

# Simulated metal and H I absorption lines at the conclusion of reionization

L. A. García,<sup>1,2★</sup> E. Tescari,<sup>2,3</sup> E. V. Ryan-Weber<sup>1,2</sup> and J. S. B. Wyithe<sup>2,3</sup>

<sup>1</sup>Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia

<sup>2</sup>ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO)

<sup>3</sup>School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia

Accepted 2017 June 1. Received 2017 May 2; in original form 2016 September 30

## ABSTRACT

We present a theoretical study of intergalactic metal absorption lines imprinted in the spectra of distant quasars during and after the Epoch of Reionization (EoR). We use high-resolution hydrodynamical simulations at high redshift ( $4 < z < 8$ ), assuming a uniform UV background Haardt–Madau 12, post-processing with CLOUDY photoionization models and Voigt profile fitting to accurately calculate column densities of the ions C II, C IV, Si II, Si IV and O I in the intergalactic medium (IGM). In addition, we generate mock observations of neutral hydrogen (H I) at  $z < 6$ . Our simulations successfully reproduce the evolution of the cosmological mass density ( $\Omega$ ) of C II and C IV, with  $\Omega_{\text{C II}}$  exceeding  $\Omega_{\text{C IV}}$  at  $z > 6$ , consistent with the current picture of the tail of the EoR. The simulated C II exhibits a bimodal distribution with large absorptions in and around galaxies, and some traces in the lower density IGM. We find some discrepancies between the observed and simulated column density relationships among different ionic species at  $z = 6$ , probably due to uncertainties in the assumed UV background. Finally, our simulations are in good agreement with observations of the H I column density distribution function at  $z = 4$  and the H I cosmological mass density  $\Omega_{\text{H I}}$  at  $4 < z < 6$ .

**Key words:** methods: numerical – intergalactic medium – cosmology: theory.

## 1 INTRODUCTION

Understanding the Epoch of Reionization (EoR) is one of the current challenges of extragalactic astronomy. It will complete the big picture of the thermal history of the Universe (e.g. McQuinn 2015). The EoR commenced when the first stars switched on, producing UV radiation that transformed the neutral hydrogen (H I) in the surrounding circumgalactic and intergalactic media (CGM and IGM) into ionized hydrogen (H II). Observations of Ly  $\alpha$  photons (Ly  $\alpha$  1216 Å) in absorption towards  $z_{\text{em}} \gtrsim 6$  quasars (QSOs) show fluctuations in flux consistent with reionization concluding at  $z \lesssim 6$  (e.g. Fan 2006; Becker et al. 2015b). Measuring Lyman series absorption or UV emissivity at wavelengths bluewards of Ly  $\alpha$  becomes almost impossible at redshifts greater than 5.5 due to the increasing density of matter and neutral hydrogen fraction (Becker, Bolton & Lidz 2015a).

Metal absorption lines are an alternative proxy to the Lyman series for probing the evolution of the IGM during the EoR and offer many advantages at high redshift. For example, they can be detected even when hydrogen is completely saturated in the spectra of the background quasars, since some of these transitions occur redwards of Ly  $\alpha$  and are thus unaffected by Lyman series absorption

and Gunn–Peterson troughs (Gunn & Peterson 1965). Also, ionic transitions in the QSO spectra at high redshift provide important constraints on the model of the ionizing background that drove reionization (e.g. Furlanetto & Mesinger 2009; Becker et al. 2011, 2015b; Finlator et al. 2016). Low-ionization transitions (O I, C II, Si II, Mg II, Fe II) trace the location of neutral hydrogen at high redshift. On the other hand, high-ionization states (C IV, Si IV, O VI) have a comparably larger ionization potential than H, therefore the energy required to produce these transitions is reached in regions where hydrogen is highly ionized at early times.

The direct measurement of absorption features imprinted on the spectra of quasars at high redshift is the best method to infer the ionic ratios, but the observations depend on the wavelength bands where each transition can be found and the observational sensitivity is limited by the decreasing likelihood of an absorption line to be detected. In order to complement the observational techniques, hydrodynamical simulations are used to simulate the physical environment at high redshift where the absorption occurs. The numerical approach can produce a large number of sightlines, improving the statistical estimation of the column densities at redshifts that are not accessible due to current observational limits. Although the field is rapidly expanding, this approach is computationally expensive and the large dynamical range of the underlying physical phenomena makes a true self-consistent simulation impossible. Current simulations are still not able to simultaneously resolve the small scales

\* E-mail: luzangelagarcia22@gmail.com

(turbulence, shocks, fluctuations in the background density, etc.) and large scales (cosmological structures such as clusters and filaments) involved in the progression of the EoR.

An extensive observational effort has been pursued with C IV, the triply ionized state of carbon. C IV offers many advantages that make its detection easier than other ionic species: large oscillator strength, wavelength redwards of Ly  $\alpha$  emission and a doublet transition. Observations of C IV in the foreground of quasar spectra at high redshift by Songaila (2001, 2005), Pettini et al. (2003), Ryan-Weber, Pettini & Madau (2006); Ryan-Weber et al. (2009), Simcoe (2006), Simcoe et al. (2011), D’Odorico et al. (2010, 2013, 2016) and Boksenberg & Sargent (2015) have built the largest sample of metal absorbers as a function of redshift to date. The cosmological mass density  $\Omega_{\text{CIV}} = \frac{\rho_{\text{CIV}}}{\rho_{\text{crit}}}$  (the density of C IV with respect to the critical density today) drops at high redshifts during the progression of the EoR, due to the changing ionization state of the IGM and a decrease of its metallicity.

Numerical simulations by Oppenheimer & Davé (2006), Oppenheimer, Davé & Finlator (2009), Tescari et al. (2011), Cen & Chisari (2011), Pallottini et al. (2014), Finlator et al. (2015) and Rahmati et al. (2013) and a comparison by Keating et al. (2016) have tried to reproduce the evolution of  $\Omega_{\text{CIV}}$ , taking into account different feedback prescriptions, photoionization modelling and variations in the UV ionizing background at high redshift. These theoretical efforts provide insight into the physical environments where C IV absorptions occur and show that C IV traces the distribution of the IGM at high temperature up to a few hundred kpc away from galaxies (e.g. Oppenheimer et al. 2009).

On the other hand, low-ionization state ions, such as neutral oxygen O I, should trace the distribution of H I at high redshift. In fact, since the ionization potential of O I differs from that of the neutral hydrogen by 0.02 eV, the two ions sit in tight charge exchange. The detections of O I towards high-redshift QSOs in Becker et al. (2006, 2011) are consistent with large variations from one line of view to another, i.e. with an inhomogeneous distribution of the ionizing sources as well as the absorbers. Numerical results from Finlator et al. (2013) showed a tight correlation between H I and O I column densities and studies by Keating et al. (2014) at  $N_{\text{HI}} > 10^{17} \text{ cm}^{-2}$  (where the gas is self-shielded) revealed an excellent agreement between O I and H I fractions, regardless of the photoionization model assumed, as well as an increasing incidence rate of O I at higher redshift, consistent with the IGM being more neutral when approaching the EoR (Becker et al. 2011).

Likewise, C II has been observed by Becker et al. (2006) and modelled by D’Odorico et al. (2013) at  $z = 5.7$  to be best fitted by low-density gas,  $\delta = \frac{\rho_{\text{gas}}}{\langle \rho \rangle} - 1 = 10$ , where  $\langle \rho \rangle$  is the mean density at the considered redshift. Low-ionization metals as C II and Mg II are routinely detected at low redshifts in regions close to the centre of galaxies or in damped Ly  $\alpha$  systems (DLAs, i.e. absorption systems with  $N_{\text{HI}} > 10^{20.3} \text{ cm}^{-2}$ ; Wolfe et al. 1986; Wolfe, Gawiser & Prochaska 2005). The latter are optically thick structures that constitute the main reservoir of neutral hydrogen after the EoR. DLAs have been studied both observationally (e.g. Péroux et al. 2003; Prochaska, Herbert-Fort & Wolfe 2005; O’Meara et al. 2007; Prochaska & Wolfe 2009; Crighton et al. 2015) and theoretically/numerically (e.g. Nagamine, Springel & Hernquist 2004; Pontzen et al. 2008; Barnes & Haehnelt 2009; Tescari et al. 2009; Bird et al. 2014; Maio & Tescari 2015; Rahmati et al. 2015) in order to understand their physical properties and statistical distribution. However, many open questions still remain on the nature of these systems and their connection with the chemical enrichment of the Universe.

The aim of this work is to (i) compare high-redshift observations of metal and H I absorbers with synthetic spectra and (ii) use simulations to gain insight into the underlying physical explanation for the observations. This paper is divided as follows: in Section 2, we introduce our numerical simulations. In Section 3, we explain the method used to include the UV background Haardt & Madau 2012 (hereafter HM12 UVB), the photoionization modelling and the column density calculation procedure. In Section 4, we show the evolution with redshift of the physical properties that describe the IGM. Sections 5 and 6 analyse the ionic species in the simulations focusing on the cosmological mass density and the relationships among different ions. In the subsequent Section 7, we discuss the evolution of the cosmological mass density of neutral hydrogen at  $z < 6$ . Finally, in Sections 8 and 9 we summarize our results and the implications of the models. Throughout the paper, we use the prefix c for comoving and p for physical distances.

## 2 COSMOLOGICAL SIMULATIONS

The numerical simulations used in this work reproduce representative volumes of the Universe at redshift  $4 < z < 8$  and are run with the smoothed particle hydrodynamics (SPH) code P-GADGET3(XL) – a customized version of GADGET-3 (Springel 2005). The model is an extension at high redshift of the AustraliaN GADGET-3 early Universe Simulations (ANGUS) project (Tescari et al. 2014; Katsianis, Tescari & Wyithe 2015, 2016; Maio & Tescari 2015; Katsianis et al. 2017). Among other technical improvements, the suite of hydrodynamical simulations takes into account: a multiphase star formation criterion from Springel & Hernquist (2003), self-consistent stellar evolution and chemical enrichment modelling (Tornatore et al. 2007), supernova (SN) momentum- and energy-driven galactic winds<sup>1</sup> (MDW and EDW; Springel & Hernquist 2003; Puchwein & Springel 2013), metal-line cooling (Wiersma et al. 2009) and low-temperature cooling by molecules/metals (Maio et al. 2007). In addition, the code is supported by a parallel Friends-of-Friends (FoF) algorithm to identify collapsed structures and SUBFIND to classify substructures within FoF haloes.

Each simulation generates a cosmological box (including periodic boundary conditions) with the same initial number of gas and dark matter (DM) particles for a total of  $2 \times 512^3$ . Initial masses of the gas and DM particles are given in Table 1. Whenever the gas density is above the threshold  $\rho_{\text{th}}$ , there is a probability that a gas particle will turn into a star.<sup>2</sup> Stochastically, a new star-type particle is introduced in the simulation. Each star particle represents a simple stellar population with mass  $0.1 M_{\odot} \leq m \leq 100 M_{\odot}$ . The stars with mass  $m \leq 40 M_{\odot}$  explode as SNe before turning into a black hole, while stars that have masses above this threshold collapse into a black hole without passing through the SN stage.

A flat  $\Lambda$  cold dark matter model with cosmological parameters from the latest release of the Planck Collaboration XIII (2016) is assumed:  $\Omega_{\text{om}} = 0.307$ ,  $\Omega_{\text{ob}} = 0.049$ ,  $\Omega_{\Lambda} = 0.693$ ,  $n_{\text{s}} = 0.967$ ,  $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (or  $h = 0.6774$ ) and  $\sigma_8 = 0.816$ . The simulations were calibrated according to the parameters used in

<sup>1</sup> Although active galactic nucleus (AGN) feedback is implemented in the code (Springel, Di Matteo & Hernquist 2005; Fabjan et al. 2010; Planelles et al. 2013), in this work we do not consider it. We stress that for the analysis presented in the paper, the role of AGN feedback is expected to be negligible (see e.g. Keating et al. 2016 and references therein).

