha} \approx -1.6$. However, because of self-shielding, the HI CDDF flattens to $\beta_{\mathrm{LLS}} \approx -1.1$ at $10^{18} \, \mathrm{cm}^{-2} \lesssim N_{\mathrm{HI}} \lesssim 10^{20} \, \mathrm{cm}^{-2}$ in the RT simulation (solid curve). These predicted slopes are in excellent agreement with the latest observational constraints of $\beta_{\mathrm{Ly}\alpha} \gtrsim -1.6$ for $10^{15} \, \mathrm{cm}^{-2} < N_{\mathrm{HI}} < 10^{17} \, \mathrm{cm}^{-2}$ to $\beta_{\mathrm{LLS}} \approx -1$ in the LLS regime (O'Meara et al., 2013). We also note that $\beta_{\mathrm{Ly}\alpha} \gtrsim -1.6$ 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 $n_{\mathrm{H}} \gtrsim 0.1 \, \mathrm{cm}^{-3}$ the gas is nearly neutral and the Jeans scaling in equation (2.10) controls the $n_{\mathrm{H}} - N_{\mathrm{HI}}$ relation. Consequently, the rate at which N_{HI} responds to changes in n_{H} slows down, causing a steepening in the resulting $f(N_{\mathrm{HI}}, z)$ in the DLA range (i.e., $N_{\mathrm{HI}} \gtrsim 10^{21} \, \mathrm{cm}^{-2}$). However, as the thick solid

---

[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 gray dotted line which is based on equation (2.10) shows, the simulated N_{HI} for a given n_{H} is ≈ 0.5 dex higher than implied by the Jeans scaling for the nearly neutral case (i.e., steep, gray solid line).

curve in Figure 2.4 illustrates, the slope of $f(N_{HI}, z)$ remains constant for $N_{HI} = 10^{21} - 10^{22}$ cm^{-2}. This is in contrast with observed trends indicating a sharp cut-off at $N_{HI} \gtrsim 3 \times 10^{21}$ cm^{-2} (Prochaska et al., 2010; O'Meara et al., 2013; but see Noterdaeme et al., 2012). At those column densities a large fraction of hydrogen is expected to form H_2 molecules and be absent from HI observations (Schaye, 2001b; Krumholz et al., 2009; Altay et al., 2011). As the thin solid line in Figure 2.4 shows, accounting for H_2 using the empirical relation between 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 sharply (Noterdaeme et al., 2012), then this may imply that 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 HI column densities (Altay et al. in prep.).

To first order, one can mimic the effect of RT by assuming gas with $n_H < n_{H,SSh}$ to be optically thin (i.e., Case A recombination) and gas with $n_H > n_{H,SSh}$ to be fully neutral. Simulations with three different self-shielding density thresholds are shown in Figure 2.3 & 2.4. The dot-dashed, dot-dot-dot-dashed and long dashed curves correspond to $n_{H,SSh} = 10^{-1}$, 10^{-2} and 10^{-3} cm^{-3}, respectively. Although all of these simulations predict the flattening of $f(N_{HI}, z)$, they produce a transition between optically thin and neutral gas that is too steep. In contrast, the RT results show a transition between highly ionized and highly neutral gas that is more gradual, as observed.

2.3.3 Photoionization rate as a function of density

Figure 2.5 illustrates the RT results for neutral fractions and photoionization rates as a function of density in the presence of UVB radiation and diffuse recombination radiation 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 $n_H \gtrsim 0.01$ cm^{-3}, consistent with equation (2.13) and previous studies (Tajiri & Umemura, 1998; Razoumov et al., 2006; Faucher-Giguère et al., 2010; Nagamine et al., 2010; Fumagalli et al., 2011; Altay 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 §2.3.5, 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
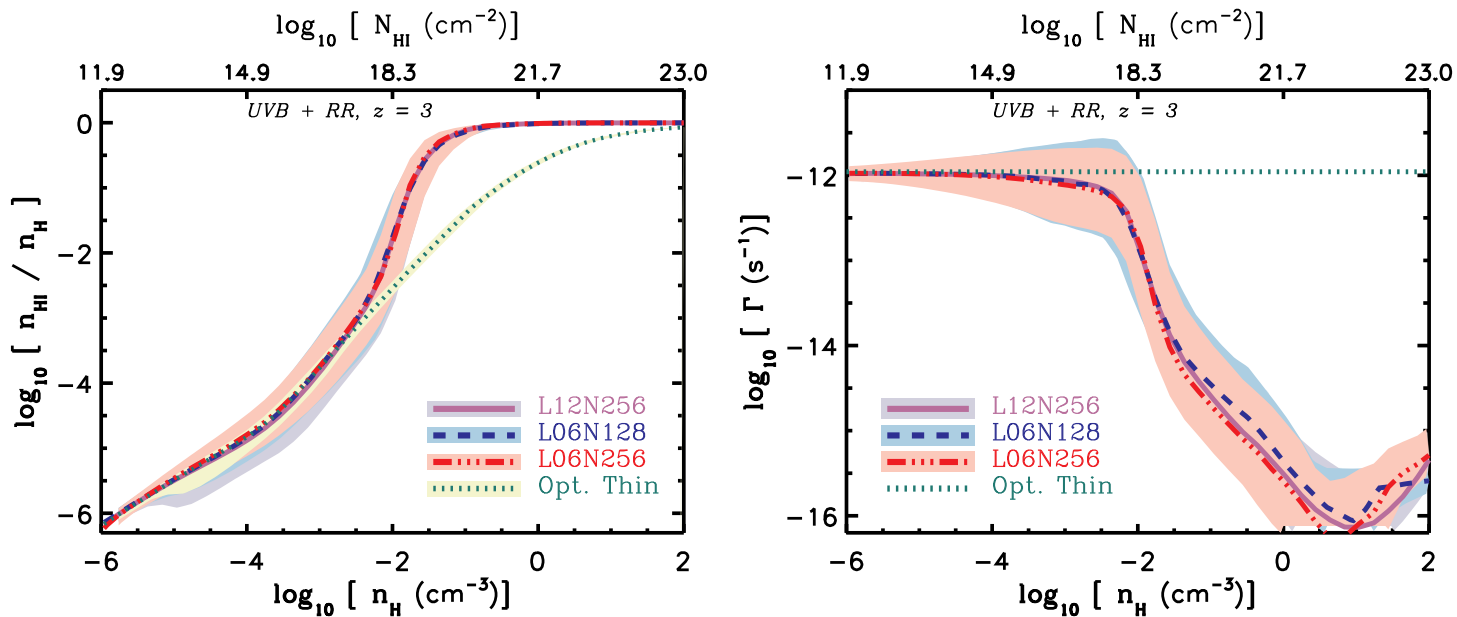
Figure 2.5: The hydrogen neutral fraction (*left*) and the photoionization rate (*right*) as a function of hydrogen number density do not change by varying the simulation box size or mass resolution. This is shown for different simulations at z = 3 in the presence of the UVB and recombination radiation. Purple solid, blue dashed and red dot-dashed lines show, respectively, the results for *L12N256*, *L06N128* and *L06N256*. The green dotted line indicates the results for the *L06N128* simulation if the gas is assumed to be optically thin to the UVB radiation (i.e., no RT calculation is performed). The deviation between the optically thin hydrogen neutral fractions and RT results at $n_H \gtrsim 10^{-2}\,\mathrm{cm}^{-3}$ shows the impact of self-shielding. The lines show the medians and the shaded areas indicate the $15\% - 85\%$ percentiles. At the top of each panel we show H$_I$ column densities corresponding to each density.