<sup>2</sup> This probability is calibrated to reproduce the Kennicutt–Schmidt law (Schmidt 1959).

**Table 1.** Summary of the simulations used in this work. Column 1: run name. Column 2: box size. Column 3: Plummer-equivalent comoving gravitational softening length. Columns 4 and 5: mass of gas and DM particles. All the simulations have the same initial number of gas and DM particles ( $2 \times 512^3$ ). Column 6: feedback model. Column 7: inclusion of low-temperature metal and molecular cooling (Maio et al. 2007; Maio & Tesfari 2015). The first run, Ch 18 512 MDW, is the fiducial model. The second one in the list, Ch 18 512 MDW mol, has exactly the same configuration as the reference run, but includes low-T metal and molecular cooling.

Simulation	Box size (cMpc $h^{-1}$ )	Comoving softening (ckpc $h^{-1}$ )	$M_{\text{gas}}$ ( $\times 10^5 M_{\odot} h^{-1}$ )	$M_{\text{DM}}$ ( $\times 10^6 M_{\odot} h^{-1}$ )	Model for SN-driven winds	Low-T metal & molecular cooling
<b>Ch 18 512 MDW</b>	<b>18</b>	<b>1.5</b>	<b>5.86</b>	<b>3.12</b>	<b>Momentum-driven</b>	
Ch 18 512 MDW mol	18	1.5	5.86	3.12	Momentum-driven	✓
Ch 18 512 EDW	18	1.5	5.86	3.12	Energy-driven	
Ch 18 512 EDW mol	18	1.5	5.86	3.12	Energy-driven	✓
Ch 12 512 MDW mol	12	1.0	1.74	0.925	Momentum-driven	✓
Ch 25 512 MDW mol	25	2.0	15.73	8.48	Momentum-driven	✓

Tescari et al. (2014) and Katsianis et al. (2015), and are compatible with observations of the cosmic star formation rate (SFR) density history and the galaxy stellar mass function at  $z = 6-8$ .

The subgrid scheme takes into account the lifetimes of stars of different mass and follows the evolution of hydrogen, helium, molecules and metals (C, Ca, O, N, Ne, Mg, S, Si and Fe) released from SNIa, SNII and low- and intermediate-mass stars. It is possible to vary the initial mass function (IMF), the lifetime function and stellar yields. The algorithm is ideally suited to model IGM enrichment produced by galactic winds blown by ‘starburst’ galaxies at high redshift. Radiative cooling and heating processes are included according to Wiersma et al. (2009) and Maio et al. (2007). The IMF considered is the Chabrier multisloped (Chabrier 2003), as described in equation (1), that produces a large number of intermediate- and high-mass stars, expected to play an important role during reionization,

$$\zeta(m) = \begin{cases} 0.497 \times m^{-0.2} & 0.1 M_{\odot} \leq m < 0.3 M_{\odot}, \\ 0.241 \times m^{-0.8} & 0.3 M_{\odot} \leq m < 1 M_{\odot}, \\ 0.241 \times m^{-1.3} & m \geq 1 M_{\odot}. \end{cases} \quad (1)$$

We adopted the following stellar yields.

(i) SNIa: Thielemann et al. (2003). The mass range for the SNIa originating from binary systems is  $0.8 M_{\odot} < m \leq 8 M_{\odot}$ , with a binary fraction of 7 per cent.

(ii) SNII (massive stars): Woosley & Weaver (1995). The mass range for SNII is  $m > 8 M_{\odot}$ .

(iii) Asymptotic giant branch (low- and intermediate-mass) stars: van den Hoek & Groenewegen (1997).

## 2.1 Feedback mechanisms

In order to regulate the star formation and to chemically enrich the IGM, the simulations have been set up to account for kinetic SN-driven winds. According to the nature of the feedback and its effectiveness, galactic winds are classified as momentum- or energy-driven winds.

### 2.1.1 Momentum-driven galactic winds

MDW (Puchwein & Springel 2013) are the reference feedback model used in this work. We assume that the mass-loss rate associated with the winds  $\dot{M}_w$  is proportional to the SFR  $\dot{M}_*$  through  $\eta$ , the wind mass-loading factor that accounts for the efficiency of the wind

$$\dot{M}_w = \eta \dot{M}_*. \quad (2)$$

In this model, the velocity of the winds  $v_w$  is regulated by the mass of the host halo and  $R_{200}$ , the radius where the density is 200 times larger than the critical density at redshift  $z$ :

$$v_w = 2 \sqrt{\frac{GM_h}{R_{200}}} = 2 \times v_{\text{circ}}, \quad (3)$$

where  $v_{\text{circ}}$  is the circular velocity and  $R_{200}$  is defined as follows:

$$R_{200} = \sqrt[3]{\frac{3}{4\pi} \frac{M_h}{200\rho_c\Omega_{0m}} (1+z)^{-1}}, \quad (4)$$

with  $\Omega_{0m}$  and  $\rho_c$  the matter density and the critical density today. The conservation of momentum of the winds imposes that  $\eta$  is proportional to the inverse of the wind velocity  $v_w$ :

$$\eta = 2 \times \frac{600 \text{ km s}^{-1}}{v_w}. \quad (5)$$

The normalization factor was chosen to reproduce the observed cosmic SFR density and galaxy stellar mass function up to  $z = 8$ . According to the model, weak winds ( $v_w < 600 \text{ km s}^{-1}$ ) lead to a large efficiency and more material is expelled from SN with respect to strong winds ( $v_w > 600 \text{ km s}^{-1}$ ) that reach further but load less material.<sup>3</sup> Stochastically, some gas particles are selected to be part of the wind, and subsequently, decouple from the hydrodynamical scheme for a given amount of time  $t_{\text{dec}}$  (Tescari et al. 2014).

### 2.1.2 Energy-driven galactic winds

To test the impact of the adopted feedback model on our results, we also use energy-driven winds (Springel & Hernquist 2003). The only remarkable difference with momentum-driven winds is the scaling form of  $\eta$ . As for MDW, the speed of the winds is regulated by  $v_w = 2 \times v_{\text{circ}}$ , but in this case due to conservation of the energy of the winds, the wind mass-loading factor  $\eta$  scales as the square of the inverse of  $v_w$ ,

$$\eta = 2 \times \left( \frac{600 \text{ km s}^{-1}}{v_w} \right)^2, \quad (6)$$

<sup>3</sup> Please note that in these simulations the normalization of the wind mass-loading factor,  $\eta_0 = 600 \text{ km s}^{-1}$ , is slightly larger than in previous ANGUS runs, such as those presented in Tescari et al. (2014) and Katsianis et al. (2015), where  $\eta_0 = 450 \text{ km s}^{-1}$ . This is due to the fact that these simulations do not include AGN feedback (and in particular the early AGN feedback in low-mass galaxies model), and therefore a recalibration of the wind model was necessary for properly matching the cosmic SFR density history and galaxy stellar mass function up to  $z = 8$ .

making EDW more aggressive than MDW, especially in low-mass galaxies. Thermal feedback produced by SNIa and SNII is also considered, in addition to the kinetic feedback just described.

### 3 METHODOLOGY

The set of simulations was run to  $z = 4$ , with initial conditions at  $z = 125$ , using the Raijin supercluster from the National Computational Infrastructure (NCI) facility.<sup>4</sup> These simulations describe the physical conditions of the gas and the evolution of H, He, C, Ca, O, N, Ne, Mg, S, Si and Fe released from SNIa and SNII. We post-process the simulations introducing a uniform UV ionizing background: a field radiation due to the cosmic microwave background and the Haardt & Madau (2012) ultraviolet/X-ray background from quasars and galaxies with saw-tooth attenuation that evolves with redshift (hereafter HM12).

Assuming this UVB, we compute the ionization states of each element with CLOUDY photoionization code v8.1 (Ferland et al. 2013) for optically thin gas in ionization equilibrium, focusing in particular on the following ions: H I, C II, C IV, Si II, Si IV and O I. We consider only the transition with the highest oscillator strength and rest-frame wavelength  $\lambda_{\text{rest}} \geq 1216 \text{ \AA}$ , which can be observed at high redshift. The assumed metallicity in the CLOUDY tables is solar.

In addition, an effective prescription for H I self-shielding was introduced to accurately describe the regions at the centre of the galaxies, where  $N_{\text{HI}}$  is significantly higher than that in the IGM due to the shielded bubbles that contain pristine hydrogen. We adopt the parametric function of Rahmati et al. (2013):

$$\frac{\Gamma_{\text{UVB}}^{\text{phot}}}{\Gamma_{\text{UVB}}} = (1 - f) \left[ 1 + \left( \frac{n_{\text{H}}}{n_0} \right)^\beta \right]^{\alpha_1} + f \left[ 1 + \frac{n_{\text{H}}}{n_0} \right]^{\alpha_2}, \quad (7)$$

whose best-fitting parameter values reproducing the radiative transfer results at  $1 < z < 5$  are:  $\alpha_1 = -2.28 \pm 0.31$ ,  $\alpha_2 = -0.84 \pm 0.11$ ,  $\beta = 1.64 \pm 0.19$ ,  $n_0 = (1.003 \pm 0.005) n_{\text{H,SSH}}$  (where  $n_{\text{H,SSH}}$  is the self-shielding density threshold) and  $f = 0.02 \pm 0.0089$ . The equilibrium neutral hydrogen density is obtained using a recombination rate given by Hui & Gnedin (1997) and the photoionization rate as a function of the temperature as given by Theuns, Leonard & Efstathiou (1998).

In order to mimic real observations and avoid the introduction of bias in the data, we generate random lines of sight inside the simulated box along the three perpendicular directions and extract the relevant physical information taking into account positions, velocities, densities and temperatures of the SPH particles inside each line of sight. Subsequently, we compute a synthetic spectra in density and optical depth/flux as a function of the velocity width (1024 pixels) for each ion in the simulation, according to the procedure introduced by Theuns et al. (1998). For a bin  $j$  at a position  $x(j)$ , the density and density weighted temperature and velocity are calculated from

$$\rho_X(j) = a^3 \sum_i X(i) W_{ij}, \quad (8)$$

$$(\rho T)_X(j) = a^3 \sum_i X(i) T(i) W_{ij}, \quad (9)$$

$$(\rho v)_X(j) = a^3 \sum_i X(i) (a \dot{x}(i) + \dot{a} [x(i) - x(j)]) W_{ij}, \quad (10)$$

where  $a$  is the scale factor,  $X(i)$  the abundance of the species  $X$  of the SPH particle  $i$  and  $W_{ij} = m W(q_{ij})/h_{ij}^3$  the normalized SPH kernel.  $W$  is the SPH kernel,  $m$  the particle mass and:

$$q_{ij} = \frac{a|x(i) - x(j)|}{h_{ij}}, \quad (11)$$

$$h_{ij} = \frac{1}{2} [h(i) + h(j)], \quad (12)$$

with  $h$  the physical softening scale.

With this information, it is possible to calculate the number density of the ion transition considered  $n_{\text{ion}}$ . The synthetic flux for any transition at the redshift–space coordinate  $u$  is given by  $F(u) = \exp[-\tau(u)]$ , with  $\tau(u)$ :

$$\tau(u) = \frac{\sigma_{0,1} c}{H(z)} \int_{-\infty}^{\infty} n_1(x) V \left[ u - x - v_{\text{pec,||}}^{\text{IGM}}(x), b(x) \right] dx, \quad (13)$$

where  $\sigma_{0,1}$  is the cross-section of the ion transition,  $H(z)$  the Hubble parameter at  $z$ ,  $x$  the space coordinate in  $\text{km s}^{-1}$ ,  $b$  the velocity dispersion (in units of  $c$ ) and  $V$  the Voigt profile. The spectra can be converted from the velocity space  $v$  to the observed wavelength using  $\lambda = \lambda_0(1+z)(1+v/c)$ .

In numerical works, it is a common practice to normalize H I fluxes averaged over a large number of random lines of sight to the observed mean normalized flux of the Ly  $\alpha$  forest at a given redshift,  $\langle F(z) \rangle = \exp(-\tau_{\text{eff}})$ , through a constant rescaling factor  $A_{\text{HI}}$ . For consistency, the fluxes of metal ions are also rescaled by the same constant factor. However, at  $z > 5$ , reliable measurements of  $\tau_{\text{eff}}$ , and therefore  $A_{\text{HI}}$ , are not possible due to the thickening of the Ly  $\alpha$  forest. To overcome this problem, we adopted a reverse approach. First, we calculated the C IV column density distribution function (CDDF) at  $z = 4.8$  and  $5.6$  (see Section 5.1) using three different simulations (Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW) and a range of scaling factors for the C IV optical depths,  $\tau_{\text{CIV}}$ . Then, we selected the scaling factor,  $A_{\text{CIV}}$ , which provided the best chi-by-eye agreement between the simulated and observed distribution functions. Finally, we rescaled all the ionic fluxes (including H I) in all the simulations and at all the redshifts considered (i.e. also at  $z \leq 5$ , for consistency) by  $A_{\text{CIV}}$ . The best-fitting value is  $A_{\text{CIV}} = 0.85$ , very close to unity.

The individual spectra are convolved with Gaussian noise profiles with full width at half-maximum, FWHM =  $7 \text{ km s}^{-1}$  to produce final synthetic spectra with a signal-to-noise ratio,  $S/N = 50$ , comparable with observations obtained using the UVES spectrograph mounted at the Unit 2 of the Very Large Telescope. Finally, to obtain a fair comparison with real data, we fit all the individual absorption features in the spectra through Voigt profile components in a range of  $100 \text{ km s}^{-1}$  among contiguous systems with the code VPFIT v.10.2 (Carswell & Webb 2014).

The column density  $N$ , equivalent width EW and Doppler parameter  $b$  from each spectrum is used to produce a sample for each ion. The estimated errors of the column density  $\Delta N$ , Doppler parameter  $\Delta b$  and redshift  $\Delta z$  of the absorption features are also calculated with VPFIT. The individual fits are selected if the conditions  $N > \Delta N$  and  $b > \Delta b$  are fulfilled simultaneously. Otherwise, the component is rejected.

As a final remark, we note that radiative transfer effects are not included in our simulations. The models use the evolving HM12 uniform UVB to quantify the ionization state of the CGM/IGM at a given redshift. The aim of this work is not to follow the progression of reionization, nor the evolution of the H II bubbles or their topology. The implicit assumption is that our boxes (that are small compared to the size of the H II bubbles at the redshifts of

<sup>4</sup> <http://nci.org.au>

interest) represent a region of the Universe already reionized at a level given by the HM12 UVB. At  $6 < z < 8$ , chemical enrichment occurs mostly inside and in close proximity of galaxies (interstellar medium, CGM and high-density IGM) where, assuming an inside-out progression of reionization, the gas in which metals lie should be ionized. Although proper Radiative Transfer (RT) calculations would be more accurate, they are extremely expensive from the computational point of view. Using a uniform (evolving) UVB is a common approach in the literature (see e.g. Oppenheimer et al. 2009; Keating et al. 2016). Moreover, Finlator et al. (2015) studied the evolution of carbon absorption from  $z = 10$  to 5 with cosmological hydrodynamic simulations that include a self-consistent multifrequency, inhomogeneous UVB. They found that the difference between their more realistic UVB and the uniform HM12 is within  $\sim 2$ – $4$  times, which according to them is fair agreement given the uncertainties.

We stress that in this work we do not present results on the neutral hydrogen fraction at  $z > 6$ . In this case, a uniform background slowly ionizing the IGM is a completely wrong picture. For this reason, we only study the H I fraction at  $z < 6$ . As already mentioned before, to accurately describe the regions at the centre of galaxies we also include an effective prescription for H I self-shielding (Rahmati et al. 2013).

#### 4 EVOLUTION OF THE IGM DURING REIONIZATION

Fig. 1 shows the distribution of the gas particles in the fiducial simulation Ch 18 512 MDW in diagrams of overdensity–temperature at redshifts  $z = 8, 6$  and 4, comparing total metallicity, number of particles and H I fraction (left-hand, middle and right-hand panels, respectively). These redshifts are chosen to study the IGM at (i)  $z = 8$  when reionization is progressing; (ii)  $z = 6$ , when the overlap of the H II bubbles is nearly complete at the end of the EoR and (iii)  $z = 4$  when the reionization of hydrogen has fully concluded and the only remaining reservoirs of neutral hydrogen are self-shielded regions of H I and DLA systems.

The diagrams are split in well-defined regions whose sizes evolve with time in terms of the overdensity,  $\Delta = \delta + 1 = \frac{\rho_{\text{gas}}}{\langle \rho \rangle}$ , and the temperature  $T$ :

- (i) voids: gas particles at very low densities ( $\Delta < 1$ );
- (ii) IGM: the gas in this regime follows an adiabatic relation  $T = T_0 \Delta^\gamma - 1$  and is in ionization equilibrium: cooling is counterbalanced by photoheating. The conditions of the particles in the IGM are  $T \leq 10^4$  K and  $1 \leq \Delta < 10$ ;
- (iii) circumgalactic/interstellar media (CGM/ISM) represent the transition between the IGM and the inside of galaxies, with typical densities of  $\Delta \sim 10^{1-3}$ . This environment is heated by photoionization fronts coming from galaxies. The sharp feature at  $\Delta \sim 10^3$  represents the density threshold where gas particles turn stochastically into star particles in the simulation;
- (iv) star-forming region: particles in this regime are above  $\Delta \sim 10^3$ , have temperatures higher than  $10^4$  K and follow an effective equation of state imposed by the subgrid star formation model (Springel & Hernquist 2003);
- (v) shock-heated gas: a growing region of the IGM that is heated by feedback processes at late times ( $T > 10^5$  K and low densities).

At  $z = 6$ – $8$ , most of the gas in the diagrams is located in the so-called diffuse phase – low  $T$  ( $\leq 10^4$  K) and low  $\Delta$  ( $\leq 10^2$ ) – the first stars are being formed and the main processes that increase the temperature of the IGM (and voids) are taking place.

The leftmost panels show the metallicity with respect to solar. The enriched sections of the phase diagram are mainly star-forming regions and some gas at very high temperature, expelled from galaxies through SN-driven winds. On the other hand, voids exhibit zero metallicity, because the galactic outflows do not reach regions that far from the centre of the galaxies. The CGM and ISM are chemically enriched at high redshift. Instead, the IGM is very metal-poor at  $z = 8$ . At late stages (bottom panels), feedback prescriptions spread out material from the high-density gas to regions originally empty at higher redshifts. Chemical enrichment contributes progressively more to the cooling processes.

The right-hand panels show the distribution of hydrogen in the reference simulation. Yellow indicates completely neutral hydrogen ( $X_{\text{HI}} \geq 10^{-1}$ ), dark green particles contain H II ( $X_{\text{HI}} < 10^{-7}$ ) and the intermediate colours are used to illustrate the transition between the two regimes. It is clear that most of the reservoirs of H I are located at low temperature and intermediate- to high-density regions, which are presumably where DLAs or isolated self-shielded regions are located. On the other hand, the shock-heated gas with temperatures  $T > 10^5$  K contains a large amount of ionized gas. The fraction of this gas that has been chemically enriched is expected to be found in high-ionization states, such as C IV or Si IV, consistent with Cen & Chisari (2011).

Please note that, for the reasons discussed at the end of Section 3, the top-right panel should not be regarded as an accurate representation of the neutral hydrogen content in the Universe at  $z = 8$ , since a uniform UVB does not properly describe the progression of reionization. We included this panel just for illustrative purposes.

#### 5 EVOLUTION OF CARBON

In order to study the evolution of the IGM at high redshift, we use synthetic spectra and analyse the evolution of metal absorption lines.

##### 5.1 C IV CDDF

C IV, the triply ionized state of carbon, is the most detected transition and studied in the foreground of quasar spectra at high redshift. We devote this section to its cosmological evolution and compare our theoretical predictions with the available observational archive.

The CDDF is an observable that takes into account the statistical distribution of absorption systems with respect to their column densities. Once the absorptions are detected and the redshift path is confirmed to be complete, the CDDF can be used to build the cosmological mass density of a particular ion. By construction, the CDDF, or  $f(N, X)$ , quantifies the number of absorption systems  $n_{\text{sys}}$  in the column density interval  $(N, N + \Delta N)$  in an absorption path  $\Delta X$ :

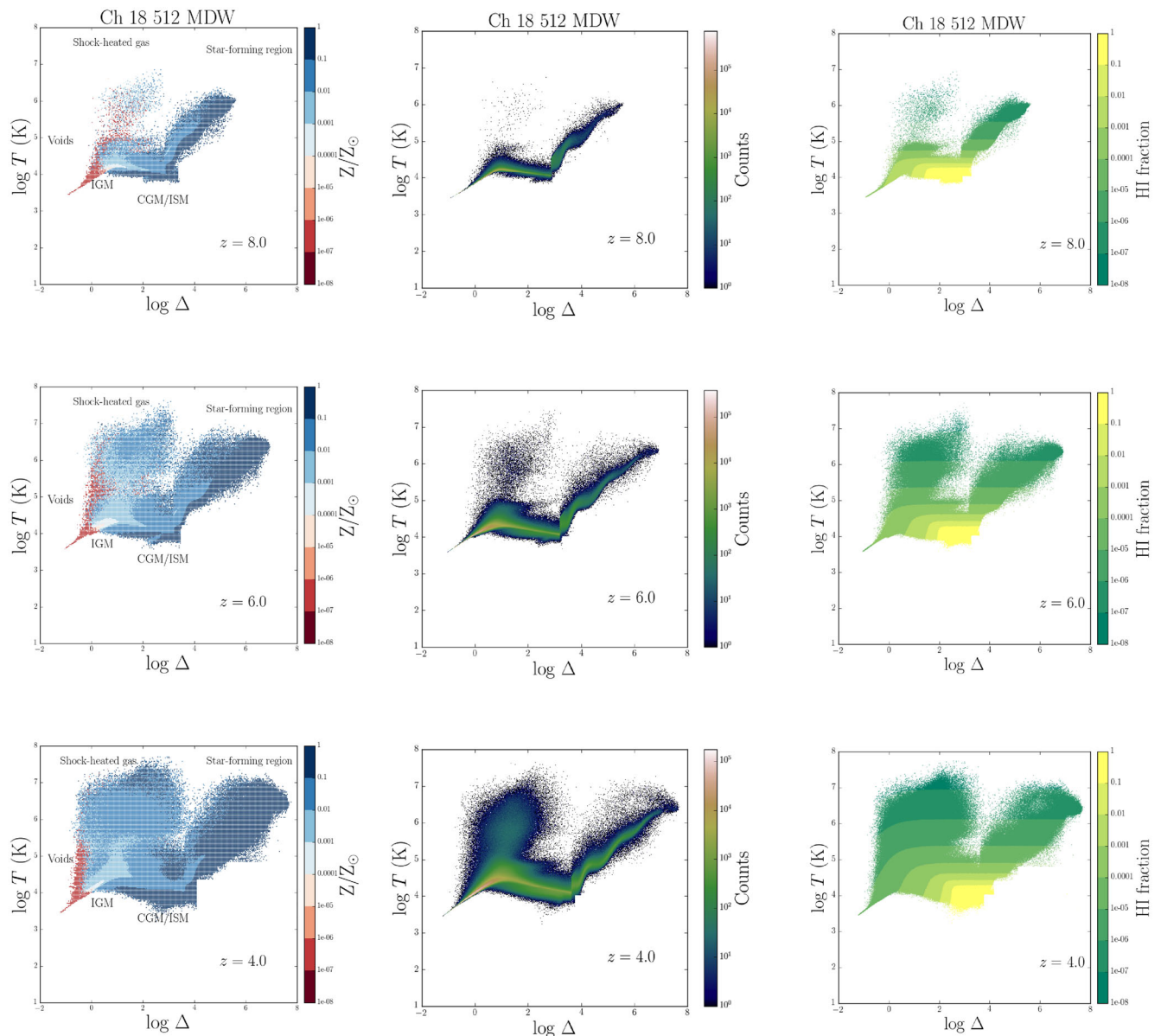
$$f(N, X) = \frac{n_{\text{sys}}(N, N + \Delta N)}{n_{\text{lov}} \Delta X}, \quad (14)$$