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 ~ 2 dex from its maximum value which is caused by the RR photoionization. These features are captured by the following fitting formula:

$$\frac{\Gamma_{\text{Phot}}}{\Gamma_{\text{UVB}}} = 0.98 \left[1 + \left(\frac{n_{\text{H}}}{n_{\text{H,SSh}}} \right)^{1.64} \right]^{-2.28}$$
$$+ 0.02 \left[1 + \frac{n_{\text{H}}}{n_{\text{H,SSh}}} \right]^{-0.84}, \qquad (2.14)$$

where Γ_{UVB} is the background photoionization rate and Γ_{Phot} is the total photoionization rate. Moreover, the self-shielding density threshold, $n_{\text{H,SSh}}$, is given by equation (2.13) and is thus a function of Γ_{UVB} and $\bar{\sigma}_{\nu_{\text{HI}}}$ 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 Γ_{UVB} and $\bar{\sigma}_{\nu_{\text{HI}}}$ in order to determine the corresponding $n_{\text{H,SSh}}$ from (2.13) (see also Table 2.2). Then, after using equation (2.14) 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 (2.14) are only accurate for photoionization dominated cases. As we show in §2.3.5, at $z \sim 0$ the collisional ionization rate is greater than the total photoionization rate around the self-shielding density threshold. Consequently, equation (2.13) 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 (2.14) reproduce the $f(N_{\text{HI}}, z)$ accurately to within 10% 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).
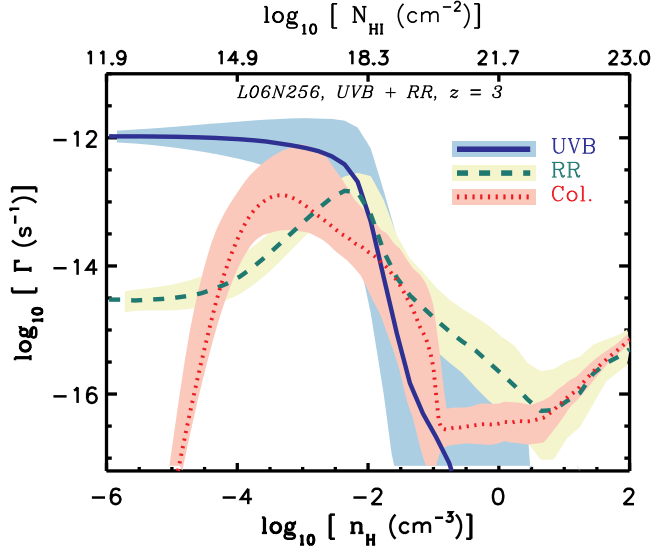
Figure 2.6: Ionization rates due to different sources of ionization as a function of hydrogen number density. Blue solid, green dashed and red dotted curves show, respectively, the UVB photoionization rate, the recombination radiation photoionization rate and the collisional ionization rate. The curves show the medians and the shaded areas around the medians indicate the $15\% - 85\%$ percentiles. HI column densities corresponding to each density are shown along the top x-axis. While the UVB is the dominant source of ionization below the self-shielding (i.e., $n_H \lesssim 10^{-2}\,\mathrm{cm}^{-3}$), recombination radiation dominates the ionization at higher densities.

2.3.4 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 HI, 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 §2.3.5.