where  $n_{\text{lov}}$  is the number of lines of view (lov) considered. The absorption path relates the Hubble parameter at a given redshift  $z$  with the correspondent redshift path  $\Delta z$  as follows:

$$\Delta X = \frac{H_0}{H(z)} (1+z)^2 \Delta z. \quad (15)$$

The equivalent redshift path of our cosmological boxes is defined as

$$\Delta z = (1+z) \frac{\Delta v}{c}, \quad (16)$$



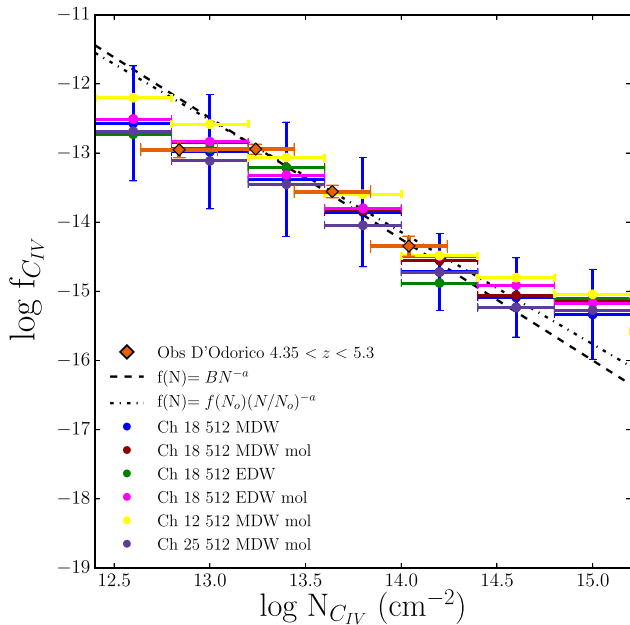
**Figure 1.** Density–temperature diagram for gas particles at redshifts  $z = 8, 6$  and  $4$  for the fiducial model Ch 18 512 MDW. The left-hand panels corresponds to the distribution of the total metallicity (in solar units). The dark red region represents zero metallicity. As a reference, the different regions are tagged (voids, IGM, CGM/ISM, shock-heated gas and star-forming region). In the middle panels, the colour maps represent the number of particles with overdensity and temperature in a given  $(\log \Delta_i, \log T_i)$  bin. On the right-hand side, we show the distribution of H I fraction. The top-right panel was included for illustrative purposes and does not represent a realistic description of the neutral hydrogen content of the Universe at  $z = 8$  (see the discussion at the end of Section 3).

where  $\Delta v$  is the box size in  $\text{km s}^{-1}$  at a given redshift  $z$ . We impose a cut-off in column density,  $\log N_{\text{th}} (\text{cm}^{-2}) = 12.5$ , to the synthetic data to mimic the sensitivity of available observations and to avoid including poorly sampled features. Although a large number of systems below this column density threshold is present ( $\sim 80$  per cent), these have not been taken into account in the statistics. Moreover, we limit the calculation of the C IV–CDDF to  $N_{\text{C IV}} (\text{cm}^{-2}) < 10^{15.2}$ , consider bins of 0.4 dex and introduce Poissonian errors for the theoretical sample as  $\sqrt{n_{\text{sys}}}$  to fairly compare with the observations of D’Odorico et al. (2013).

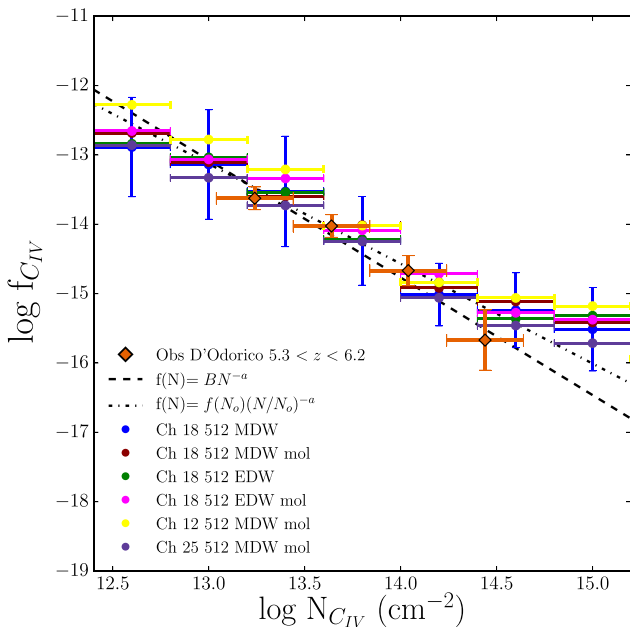
Fig. 2 shows the predicted CDDF of C IV at  $z = 4.8$  and compares with observations by D’Odorico et al. (2013) and two fitting functions proposed by the authors:  $f(N) = BN^{-\alpha}$  with  $B = 10.29 \pm 1.72$  and  $\alpha = 1.75 \pm 0.13$  and  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 13.56$

and  $\alpha = 1.62 \pm 0.2$ . Although we used this statistics as a rough guide to calibrate the C IV optical depths in three of our simulations (see the final part of Section 3), there is good agreement among *all* the simulations and the observational data at this redshift, and small deviations from the observations are within the error bars. The error bars in C IV–CDDF shown correspond to the reference run and are representative for all the simulations.

The same observable is computed at  $z = 5.6$  in Fig. 3. The values from the numerical simulations are shown in points and they are compared to systems detected by D’Odorico et al. (2013). The observational fitting function parameters for  $f(N) = BN^{-\alpha}$  are  $B = 8.96 \pm 3.31$  and  $\alpha = 1.69 \pm 0.24$  and for  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  are  $f(N_0) = 14.02$  and  $\alpha = 1.44 \pm 0.3$ . There is good agreement between the observational and theoretical points except at large



**Figure 2.** C IV CDDF at  $z = 4.8$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 10.29 \pm 1.72$  and  $\alpha = 1.75 \pm 0.13$  and the dot-dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 13.56$  and  $\alpha = 1.62 \pm 0.2$ , from the same observational work. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. Hereafter, these colours are used to represent the simulations.



**Figure 3.** C IV CDDF at  $z = 5.6$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 8.96 \pm 3.31$  and  $\alpha = 1.69 \pm 0.24$  and the dot-dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 14.02$  and  $\alpha = 1.44 \pm 0.3$ , from the same observational work. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in other models. At large column densities, simulations predict values of the C IV–CDDF slightly higher than observations. However, the error bars on all of the theoretical models overlap with the observational data.

column densities, where simulations deviate to values higher than observations. We point out that the function fits at  $z = 4.8$  and  $5.6$  extend to high column densities, even though observations are only available up to  $N_{\text{C IV}} = 10^{14.4} \text{ cm}^{-2}$ . Therefore, the outcome of the simulations gives additional information in a range of  $N_{\text{C IV}}$  where the C IV–CDDF has not been measured.

We stress that, at these redshifts, the C IV–CDDF cannot be used to disentangle different physical prescriptions in the simulations, since all the runs are in agreement within the error bars.

Finally, it is important to note that the run Ch 12 512 MDW mol predicts the highest C IV–CDDF at low column density both at  $z = 4.8$  and  $5.6$ . The higher resolution of this run with respect to all the others leads to a more complete and precise estimation of the CDDF at low  $N_{\text{C IV}}$ .

In similar works by Finlator et al. (2015, 2016) and Keating et al. (2016), strong absorbers are highly disfavoured for different simulated UVB and feedback prescriptions. In Finlator et al. (2015, 2016), the feedback mechanisms are quite efficient in enriching the IGM with metals. However, their simulated boxes are relatively small. Therefore, galaxies with large halo masses are suppressed. If these structures host the high column density systems, such simulations will also lack strong absorptions. On the other hand, Keating et al. (2016) explore various feedback prescriptions. Their SHERWOOD run has a similar hydrodynamics and configuration as our models, whereas HVEL implements a more aggressive version of the energy-driven winds used in this work (also including a minimum wind speed of  $600 \text{ km s}^{-1}$ ). Nevertheless, all their models struggle to reproduce the high C IV column density absorbers. This issue is not alleviated by varying the UVB. The presence of rare strong C IV absorbers in our theoretical models results from an appropriate level of resolution on the scale of the absorbers and a post-processing pipeline that closely mimics the method used by observers and accounts for individual features to calculate the column densities from Voigt profile fits to the absorption lines. The latter increases the accuracy in the estimation of the column densities.

## 5.2 C IV cosmological mass density

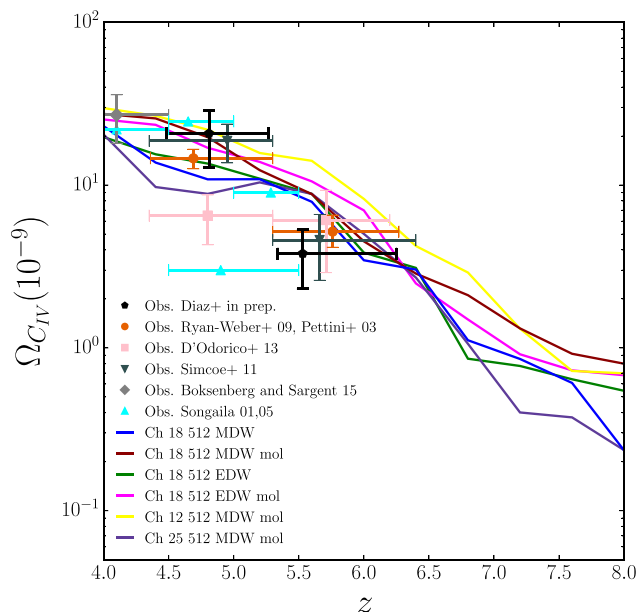
The evolution of the total density in C IV ions with respect to the critical density is described by the comoving mass density  $\Omega_{\text{C IV}}$ . For any ion,  $\Omega_{\text{ion}}$  can be derived by summing the column densities of identified absorbers:

$$\Omega_{\text{ion}}(z) = \frac{H_0 m_{\text{ion}}}{c \rho_{\text{crit}}} \frac{\sum N(\text{ion}, z)}{n_{\text{lov}} \Delta X}, \quad (17)$$

where  $m_{\text{ion}}$  is the mass of the ionic species,  $n_{\text{lov}}$  is the number of lov,  $\rho_{\text{crit}}$  is the critical density today and  $\Delta X$  is defined in equation (15).

In the case of C IV, we restricted the column densities in the sum to the range  $13.8 < \log N_{\text{C IV}}(\text{cm}^{-2}) < 15.0$ , to directly compare with the results of D’Odorico et al. (2013). The results for all of the simulations are shown in Fig. 4. The theoretical results are compared to observations by Pettini et al. (2003), Ryan-Weber et al. (2009), Songaila (2001, 2005), Simcoe et al. (2011), D’Odorico et al. (2013), Boksenberg & Sargent (2015) and Díaz et al. (in preparation). It is worth mentioning that the observations by Pettini, Ryan-Weber and Díaz have been recalibrated to the cosmology adopted in this paper. In the other cases, missing details of the precise pathlength probed do not allow us to convert the observations to the Planck cosmology.

The rise of C IV over the redshift period from 8 to 4 reflects both the increase in the chemical enrichment of the IGM through



**Figure 4.**  $C_{IV}$  cosmological mass density at  $4 < z < 8$ . Comparison between the simulated data and observations by Pettini et al. (2003) and Ryan-Weber et al. (2009) in orange circles, Songaila (2001, 2005) in cyan triangles, Simcoe et al. (2011) in dark green inverted triangles, D’Odorico et al. (2013) in pink squares, Boksenberg & Sargent (2015) in grey diamond and Díaz et al. (in preparation) in black pentagons. Pettini, Ryan-Weber and Díaz measurements are converted to the Planck cosmology, while for the others this recalibration was not possible due to missing details of the precise pathlength probed.

galactic feedback mechanisms, and the evolution of the ionization state of the IGM. At  $z = 8$ ,  $C_{IV}$  is largely suppressed and most of the carbon in the box is in its neutral state. As time passes, more SN events take place and pollute the IGM with metals (as shown in Section 5.4). Simultaneously, the specific intensity of the HM12 UVB around the wavelength where the  $C_{IV}$  transition occurs ( $\lambda_{CIV} = 192 \text{ \AA}$ ) increases by  $\sim 4$  orders of magnitude between  $z = 8$  and 4. At all redshifts, each simulation is broadly consistent with current observational data.

Using results from a hybrid wind model, Finlator et al. (2015, 2016) integrated a power-law fit to the CDDF to calculate the evolution of the comoving mass density of  $C_{IV}$ . Instead, we calculate  $\Omega_{CIV}$  by summing the column density of the absorbers using equation (17). This method is more accurate when there are uncertainties in the slope of the  $C_{IV}$ -CDDF, especially at high column densities.

### 5.3 $C_{II}$ cosmological mass density

We extend the analysis to  $C_{II}$  and compare with the observational lower limits of Becker et al. (2006) in Fig. 5, taking into account systems whose column densities are in the range  $13.0 < \log N_{CII}(\text{cm}^{-2}) < 15.0$ . In the context of reionization, it is interesting to consider the evolution of  $C_{II}$  (a low-ionization state of carbon with ionization energy of 11.26 eV) with respect to  $C_{IV}$ , whose ionization energy is more than four times larger (47.89 eV). Since the UVB is softer at earlier times,  $C_{II}$  is expected to be more dominant than higher ionization states of carbon at high redshifts.

On the left-hand and central panels of Fig. 5, we draw a comparison between different physical scenarios among the simulations with box size  $18 \text{ cMpc } h^{-1}$ : the further left relates simulations

with the same feedback prescription (MDW) with and without low-temperature metal and molecular cooling. The run with low-T cooling included produces slightly more  $C_{IV}$  (and less  $C_{II}$ ) than the fiducial run, but the difference is only significant at  $z > 6.5$ . In the middle panel, we contrast simulations without low-temperature metal and molecular cooling with different feedback prescriptions: MDW and EDW. The different physical scenarios do not have a strong impact on the evolution of  $C_{IV}$ , as seen also in Fig. 4. The trends for  $C_{II}$  are more dependent on the model, because this low-ionization state is very sensitive to the densities of the absorbers (which are correlated to the strength of the feedback prescription), and the Ch 18 512 EDW run is in good agreement with the observations of Becker et al. (2006).

Although the run with lowest resolution, Ch 25 512 MDW mol, produces slightly degraded results, the evolution of  $\Omega_{CIV}$  does not change significantly with the spatial resolution or size of the simulations (see the rightmost panel of Fig. 5).

On the other hand,  $C_{II}$  trends are more influenced by the resolution. Ch 12 512 MDW mol better reproduces the observed evolution of this ion. The absorption features are affected by the box size (and the comoving softening) for different reasons: the absorption path calculated to create the synthetic spectra grows with the box size, as well as the size of the pixels, while the definition of the individual features decreases for larger configurations. The Ch 12 512 MDW mol run is in better agreement with the observational limits for  $C_{II}$  at high redshift indicating that higher resolution in the simulations helps to better describe the absorption features, in particular for low-ionization states.

The most interesting part of the analysis arises when the  $C_{II}$  and  $C_{IV}$  curves are compared for each model: in the fiducial simulation, the amount of  $C_{IV}$  exceeds the amount of  $C_{II}$  at  $z \sim 6$ , consistent with the fact that the Universe is approaching the tail of the EoR. This result is seen in all of the simulations, with a crossover in the range of  $z \sim 6-6.5$ .

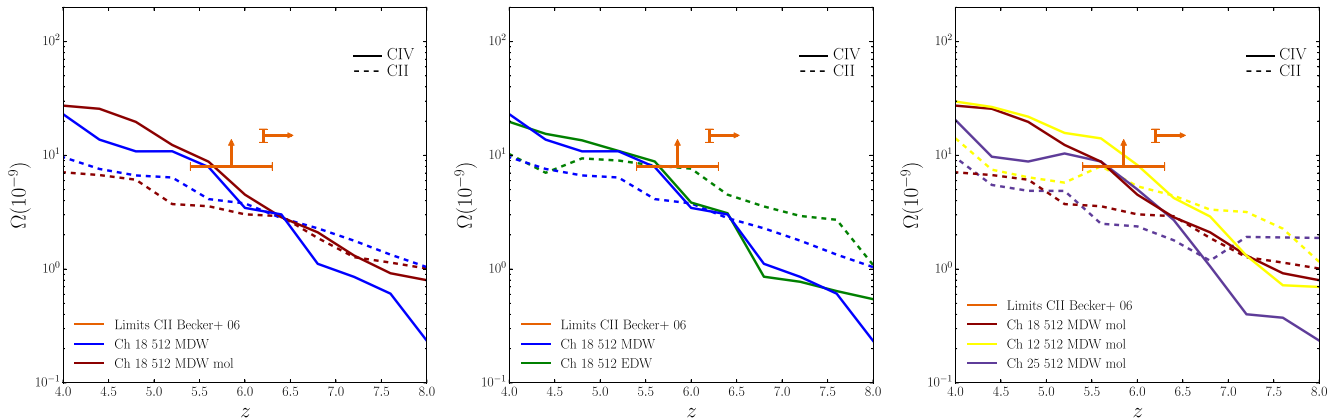
Despite the fact that in some runs the synthetic  $C_{II}$  does not reach the lower limits predicted in Becker et al. (2006), the crossover of these ions at  $z \sim 6-7$  is very promising. Other theoretical works that also follow the trend of these ions (e.g. Finlator et al. 2015) obtained a crossover between  $\Omega_{CII}$  and  $\Omega_{CIV}$  at  $z \sim 8$ . However, a later crossover of  $C_{II}$  and  $C_{IV}$  at  $z \sim 6$  is more consistent with the observations and the current paradigm of the tail of reionization.

### 5.4 Total carbon cosmological mass density

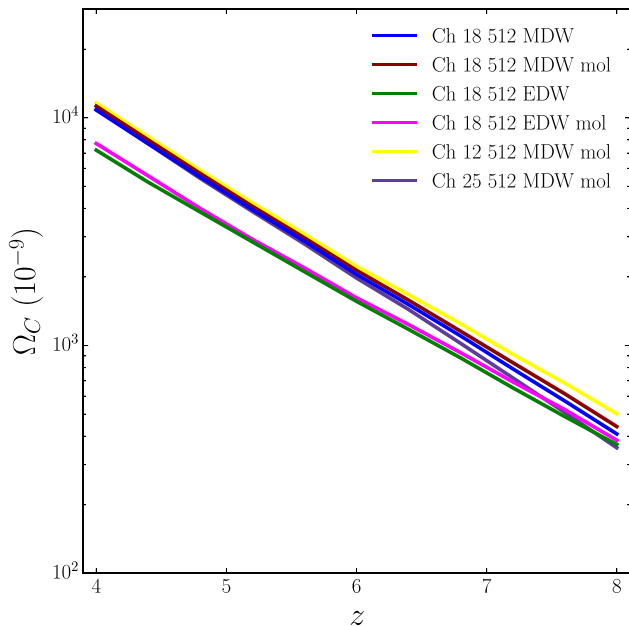
To conclude Section 5, in Fig. 6 we analyse the evolution of the cosmological mass density of (total) carbon in the simulated boxes,  $\Omega_C$ . This is calculated differently than  $\Omega_{CIV}$  and  $\Omega_{CII}$ ; here, we sum the amount of carbon in each gas particle in the simulation and divide by the comoving volume. Thus,  $\Omega_C$  gives an estimation of the total amount of carbon at high redshift. The trend is consistent with the enrichment history of the Universe, increasing by 1.5 orders of magnitude from  $z = 8$  to 4. At  $z = 4$ , our results are compatible with those of Tescari et al. (2011). Given the different methods used to calculate  $\Omega_C$  and  $\Omega_{CIV}$  &  $\Omega_{CII}$ , the first quantity should not be directly compared with the other two. Moreover, the evolution of the total carbon cosmological mass density is a purely theoretical prediction of our simulations, since  $\Omega_C$  is not (directly or indirectly) observable.

The overall evolution of C is nearly independent on the nature of the cooling (blue versus red and green versus magenta lines), but depends at low redshift on the feedback mechanism. The EDW





**Figure 5.** Evolution of the C II and C IV cosmological mass density. The left-hand and central panel compare simulations with 18 cMpc  $h^{-1}$  box size. On the left is displayed the case of simulations with MDW feedback with and without low-temperature metal and molecular cooling included. In the middle, we show the configuration without low-T metal and molecular cooling for the runs with MDW and EDW feedback. On the right-hand side, a resolution and box size test with the following models is shown: Ch 12 512 MDW mol, Ch 18 512 MDW mol and Ch 25 512 MDW mol. In all the panels, the solid lines show the evolution of  $\Omega_{\text{CIV}}$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , and the dashed lines  $\Omega_{\text{CII}}$  in the range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$ . The orange points with errors represent the observational lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006).



**Figure 6.** Evolution of the total carbon cosmological mass density for our simulations. This is a purely theoretical prediction, since  $\Omega_{\text{C}}$  is not (directly or indirectly) observable.

models quench the formation of stars in galaxies more effectively than MDW, therefore less carbon is produced and  $\Omega_{\text{C}}$  is lower.