The total photoionization rate profiles shown in the right panel of Figure 2.5 are almost flat at low densities and decrease with increasing density, starting at densities $n_H \sim 10^{-4}\,\mathrm{cm}^{-3}$. Just below $n_H = 10^{-2}\,\mathrm{cm}^{-3}$ self-shielding causes a sharp drop, but the fall-off becomes shallower for $n_H > 10^{-2}\,\mathrm{cm}^{-3}$ and the photoionization rate starts to increase at $n_H > 10\,\mathrm{cm}^{-3}$. As shown in Figure 2.6, 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 artifact of the imposed temperature for ISM particles (i.e., $T = 10^4$ K) which produces a rising collisional ionization rate with increasing density. As the comparison between the UVB and RR photoionization profiles shows (see Figure 2.6), RR only starts to dominate the
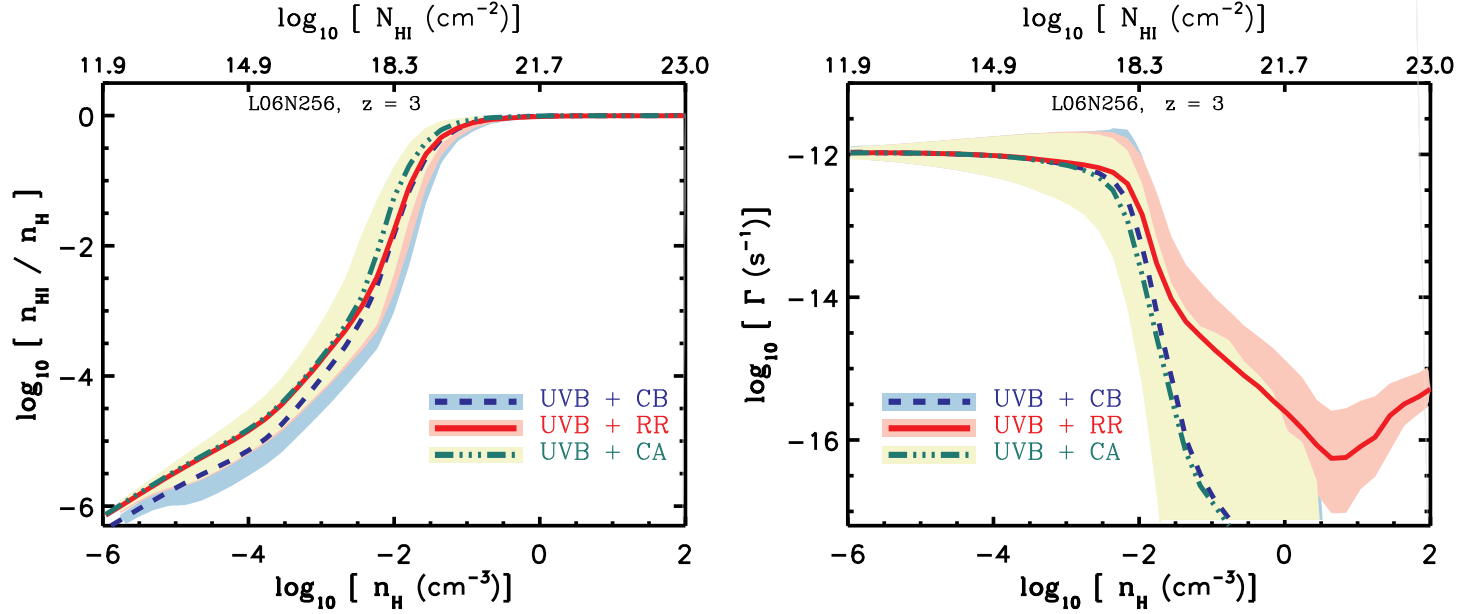
Figure 2.7: The hydrogen neutral fractions (*left*) and the UVB photoionization rate profiles (*right*) as a function of density in RT simulations with different models for recombination radiation for the *L06N256* simulation at z = 3. The red dashed curve shows the reference simulation where recombination radiation is modeled self-consistently. The blue solid and green dot-dashed curves show simulations in which recombination radiation is substituted by the use of Case A and Case B recombination rates, respectively. The curves show the medians and the shaded areas around the medians indicate the 15% − 85% percentiles. H I column densities corresponding to each density are shown along the top x-axes. The effect of recombination radiation on the hydrogen neutral fractions is similar to the use of Case A recombination at low densities (i.e., $n_{\rm H} \lesssim 10^{-3}\,{\rm cm}^{-3}$) and to the use of Case B recombination at higher densities (i.e., $n_{\rm H} \gtrsim 10^{-1}\,{\rm cm}^{-3}$). However, recombination radiation can penetrate into the self-shielded regions, an effect that is not captured by the use of Case B recombination.

total photoionization rate at $n_H > 10^{-2}\,\text{cm}^{-3}$, 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 Hi content of high-density gas by $\approx 20\%$. 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 $\lesssim 20\%$.

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 ~ 100 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 Figure 2.7 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 modeled by switching from Case A to Case B recombination at $n_H \sim n_{H,\text{SSh}}$ (e.g., Altay et al., 2011; McQuinn et al., 2011).

2.3.5 Evolution

The general trends in the profile of the photoionization rates with density and their influences on the distribution of Hi are not very sensitive to redshift. However, as shown in table 2.2, 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 Figure 2.8, the evolution of the hydrogen neutral fraction is il-
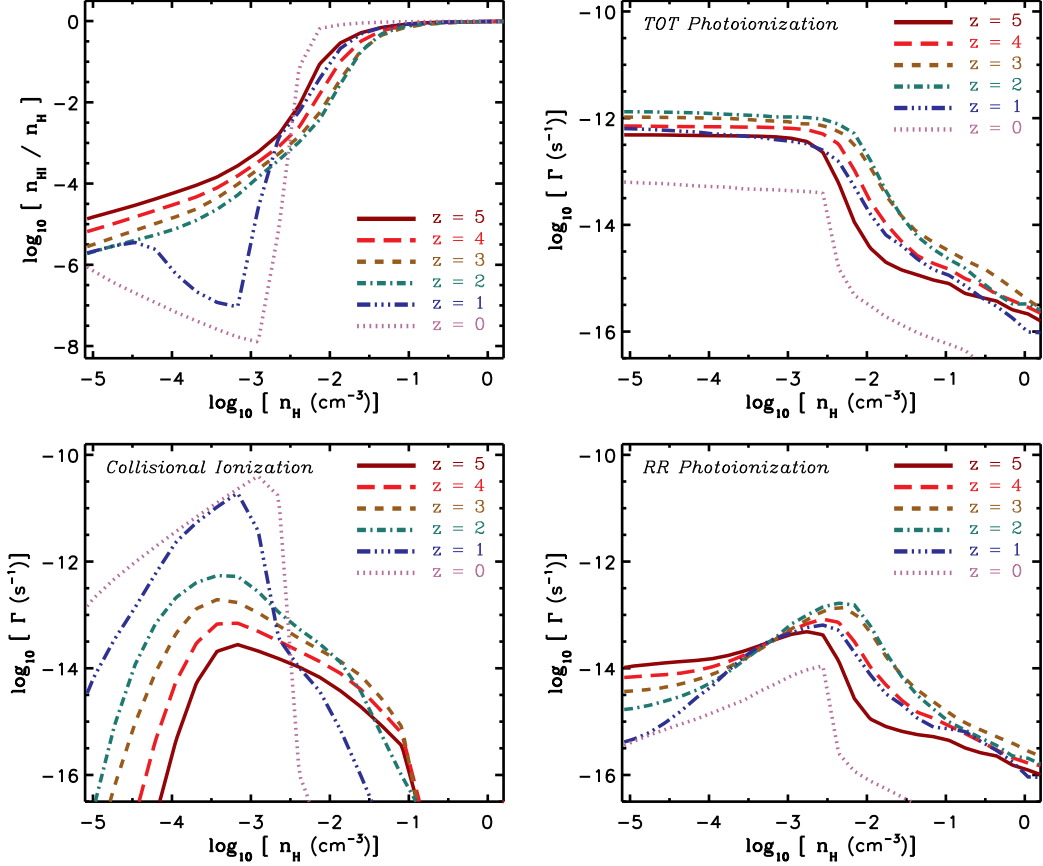
Figure 2.8: Evolution of the hydrogen neutral fraction profile (*top-left*) and various ionization rates as a function of density. *Top-right*, *bottom-left* and *bottom-right* panels show, respectively, the total photoionization rates, collisional ionization rates and RR photoionization rates. All the HI 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 become 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 $n_H \lesssim 10^{-3} \, \mathrm{cm}^{-3}$.