Fig. 6 also compares simulations with different spatial resolutions (the comoving softening of the different configurations scales with the size of the box). The Ch 12 512 MDW mol, Ch 18 512 MDW mol and Ch 25 512 MDW mol runs have comoving softening of 1.0, 1.5 and 2.0 cMpc  $h^{-1}$  (yellow, red and purple lines), respectively. These simulations predict a similar trend for the mass density of carbon and at low redshift converge to the same  $\Omega_{\text{C}}$ . However, at high redshift the resolution affects the amount of metals. The run with the highest resolution (Ch 12 512 MDW mol, yellow line) can resolve higher densities at earlier times and therefore better describes star formation in galaxies and produces more carbon than

Ch 18 512 MDW mol (red) and, particularly, Ch 25 512 MDW mol (purple).

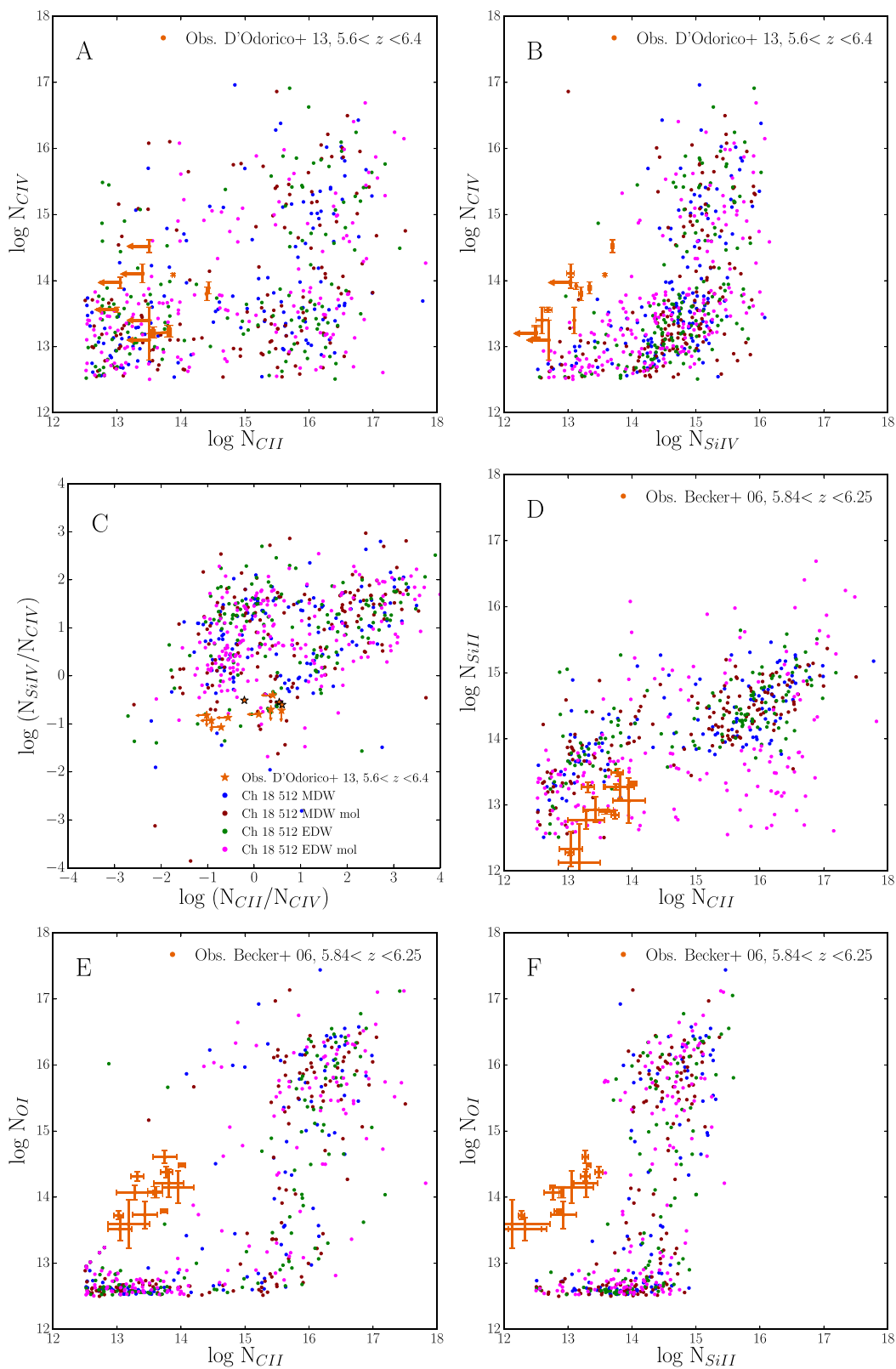
## 6 COLUMN DENSITY RELATIONSHIPS OF METAL ABSORPTION LINES

One of the advantages of the numerical approach is that it reproduces a large number of sightlines, improving the statistical estimation of the column densities at redshifts that are hard to access with current observations. By convolving our synthetic spectra with Gaussian noise and extracting column densities using Voigt profile fitting, our simulated results closely mimic the observed spectra.

As mentioned in Section 5.1, this procedure produces a very large number of systems with low column densities, which display non-physical linear correlations among the metal absorption features. To avoid this issue, a cut-off in column density,  $\log N_{\text{th}}(\text{cm}^{-2}) \geq 12.5$ , has been imposed to the synthetic data, and systems below this noise threshold have not been taken into account in the statistics.

We explore relationships in column densities among different ionic species in a uniform UVB HM12 at  $z = 6$  using random lines of sight to emulate the observational method and compare with the observations available to date in Fig. 7. In panel A,  $N_{\text{CII}}$  versus  $N_{\text{CIV}}$  are compared, panel B displays  $N_{\text{SiIV}}$  versus  $N_{\text{CIV}}$  and panel C the ratio of  $N_{\text{CII}}/N_{\text{CIV}}$  versus  $N_{\text{SiIV}}/N_{\text{CIV}}$ . All these quantities are compared to systems detected by D’Odorico et al. (2013), while panels D, E and F draw a comparison with low-ionization states observed by Becker et al. (2006). The comparisons  $N_{\text{CII}}$  versus  $N_{\text{SiII}}$ ,  $N_{\text{CII}}$  versus  $N_{\text{OI}}$  and  $N_{\text{SiII}}$  versus  $N_{\text{OI}}$  are shown in panels D, E and F, respectively. In all cases, we display the simulations with 18 cMpc  $h^{-1}$  box size to analyse possible variations among the predicted values with different prescriptions of feedback and cooling. However, the dispersion of the data points do not allow clear distinctions between the models.

The high-ionization states are relatively well represented by the simulated data with the UV ionizing background implemented, especially  $N_{\text{CIV}}$  as seen in panels A and C. On the other hand, the low-ionization species are produced far in excess by the simulations and in all the different configurations considered the column densities are overestimated for O I and Si II.



**Figure 7.** Column density relationships among metal absorption lines at  $z = 6$ . Panels A, B and C compare with systems observed by D’Odorico et al. (2013) in the redshift range of (5.6–6.4) and D, E and F with absorbers detected by Becker et al. (2006) in the redshift range of (5.84–6.25), both in orange. From the synthetic spectra of each ion, the column densities are calculated and displayed for the simulations with box size of  $18 \text{ Mpc } h^{-1}$  (the label keys are shown in the panel C), but there is not a remarkable difference between the different prescriptions (feedback and cooling) imposed in each case. The simulated data displayed are above the threshold noise at  $\log N_{\text{H}} = 12.5 \text{ (cm}^{-2}\text{)}$ . This selection criterion results in an effective number of synthetic systems of  $\sim 18$  per cent from the initial 1000 random lines of sight traced. In panel C, the observations by D’Odorico et al. (2013) produced three systems with column densities reported (represented by the black stars) and eight with upper limits in  $N_{\text{CII}}$  and/or  $N_{\text{SiIV}}$ , represented by arrows.

$C_{\text{II}}$  is a very interesting case in this study. The numerical estimation of  $C_{\text{II}}$  column densities do not seem to be well reproduced by the simulations, although in most of the cases the observational data are just a lower limit, and a bimodality in the distribution of column densities appears, which will be analysed in Section 6.1.

The scatter plots in many cases struggle to reproduce the observed column densities, regardless of the feedback model used, possibly due to the UVB implemented. The parameters that characterize HM12 can suppress the low-ionization states. Bolton & Viel (2011) claimed that  $N_{\text{SiIV}}/N_{\text{CIV}}$  is sensitive to the spectral shape of the UVB, whereas Finlator et al. (2015) showed that  $N_{\text{CII}}/N_{\text{CIV}}$  is sensitive to the overall intensity of the UVB and its normalization. We are currently studying the effect of variations in the normalization and hardness of the HM12 UVB. The results will be presented in a follow-up paper.

$O\text{I}$  receives special attention because its ionization energy is similar to the one of  $\text{H I}$ . Besides the overprediction of  $N_{\text{OI}}$  in our simulations,<sup>5</sup> the large column densities in this ionic species may indicate that there is a huge amount of  $O\text{I}$  self-shielded that it is not being detected due to the low probability of tracing a high-density region with the current observational methods. It is worth noting that, to date, there is not a self-consistent scheme in the simulation literature that accounts for the self-shielding of the ions (corrections for self-shielding at  $z < 5$  can be found in Bird et al. 2014; Bolton et al. 2017). This could make the estimated column densities of the low-ionization states less accurate.

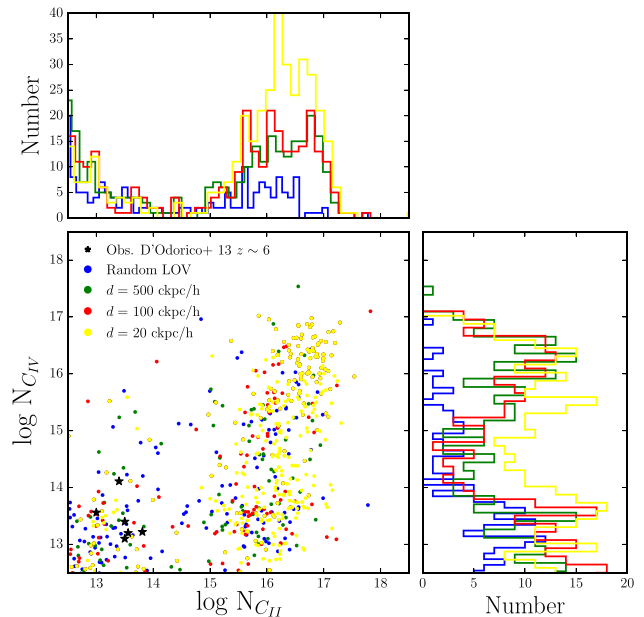
### 6.1 Bimodality in $C_{\text{II}}$ column densities

The location of the low-ionization states of metals at  $z \sim 3$ , for instance  $C_{\text{II}}$  and  $O\text{I}$ , traces the regions close to the centre of the galaxies or DLA systems. This correlation of gas ionization state and proximity to the centre of the nearest galaxy may no longer be valid at high redshift, as  $C_{\text{II}}$  may be tracing either DLAs or the IGM at low temperature.

Observations at  $z \sim 6$  exhibit  $N_{\text{CII}} \leq 10^{14} \text{ cm}^{-2}$ , while at higher redshift only lower limits in  $C_{\text{II}}$  are available. In all the simulations, we found a remarkable bimodal distribution in the  $N_{\text{CII}}$  calculated using random sightlines through each box (see e.g. panel A in Fig. 7). In order to test the nature of this bimodality, in Fig. 8 we plot the column densities of  $C_{\text{II}}$  and  $C_{\text{IV}}$ , using lines of sight at different impact parameters from the centre of the galaxies in the fiducial run Ch 18 512 MDW ( $d = 20, 100$  and  $500 \text{ ckpc } h^{-1}$  in yellow, red and green points, respectively). At each  $d$ , we extracted a thousand low around haloes with masses in the range  $10^9\text{--}10^{10} M_{\odot} h^{-1}$ . The blue points represent the column densities along 1000 random lines of sight through the box shown in panel A of Fig. 7. The black stars are observational data from D’Odorico et al. (2013). The top panel in Fig. 8 displays the distribution of the  $C_{\text{II}}$  absorption features and the rightmost window the corresponding distribution of  $C_{\text{IV}}$  column densities above the noise threshold.

When the impact parameter  $d$  approaches the galaxies, the peak at high column densities becomes sharper, indicating that there are more  $C_{\text{II}}$  absorbers in the CGM where the temperature and density are larger than in the IGM. For  $d = 20 \text{ ckpc } h^{-1}$ ,  $\sim 3 \text{ pkpc}$  at this redshift, the absorbers are located inside or in the outskirts of the galaxies (spectroscopic confirmation of galaxies at  $z = 6\text{--}8$  with

<sup>5</sup> There is also a constant trend of  $N_{\text{OI}}$  at low column densities, possibly indicating that in low-density environments oxygen is mostly found in higher ionization states due to the effect of the UVB.



**Figure 8.**  $C_{\text{II}}$  versus  $C_{\text{IV}}$  column densities in different environments at  $z = 6$  for the reference simulation, Ch 18 512 MDW. Using different impact parameters from the centre of the galaxies, it is possible to understand what drives a bimodality in  $N_{\text{CII}}$ . This low-ionization state of  $C$  is mostly found in the CGM, but there are some traces of it in the IGM and, so far, it has been observed in the latter, with low column densities, as it is shown in the plot (the black stars are observational data at  $z = 6$  from D’Odorico et al. 2013). The blue points are theoretical predictions using random lines of sight (cf. panel A in Fig. 7). When the sightlines are chosen to be close to the centre of the galaxies, the distribution in  $C_{\text{II}}$  is shifted to larger column densities, indicating that high-redshift galaxies are huge reservoirs of  $C_{\text{II}}$ , not detected yet due to the low likelihood of reaching these overdense regions with QSOs’ lines of sight.

the *Hubble Space Telescope* shows that the radii of these galaxies are in the range of  $0.6\text{--}1.1 \text{ pkpc}$ ; Jiang et al. 2013). Absorbers are gravitationally bounded to the haloes in the simulated sample, which have virial radii of  $3\text{--}30 \text{ pkpc}$ .

The distribution of the  $C_{\text{II}}$  column densities at  $d = 20 \text{ ckpc } h^{-1}$  indicates that the few observations towards high-redshift QSOs currently available have not yet detected rare systems with large column densities in  $C_{\text{II}}$ , due to the very low likelihood of intersecting a galaxy at a low impact parameter (that would produce a deep absorption feature). Furthermore, D’Odorico et al. (2013) have only identified  $C_{\text{II}}$  absorptions that were first detected as  $C_{\text{IV}}$  doublets. Thus, any low-ionization gas that may have produced a strong  $C_{\text{II}}$  absorber and no associated  $C_{\text{IV}}$  has not been reported. However, future observations of QSOs will raise the number of metal absorption lines detected and cover a larger range in column densities.

The results for  $N_{\text{CII}}$  with random lines of sight (blue line) agree with numerical results by Oppenheimer et al. (2009), Keating et al. (2016) and a photoionization comparison by D’Odorico et al. (2013), predicting that  $C_{\text{II}}$  lies mostly in overdense regions with  $\delta = \frac{\rho_{\text{gas}}}{\langle \rho \rangle} - 1 > 10$ , and there are some traces in low-temperature regions (mostly the IGM probed by current observations) where its column densities are of the same order of  $N_{\text{CIV}}$ .

Finally, numerical results in the CGM (yellow and red cases in Fig. 8) lead to a bimodality in  $C_{\text{IV}}$  and a clear excess of high column density systems at low impact parameters. Oppenheimer et al. (2009) showed that the properties of the  $C_{\text{IV}}$  absorbers depend

mostly on the evolution of the environment of their hosting galaxy (distance, mass and metallicity). By tracing the surroundings of the galaxies in haloes with masses  $10^9$  to  $10^{10} M_{\odot} h^{-1}$ , we confirm that there is a correlation between C IV column densities and the proximity of the galaxies that host the absorbers.

## 7 COSMOLOGICAL MASS DENSITY OF H I AT REDSHIFT LOWER THAN 6

The estimation of the amount of neutral hydrogen can be derived using DLAs. It is worth noting that at high redshift ( $z > 5$ ) the nature of DLAs is ambiguous (e.g. Simcoe et al. 2012), since the definition of a DLA is by column density only ( $N_{\text{HI}} > 10^{20.3} \text{ cm}^{-2}$ ) and does not discriminate between dense pockets of neutral hydrogen in the IGM, where reionization is still progressing, and DLAs associated with self-shielded gas in galaxies (like at  $z < 5$ ).

However, it is possible to simulate DLA systems to calculate their corresponding comoving mass density  $\Omega_{\text{DLA}}$  and from it, estimate  $\Omega_{\text{HI}}$ . We identify the haloes with masses above  $10^9 M_{\odot} h^{-1}$  (using a halo finder routine, the haloes are sorted by mass) and extract cubes of size  $200 \text{ ckpc } h^{-1}$  around the centre of mass of one thousands of them (randomly distributed in mass). In this way, we guarantee that the simulated DLAs are not affected by a selection bias in density. Then, the neutral hydrogen density in the cube is integrated along the line of sight and projected into a 2D grid of  $N_{\text{HI}}$ . With these column densities, we finally calculate the H I-CDDF,  $f_{\text{HI}}(N, X)$ , as

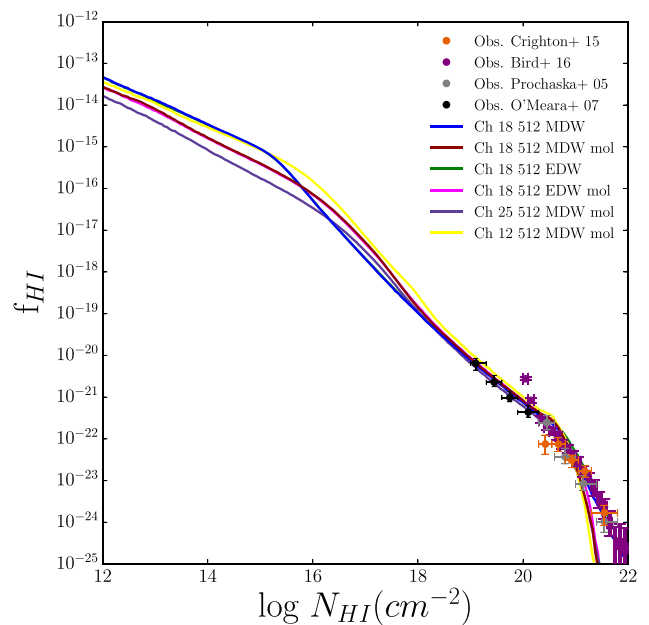
$$f_{\text{HI}}(N, X) = \frac{n_{\text{sys}}(N, N + \Delta N)}{\Delta X \cdot n_{\text{cube}} \cdot n_{\text{grid}}^2}, \quad (18)$$

where  $n_{\text{sys}}$  is the number of systems in the interval  $(N, N + \Delta N)$ ,  $\Delta X$  the absorption path defined in equation (15),  $n_{\text{cube}}$  the number of cubes generated and  $n_{\text{grid}}$  the number of bins in the grid. The parameters selected to achieve a good resolution of the systems that contain neutral hydrogen are:  $n_{\text{cube}} = 1000$  and  $n_{\text{grid}} = 64$  (i.e. the grid size is  $3.1 \text{ ckpc } h^{-1}$ ).

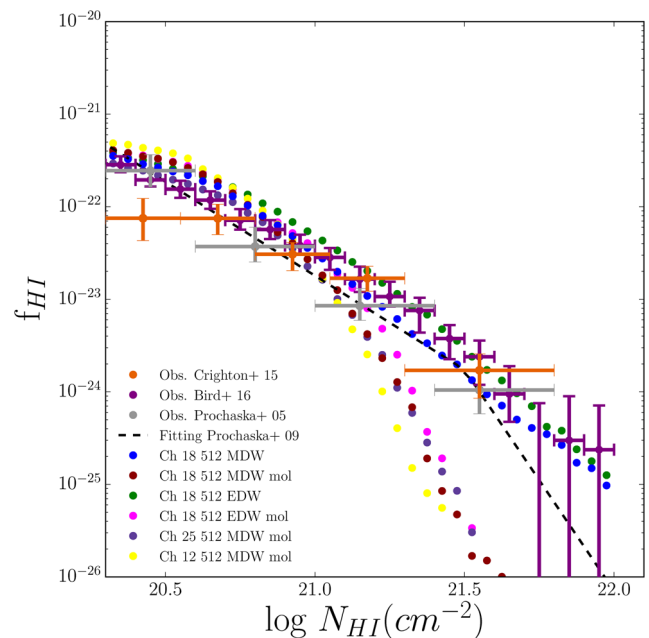
Fig. 9 shows the H I-CDDF, with column densities in the range  $12 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$ , for all the simulations at  $z = 4$  and compares with the observations of Prochaska et al. (2005), Crighton et al. (2015), Bird, Garnett & Ho (2017), all in the column densities corresponding to DLA systems at  $z \geq 4$ ) and O’Meara et al. (2007) in the super Lyman limit systems (LLS) range at  $z \sim 3.4$ .<sup>6</sup> Note that there is a very large number of systems with low column densities in all the models and DLAs are extremely rare. The agreement with the observational data is good, indicating that the method implemented to find neutral systems works well at these redshifts.

Simulations with smaller comoving softening (or smaller box size) better resolve the Ly  $\alpha$  forest. In Fig. 9, the yellow line (simulation Ch 12 512 MDW mol) is above the red line, corresponding to the run Ch 18 512 MDW mol, and the latter is above the purple line (Ch 25 512 MDW mol, the run with the lowest resolution in the suite of simulations) at  $\log N_{\text{HI}}(\text{cm}^{-2}) < 19$ .

At DLA column densities, the distributions follow a double power law, as discussed by Zwaan & Prochaska (2006) and Prochaska & Wolfe (2009), with a knee at  $\log N_{\text{HI}}(\text{cm}^{-2}) \sim 21$  in all of the cases considered. Also, at  $N_{\text{HI}} > 10^{21} \text{ cm}^{-2}$ , where molecular



**Figure 9.** H I CDDF at  $z = 4$  in the range  $12 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$  for all the simulations described in Table 1 and comparison to observations by Prochaska et al. (2005) in grey, O’Meara et al. (2007) in black, Crighton et al. (2015) in orange and Bird et al. (2017) in purple.