lustrated for the *L50N512-W3* simulation. As discussed in §2.3.3 and Appendix A1, since the photoionization rate profiles are converged with box size and resolution, we apply the profiles derived from a 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. Figure 2.8 shows that the neutral fraction profiles are similar in shape at high redshifts but that at $z \leq 1$ 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 Figure 2.8. At $z \geq 2$ and for $n_{\mathrm{H}} < 10^{-2}\,\mathrm{cm}^{-3}$, the collisional ionization rate is not high enough to compete with the UVB photoionization rate. At lower redshifts and for number densities $n_{\mathrm{H}} \lesssim 3 \times 10^{-3}\,\mathrm{cm}^{-3}$, on the other hand, collisional ionization dominates[9]. Indeed, the median collisional ionization rates are more than 100 times higher than the UVB photoionization rate at densities around the expected self-shielding thresholds. Collisional ionization therefore helps the UVB ionizing photons to penetrate to higher densities without being significantly absorbed. As a result, self-shielding starts at densities higher than expected from equation (2.13). The signature of collisional ionization on the hydrogen neutral fraction is more dramatic at low densities and partly compensates for the lower UVB intensity at z = 0. This results in a flattening of $f(N_{\mathrm{HI}}, z)$ at column densities $N_{\mathrm{HI}} \lesssim 10^{16}\,\mathrm{cm}^{-2}$ as shown in Figure 2.1.

As mentioned above, at low redshifts (e.g., $z = 0$) the collisional ionization rate peaks at densities higher than the expected self-shielding threshold against the UVB. As a result, the total photoionization rate falls off rapidly together with the drop in the collisional ionization rate. Therefore, the drop in the hydrogen ionized fraction, and hence the resulting free electron density, is much sharper at lower redshifts. This causes a steeper high-density fall-off in the collisional ionization rate as shown in the bottom-left panel of Figure 2.8.

The differences between the total photoionization rates at different redshifts shown in the top-right panel of Figure 2.8, are caused by the evolution of the UVB intensity and its hardness, which affects the self-shielding density thresholds (see equation 2.13). 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 Figure 2.8).

The filled circles in Figure 2.9 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 2.13) is also shown (green dotted line). The peaks in the RR photoionization rate in RT

---

[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.

[9]We note that at redshift $z \lesssim 1$ box sizes $L_{\mathrm{Box}} \gtrsim 25$ comoving $h^{-1}\mathrm{Mpc}$ are required to fully capture the large-scale accretion shocks and to produce converged collisional ionization rates.
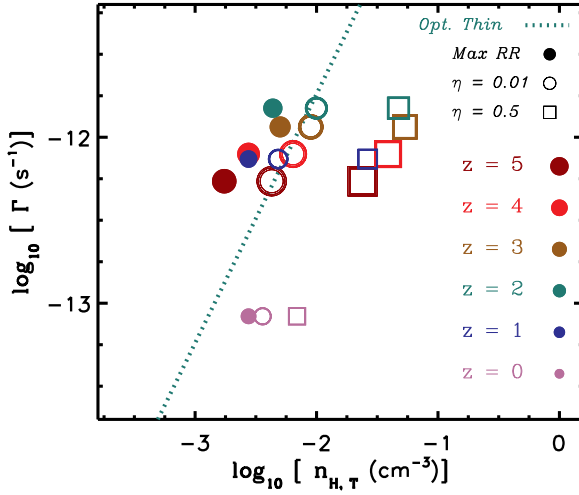
Figure 2.9: The photoionization from recombination radiation peaks at the expected self-shielding density and smooths the transition between highly ionized gas and self-shielded gas. This is illustrated by showing different characteristic densities for various redshifts. Filled circles show the density at which recombination radiation peaks, while open circles and squares show the H I number density corresponding to hydrogen neutral fractions of 10^{-2} and 0.5, respectively. The self-shielding density threshold for a given photoionization rate expected from the Jeans scaling argument (equation 2.13) is also indicated by the green dotted line. The Jeans scaling argument works well, except at $z = 0$ when collisional ionization is important.

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 (2.13).

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 median hydrogen neutral fractions are 10^{-2} and 0.5 are shown in Figure 2.9 with open circles and squares, respectively. The densities at which the hydrogen neutral fraction is 10^{-2} are slightly higher than the densities at which the RR photoionization rate peaks (filled circles). The evolution agrees with the trend expected from the Jeans scaling argument and the self-shielding density. The exception is again $z = 0$, where the large collisional ionization rate at number densities $n_{\mathrm{H}} \sim 10^{-3} - 10^{-2}\,\mathrm{cm}^{-3}$ shifts the transition to neutral fraction of 0.5 to densities that are ≈ 1 dex higher. However, the relation between the photoionization rate and the density still follows the slope expected from equation (2.13).

It is interesting to note that the UVB spectral shape at $z = 4$ is slightly harder than at $z = 1$ while the UVB intensities at these two redshifts are similar. This

results in a deeper penetration of ionizing photons at $z = 4$. Consequently, the densities corresponding to the indicated neutral fractions (i.e., 10^{-2} and 0.5) at $z = 1$ are lower than their counterparts at $z = 4$.

2.4 Conclusions

We combined a set of cosmological hydrodynamical simulations with an accurate RT simulation of the UVB radiation to compute the H I column density distribution function and its evolution. We ignored the effect of local sources of ionizing radiation, but we did include a self-consistent treatment of recombination radiation.

Our RT results for the distribution of photoionization rates at different densities are converged with respect to the simulation box size and resolution. Therefore, the resulting photoionization rate can be expressed as a function of the hydrogen density and the UVB. We provided a fit for the median total photoionization rate as a function of density that can be used with any desired UVB model to take into account the effect of H I self-shielding in cosmological simulations without the need to perform RT.

The CDDF, $f(N_{HI}, z)$, 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 $f(N_{HI}, z)$ function for a wide range of H I column densities. At low H I column densities, the CDDF is a steep function which decreases with increasing N_{HI} before it flattens at $N_{HI} \gtrsim 10^{18}\,\mathrm{cm}^{-2}$ due to self-shielding. At $N_{HI} \gtrsim 10^{21}\,\mathrm{cm}^{-3}$ on the other hand, $f(N_{HI}, z)$ is determined mainly by the intrinsic distribution of total hydrogen and the H_2 fraction.

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

The high H I column density end of the predicted $f(N_{HI}, z)$ evolves only weakly from $z = 5$ to $z = 0$, consistent with observations. In the Lyman limit range of the distribution function, the slope of $f(N_{HI}, z)$ 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 $f(N_{HI}, z)$. In contrast, the number of absorbers with lower H I column densities (i.e., the Lyα 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 recombination radiation 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., 2011; Goerdt et al., 2012). For instance, the difference in the gas density at which hydrogen is highly ionized (i.e., $n_{HI}/n_H \lesssim 0.01$) and the density at which gas is highly neutral (i.e., $n_{HI}/n_H \gtrsim 0.5$) is more than one order of magnitude (see Figure 2.9). As a result, assuming a sharp self-shielding density threshold at the density for which the optical depth of ionizing photons is ~ 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 gray 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 $f(N_{HI}, z)$ by 0.5-1 dex.

Acknowledgments

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, ApJL, 737, L37

Blitz, L., & Rosolowsky, E. 2006, ApJ, 650, 933

Cen, R., Ostriker, J. P., Prochaska, J. X., & Wolfe, A. M. 2003, ApJ, 598, 741

Cen, R. 2012, ApJ, 748, 121

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., et al. 2012, MNRAS, 420, 2799

Erkal, D., Gnedin, N. Y., & Kravtsov, A. V. 2012, arXiv:1201.3653

Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416

Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., Hernquist, L., & Zaldarriaga, M. 2010, ApJ, 725, 633

Friedrich, M. M., Mellema, G., Iliev, I. T., & Shapiro, P. R. 2012, MNRAS, 2385

Fumagalli, M., Prochaska, J. X., Kasen, D., et al. 2011, MNRAS, 418, 1796

Furlanetto, S. R., Schaye, J., Springel, V., & Hernquist, L. 2005, ApJ, 622, 7

Gardner, J. P., Katz, N., Hernquist, L., & Weinberg, D. H. 1997, ApJ, 484, 31

Gardner, J. P., Katz, N., Hernquist, L., & 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, arXiv:1205.2021

Haardt F., Madau P., 2001, in Clusters of Galaxies and the High Redshift Universe Observed in X-rays, Neumann D. M., Tran J. T. V., eds.

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, ApJL, 457, L57

Kim, T.-S., Carswell, R. F., Cristiani, S., D'Odorico, S., & Giallongo, 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, ApJL, 701, L12

Lehner, N., Savage, B. D., Richter, P., et al. 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, ApJL, 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, ApJL, 725, L219

Noterdaeme, P., Petitjean, P., Ledoux, C., & Srianand, R. 2009, A&A, 505, 1087

Noterdaeme, P., Petitjean, P., Carithers, W. C., et al. 2012, arXiv:1210.1213

O'Meara, J. M., Prochaska, J. X., Burles, S., et al. 2007, ApJ, 656, 666

O'Meara, J. M., Prochaska, J. X., Worseck, G., Chen, H.-W., & Madau, P. 2012, arXiv:1204.3093

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei, 2nd. ed. by D.E. Osterbrock and G.J. Ferland. Sausalito, CA: University Science Books, 2006,

Pawlik, A. H., & Schaye, J. 2008, MNRAS, 389, 651

Pawlik, A. H., & Schaye, J. 2011, MNRAS, 412, 1943

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

Pontzen, A., Governato, F., Pettini, M., 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, ApJL, 705, L113

Prochaska, J. X., O'Meara, J. M., & Worseck, G. 2010, ApJ, 718, 392

Rahmati, A., & van der Werf, P. P. 2011, MNRAS, 418, 176

Rahmati, A., Schaye, J., Pawlik, A. H., & Raičević, M. 2013, MNRAS, 431, 2261

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

Robbins, D. 1978, Amer. Math. Monthly 85, 278

Schaye, J. 2001a, ApJ, 559, 507

Schaye, J. 2001b, ApJL, 562, 95

Schaye, J. 2004, ApJ, 609, 667

Schaye, J. 2006, ApJ, 643, 59

Schaye, J., & Dalla Vecchia, C. 2008, MNRAS, 383, 1210

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536

Springel, V. 2005, MNRAS, 364, 1105

Tepper-García, T., Richter, P., Schaye, J., Booth, C. M.; Dalla Vecchia, C.; Theuns, T. 2012, MNRAS, 425, 1640

Tytler, D. 1987, ApJ, 321, 49

Tajiri, Y., & Umemura, M. 1998, ApJ, 502, 59

Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, MNRAS, 301, 478

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., & Smith, B. D. 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. 2011, arXiv:1112.5691

Zheng, Z., & Miralda-Escudé, J. 2002, ApJL, 568, L71

Zwaan, M. A., van der Hulst, J. M., Briggs, F. H., Verheijen, M. A. W., & Ryan-Weber, E. V. 2005, MNRAS, 364, 1467

Appendix A: Photoionization rate as a function of density

A1: Replacing the RT simulations with a fitting function

In §2.3.3 we demonstrated that the median of the simulated relation between the total photoionization rate, Γ_{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.

The left panel of Figure 2.10 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 $N_{\mathrm{HI}} \gtrsim 10^{18}\,\mathrm{cm}^{-2}$. 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 $f(N_{\mathrm{HI}}, z)$ is slightly overproduced compared to the full RT result at nearly all H I column densities.

We exploit the insensitivity of the shape of the Γ_{Phot}-density relation to the redshift, and propose the following fit to the photoionization rate, Γ_{Phot},

$$\frac{\Gamma_{\mathrm{Phot}}}{\Gamma_{\mathrm{UVB}}} = (1-f)\left[1+\left(\frac{n_{\mathrm{H}}}{n_0}\right)^{\beta}\right]^{\alpha_1} + f\left[1+\frac{n_{\mathrm{H}}}{n_0}\right]^{\alpha_2}, \qquad (2.15)$$

where Γ_{UVB} is the photoionization rate due to the ionizing background, and n_0, α_1, α_2, and β are parameters of the fit. The best-fit values of these parameters are listed in Table 2.3 and the photoionization rate-density relations they produce are compared with the RT simulations at redshifts $z = 0$ and $z = 4$ in Figure 2.11. At all redshifts, the best-fit value of n_0 is almost identical to the self-shielding density threshold, $n_{\mathrm{H,SSh}}$, defined in equation (2.13), and the characteristic slopes
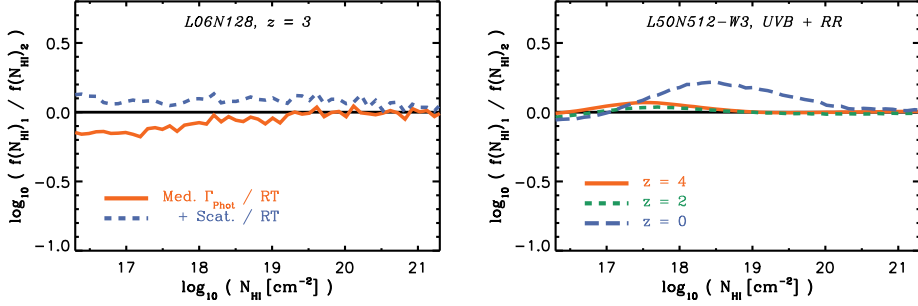
Figure 2.10: *Left:* The ratio between the HI CDDF calculated using the RT based $\Gamma_{\mathrm{Phot}} - n_{\mathrm{H}}$ relationship and the actual RT results for the *L06N128* simulation in the presence of the UVB and diffuse recombination radiation, at $z = 3$. The orange solid line shows the result 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:* HI CDDFs calculated using the $\Gamma_{\mathrm{Phot}} - n_{\mathrm{H}}$ fitting function (i.e., equation 2.14) are compared to the HI CDDFs for which the actual Γ_{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-fit parameters shown in Table 2.4. All the CDDFs are for the *L50N512-W3* simulation and in the presence of the UVB and diffuse recombination radiation.

of the photoionization rate-density relation are similar. This suggest that one can find a single set of best-fit values to reproduce the RT results at $z \gtrsim 1$. The corresponding best-fit parameter values are (see also equation 2.14) $\alpha_1 = -2.28 \pm 0.31$, $\alpha_2 = -0.84 \pm 0.11$, $n_0 = (1.003 \pm 0.005) \times n_{\mathrm{H,SSh}}$, $\beta = 1.64 \pm 0.19$ and $f = 0.02 \pm 0.0089$.

In the right panel of Figure 2.10, the ratio between the HI CDDF calculated using the fitting function presented in equation (2.14) and the RT based HI CDDF (i.e., calculated using the median of the photoionization rate-density relation in the RT simulations) is shown for the *L50N512-W3* 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 §2.3, 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 2.4.

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

Table 2.3: The best-fit parameters for equation (2.15) at different redshifts based on RT results in the *L06N128* simulation.

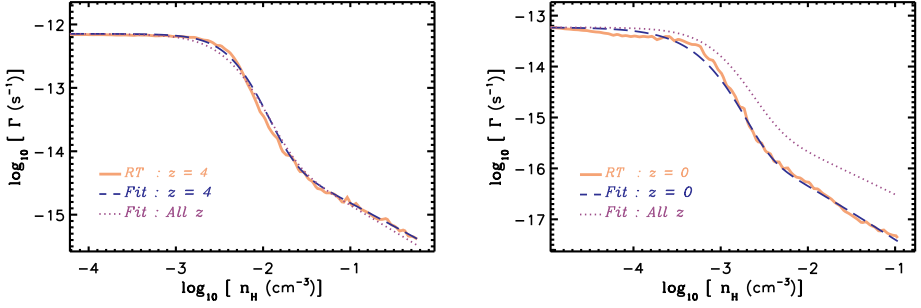| Redshift	$\log[n_0]$ (cm^{-3})	α_1	α_2	β	$1-f$
z = 1-5	$\log[n_{\mathrm{H,SSh}}]$	-2.28	-0.84	1.64	0.98
$z = 0$	-2.94	-3.98	-1.09	1.29	0.99
$z = 1$	-2.29	-2.94	-0.90	1.21	0.97
$z = 2$	-2.06	-2.22	-1.09	1.75	0.97
$z = 3$	-2.13	-1.99	-0.88	1.72	0.96
$z = 4$	-2.23	-2.05	-0.75	1.93	0.98
$z = 5$	-2.35	-2.63	-0.57	1.77	0.99

Figure 2.11: Comparisons between the total photoionization rates as a function of density in the *L06N128* simulation. Photoionization rates based on the RT simulations and best-fit 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 2.15 and Table 2.3) is shown with the blue dashed curve and the best fit to the RT results at $z = 1 - 5$ (equation 2.14) 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-fit parameters shown in Table 2.4 for $z = 0$.

Table 2.4: The best-fit parameters for equation (2.15) at $z = 0$ based on RT results for the *L50N512* 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 $50\ h^{-1}\mathrm{Mpc}$ results in converged collisional ionizations.

Redshift	$\log[n_0]$ (cm^{-3})	α_1	α_2	β	$1-f$
$z = 0$	-2.56	-1.86	-0.51	2.83	0.99

with the total number of recombinations per unit time per unit volume, we obtain

$$n_{\mathrm{HI}}\,\Gamma_{\mathrm{TOT}} = \alpha_{\mathrm{A}}\,n_{\mathrm{e}}\,n_{\mathrm{HII}}, \tag{2.16}$$

where n_{HI}, n_{e} and n_{HII} are the number densities of neutral hydrogen atoms, free electrons and protons, respectively. Γ_{TOT} is the total ionization rate per neutral hydrogen atom and α_{A} is the Case A recombination rate[10] for which we use the fitting function given by Hui & Gnedin (1997):

$$\alpha_{\mathrm{A}} = 1.269 \times 10^{-13}\, \frac{\lambda^{1.503}}{\left(1 + (\lambda/0.522)^{0.47}\right)^{1.923}}\ \mathrm{cm^3\,s^{-1}}, \tag{2.17}$$

where $\lambda = 315614/T$.

Defining the hydrogen neutral fraction as the ratio between the number densities of neutral hydrogen and total hydrogen, $\eta = n_{\mathrm{HI}}/n_{\mathrm{H}}$, and ignoring helium (which is an excellent approximation, see Appendix D2), we can rewrite equation (2.17) as:

$$\eta\,\Gamma_{\mathrm{TOT}} = \alpha_{\mathrm{A}}\,(1-\eta)^2\,n_{\mathrm{H}}. \tag{2.18}$$

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

$$\Gamma_{\mathrm{TOT}} = \Gamma_{\mathrm{Phot}} + \Gamma_{\mathrm{Col}}, \tag{2.19}$$

where $\Gamma_{\mathrm{Col}} = \Lambda_{\mathrm{T}}\,(1-\eta)\,n_{\mathrm{H}}$. The photoionization rate can be expressed as a function of density using equation (2.14). For Λ_{T}, which depends only on temperature, we use a relation given in Theuns et al. (1998):

$$\Lambda_{\mathrm{T}} = 1.17 \times 10^{-10}\, \frac{T^{1/2}\exp(-157809/T)}{1 + \sqrt{T/10^5}}\ \mathrm{cm^3\,s^{-1}}. \tag{2.20}$$

We can now rearrange equation (2.18) as a quadratic equation:

$$A\,\eta^2 - B\,\eta + C = 0, \tag{2.21}$$

with $A = \alpha_{\mathrm{A}} + \Lambda_{\mathrm{T}}$, $B = 2\alpha_{\mathrm{A}} + \frac{\Gamma_{\mathrm{Phot}}}{n_{\mathrm{H}}} + \Lambda_{\mathrm{T}}$ and $C = \alpha_{\mathrm{A}}$ which gives:

$$\eta = \frac{B - \sqrt{B^2 - 4AC}}{2A}. \tag{2.22}$$

Using the last equation one can calculate the equilibrium hydrogen neutral fraction for a given n_{H} and temperature.

---

[10]The use of Case B is more appropriate for $n_{\mathrm{H}} > n_{\mathrm{H,SSh}}$. However, we assume the photoionization due to RR is included in Γ_{TOT}, e.g., by using the best-fit function that is presented in equation 2.14. Therefore, Case A recombination should be adopted even at high densities.

Appendix B: The effects of box size, cosmological parameters and resolution on the Hi 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 Hi column densities, cannot be formed in a small cosmological box. Indeed, as shown in the top panels of Figure 2.12, one needs to use cosmological boxes larger than $\gtrsim 25$ comoving h^{-1}Mpc in order to achieve convergence in the Hi distribution (see also Altay et al., 2011). On the other hand, the bottom-right panel of Figure 2.12 shows that changing the resolution of the cosmological simulations also affects $f(N_{\mathrm{HI}}, z)$, although the effect is small.

The adopted cosmological parameters also affect the gas distribution and hence the Hi CDDF. For instance, one expects that the number of absorbers at a given density varies with the density parameter Ω_{b}, and the root mean square amplitude of density fluctuations σ_8. The bottom-left panel of Figure 2.12 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 Hi 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 $f(N_{\mathrm{HI}}, z)$).

Appendix C: RT convergence tests

C1: Angular resolution

The left-hand panel of Figure 2.13 shows the dependence of photoionization rates on the adopted angular resolution, i.e., the opening angle of the transmission cones $4\pi/N_{\mathrm{TC}}$. The photoionization rates are converged for $N_{\mathrm{TC}} = 64$ (our fiducial value) or higher.

C2: The number of ViP neighbors

The right panel of Figure 2.13 shows the dependence of the photoionization rates on the number of SPH neighbors of ViPs. As discussed in §2.2.2, ViPs distribute the ionizing photons they absorb among their $\mathrm{NGB}_{\mathrm{ViP}}$ nearest SPH neighbors. The larger the number of neighbors, 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 neighbors (our fiducial value is 5).
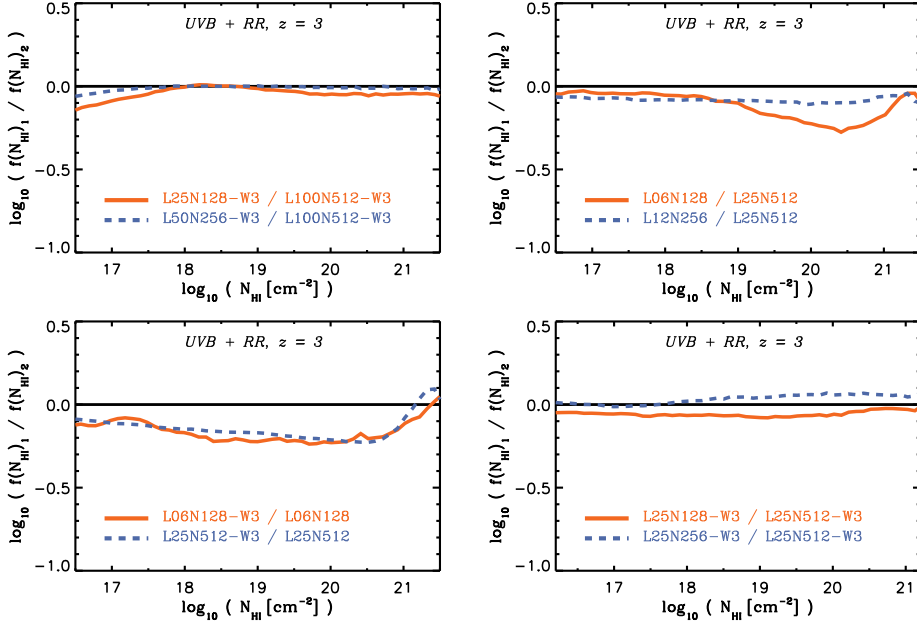
Figure 2.12: The relative changes in the Hɪ CDDF using different resolutions, box sizes and cosmologies in the presence of the UVB and diffuse recombination radiation. The *top-left* panel shows the effect of box size on $f(N_{HI}, z)$ for a fixed resolution at $z = 3$, where the orange solid (blue dashed) 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.
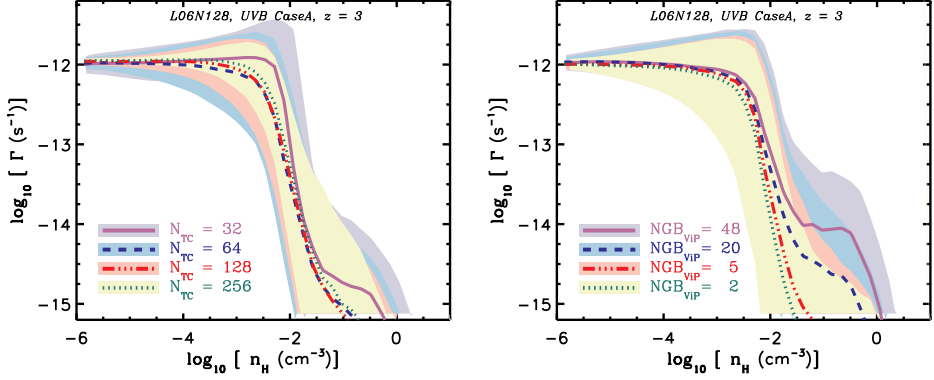
Figure 2.13: The UVB photoionization rate is converged for our adopted angular resolution, i.e., $N_{TC} = 64$, as shown in the *left* panel and our adopted number of ViP neighbors, i.e., $NGB_{ViP} = 5$, as shown in the *right* panel. Photoionization rate profiles are shown for the *L06N128* simulation 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.

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 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 recombination radiation. In Figure 2.14, we compare one of our UVB photoionization rate profiles[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

---

[11]Note that in our simulations the UVB photoionization rate is converged with the box size and the resolution as shown in §2.3.3.
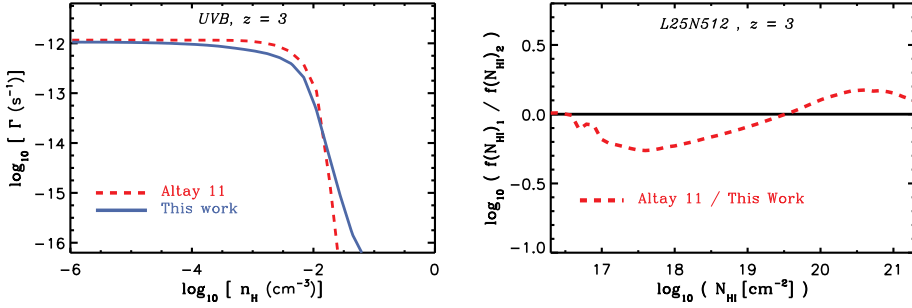
Figure 2.14: *Left:* Median UVB photoionization rate as a function of density at $z = 3$ using different RT methods. The red dashed curve shows the results based on the method that has been used in Altay et al. (2011) and the blue solid curve shows the result of this work. *Right:* The Hɪ CDDF of the *L25N512* simulation at $z = 3$ using different RT methods and without RR. The red dashed curve shows the ratio between the Hɪ CDDF given in Altay et al. (2011) and our results. This comparison shows that despite the overall agreement between our results and Altay et al. (2011), there are some important differences.

(see the left panel of Figure 2.13).

Appendix D: Approximated processes

D1: Multifrequency effects

As discussed in §2.2.3, in our RT simulations we have treated the multifrequency nature of the UVB radiation in the gray approximation (see equation 2.4). 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 Hɪ fractions by repeating the *L06N128* simulation at $z = 3$ with the UVB using 3 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 Figure 2.15, by plotting the ratio between the resulting hydrogen neutral fraction, η, and the same quantity in the original simulation that uses the gray approximation. This comparison shows that the simulation that uses multifrequency predicts hydrogen neutral fractions $< 10\%$ lower at low densities (i.e., $n_{\mathrm{H}} \lesssim 10^{-4} \, \mathrm{cm}^{-3}$). This does not change the resulting $f(N_{\mathrm{HI}}, z)$ noticeably at the column densities of interest here.

The spectral hardening captured in the simulation with 3 frequency bins is illustrated in the right panel of Figure 2.15. 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
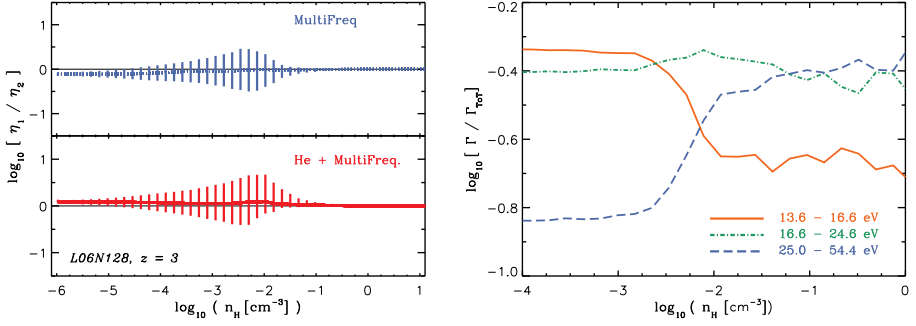
Figure 2.15: Spectral hardening and multifrequency treatment do not change the H I distribution significantly. *Left:* The ratio between hydrogen neutral fractions, η, obtained by using 3 frequency bins and by using the gray approximation is shown in the top-left apnel. The ratio between hydrogen neutral fractions resulting from a simulation with 4 frequency bins and explicit He treatment and the same quantity using the gray approximation and without explicit He treatment is shown in the bottom-left panel. The vertical lines with different lengths indicate the $15\% - 85\%$ percentiles. *Right:* The fractional contribution of different frequency bins to the total UVB photoionization rate for the simulations with 3 frequency bins. All the RT calculations are performed using the *L06N128* simulation at $z = 3$ in the presence of the UVB and assuming Case A recombination.

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 that 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 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 4 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 HeII ionization.

We adopted a helium mass fraction of 25% 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 Figure 2.15. The hydrogen neutral fractions are very close in the two simulations. However, the simulation with multifrequency and explicit He treatment results in hydrogen neutral fractions that are $< 10\%$ higher at low densities (i.e., $n_{\mathrm{H}} \lesssim 10^{-4}\,\mathrm{cm}^{-3}$). This difference is barely noticeable in the comparison between the two Hɪ CDDFs (not shown).