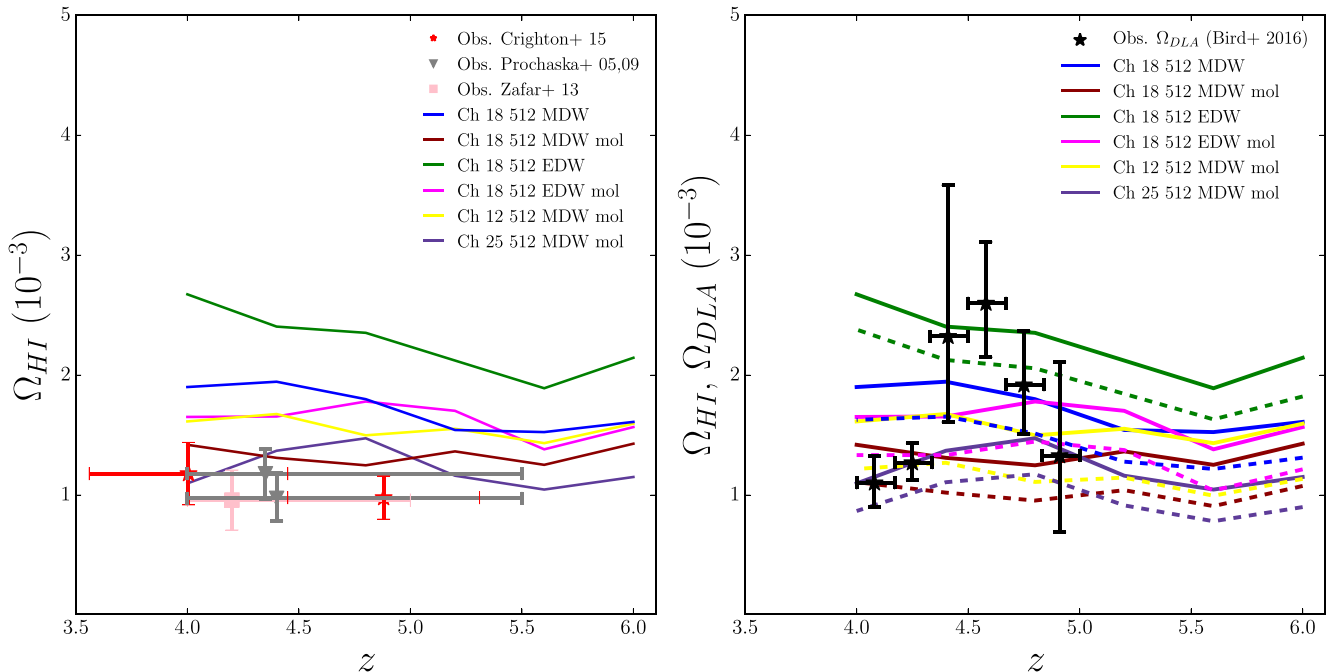


**Figure 10.** Same as Fig. 9 but zoomed in the DLA range:  $20.3 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$ . The black dashed line represents the fitting function by Prochaska & Wolfe (2009) for DLA systems at redshift 4.0–5.5.

cooling plays an important role in the star-forming regions, the H I-CDDF in models with molecular cooling implemented drops off more rapidly than in the other models.

In Fig. 10, a zoom of the DLA region is shown ( $\log N_{\text{HI}}(\text{cm}^{-2})$  in the range [20.3, 22.0]). Overplotted are the most recent observational data of H I absorbers at high redshift by Prochaska et al. (2005), Crighton et al. (2015), Bird et al. (2017) and the fitting function proposed by Prochaska & Wolfe (2009) for DLA systems

<sup>6</sup> Although these systems are at  $z < 4$ , they are a reliable benchmark for our simulated data, since there is evidence of low to no evolution in this redshift range.



**Figure 11.** Cosmological mass density of H I. In the left-hand panel, we display the theoretical prediction of  $\Omega_{\text{HI}}$  and compare with the data available at high redshift: Prochaska et al. (2005) and Prochaska & Wolfe (2009) in grey inverted triangles and Zafar et al. (2013) in pink square and Crighton et al. (2015) in red stars. The right-hand panel shows  $\Omega_{\text{HI}}$  and  $\Omega_{\text{DLA}}$  (solid and dashed lines, respectively). We compare the resulting  $\Omega_{\text{DLA}}$  with recent data from SDSS (Bird et al. 2017) in black stars. In the observational work of Crighton et al. (2015), it is proposed that more than 80 per cent of the contribution to the H I cosmological mass density at  $z < 5$  comes from DLA systems. We confirm this assumption theoretically (Table 2).

at redshift 4.0–5.5. The simulations are compatible with the observational data, especially runs with no molecular cooling.

The main distinction between the models with and without molecular cooling is driven by the conversion of neutral hydrogen to  $\text{H}_2$ , which becomes important in high-density regions where new stars are being formed. Molecules boost the cooling of the surrounding gas and the formation of stars. As a result, the amount of atomic hydrogen decreases.

At this point, we calculate the comoving mass density of H I, using the following equation:

$$\Omega_{\text{HI}}(z) = \frac{H_0 m_{\text{HI}}}{c \rho_{\text{crit}}} \int_{N_{\text{min}}}^{N_{\text{max}}} f_{\text{HI}}(N, z) N dN, \quad (19)$$

with  $m_{\text{HI}}$  the mass of the hydrogen atom and  $f_{\text{HI}}(N, z)$  defined in equation (18). In theory, the integral should be computed with an upper limit  $N_{\text{max}} = \infty$ , but in reality this value is set by the maximum column density detected. Thus, the limits considered for  $\Omega_{\text{HI}}$  are  $12 < \log N_{\text{HI}} (\text{cm}^{-2}) < 22$ .

We compare the predictions from the theoretical models to observations at high redshift by Prochaska et al. (2005), Prochaska & Wolfe (2009), Zafar et al. (2013) and Crighton et al. (2015) in the left-hand panel of Fig. 11. The amount of H I at  $z = 4$  in the Ch 18 512 MDW mol (red line) and Ch 25 512 MDW mol (purple line) runs is in excellent agreement with the observational data, probably because the inclusion of molecular cooling prevents the production of too much neutral hydrogen (by converting it into molecular hydrogen). On the other hand, simulations without molecular cooling overproduce  $\Omega_{\text{HI}}$ , especially in the EDW case (green line). All the models predict an increasing amount of H I at  $z > 5.5$ . However, we stress again that our assumption of a uniform UVB is no longer valid to describe the diffuse H I in the IGM when approaching  $z =$

**Table 2.** Contribution from DLA systems to  $\Omega_{\text{HI}}$ .

Simulation	$\alpha = \frac{\Omega_{\text{HI}}}{\Omega_{\text{DLA}}}$	per cent
Ch 18 512 MDW	$1.19 \pm 0.03$	80.9
Ch 18 512 MDW mol	$1.25 \pm 0.02$	75.3
Ch 18 512 EDW	$1.15 \pm 0.03$	84.9
Ch 18 512 EDW mol	$1.22 \pm 0.02$	77.8
Ch 12 512 MDW mol	$1.29 \pm 0.04$	70.4
Ch 25 512 MDW mol	$1.21 \pm 0.02$	78.7

6. At these redshifts, our simulated  $\Omega_{\text{HI}}$  only describes the amount of neutral hydrogen in collapsed/self-shielded systems. Please note that for this calculation we also extended the H I self-shielding prescription of Rahmati et al. (2013), which was originally proposed only up to redshift  $z = 5$ .

In order to estimate the contribution of DLAs to the overall H I budget, the right-hand panel of Fig. 11 compares the prediction of  $\Omega_{\text{HI}}$  for all the simulations (solid lines) with the DLA comoving mass density  $\Omega_{\text{DLA}}$  (dashed lines), obtained by integrating the H I–CDDF in the range  $20.3 < \log N_{\text{HI}} (\text{cm}^{-2}) < 22$ . The theoretical estimations are compared with recent data from SDSS (Bird et al. 2017) in black stars. In this case, simulations without molecular cooling included are in better agreement with observations (as for the H I–CDDF shown in Fig. 10).

We investigate the assumption  $\Omega_{\text{HI}} = \alpha \times \Omega_{\text{DLA}}$ , where  $\alpha$  is a factor that accounts for the numbers of systems that contain neutral hydrogen with  $N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$ . The values for  $\alpha$  in each simulation (averaged in the redshift range 4–6) are in Table 2. In the observational work of Crighton et al. (2015), the authors estimate a 20 per cent contribution from systems with column density below the DLA threshold (or 80 per cent contribution from DLAs) at

$z < 5$ . This assumption seems to be validated with the numerical results here reported.

## 8 DISCUSSION AND SUMMARY

The purpose of this work is to investigate the physical environment of the IGM in proximity of galaxies at the end of the EoR by producing mock spectra and measuring ionic column densities of metal transitions with a method that closely resembles current observational techniques. Taking advantage of a suite of simulations with a very robust model for chemical enrichment, SN-driven feedback that reproduces the cosmic SFR density and the galaxy stellar mass function at high redshift and a suitable comoving softening at the IGM scale ( $1\text{--}2\text{ ckpc } h^{-1}$ ), we reproduce well some observables using metal and H I absorption lines at high redshift. In particular, the comoving mass density  $\Omega$  of C IV and C II in the redshift range of  $4 \leq z \leq 8$ , the H I CDDF at  $z = 4$  and  $\Omega_{\text{HI}}$  &  $\Omega_{\text{DLA}}$  at  $z < 6$ . In all cases, we compare the theoretical results with the latest release of observational data.

### 8.1 Effect of wind feedback

The simulations we used in this work are based on the analysis of Tescari et al. (2014), where there is an extensive discussion of the feedback mechanisms implemented and how they successfully reproduce the observations of the SFR at high redshift (e.g. Madau & Dickinson 2014). When these models are used to study the evolution of different ionic species, we find that it is hard to disentangle the effect of different SN-driven galactic wind prescriptions for some of the statistics considered, in particular the C IV and H I CDDF. The reasons are mostly numerical: probably higher resolution is needed to better resolve patches of enriched gas at relatively low density in the CGM/IGM (as suggested by our resolution tests). Moreover, the SPH algorithm is not very effective in the mixing of metals. Therefore, only extreme feedback models would produce significant differences in the observables at high redshift. As a result, when calculating the column densities of metal absorption lines with Voigt profile fitting at  $z > 5$ , it is not possible to distinguish between momentum-driven and energy-driven winds (MDW and EDW). The dispersion in the relationships among the metal absorption lines (see Fig. 7) is too high to provide definitive hints on the evolution of the ionization states due to one particular wind model.

None the less, a clear difference between the numerical estimation of the total carbon content,  $\Omega_{\text{C}}$ , in EDW and MDW models arises at  $z \leq 6$  (Fig. 6). This is due to the fact that EDW quenches the overall star formation in galaxies more effectively than MDW, and consequently the production of carbon in the simulation is 0.2–0.3 dex lower at  $z = 4$ .

### 8.2 Influence of the low-temperature metal and molecular cooling

As expected, the implementation of low-temperature metal and molecular cooling affects mostly neutral hydrogen statistics. In particular, molecular cooling suppresses H I in regions where the majority of it is self-shielded, mainly because of the conversion of H I to H<sub>2</sub> at high densities. This effect is important when computing the H I CDDF and cosmological mass density. Interestingly, the introduction of molecular cooling improves the estimation of  $\Omega_{\text{HI}}$  at  $z \sim 4\text{--}5$  (left-hand panel of Fig. 11), whereas, when the H I-CDDF is computed in the DLA regime,  $20.3 < \log N_{\text{HI}} (\text{cm}^{-2}) < 22$ , runs

without molecular cooling provide a better match to new SDSS observations (Bird et al. 2017) of  $\Omega_{\text{DLA}}$  (right-hand panel of Fig. 11) and the distribution function at  $z = 4$  (Fig. 10).

Once the feedback prescription is fixed, simulations with low-T metal and molecular cooling included produce slightly more C IV than the other runs (Fig. 4 and left-hand panel of Fig. 5). The effect is less important for C II and not visible for the total carbon content and the dispersion in the relationships among the metal absorption lines (Figs 6 and 7).

### 8.3 Limitations

Even though our models are in good agreement with observations of ionic species at high redshift, there are some caveats to consider.

First of all, radiative transfer effects are not included. The simulations are run in a relatively small volume (the maximum size considered is  $25\text{ cMpc } h^{-1}$ ), whereas a typical H II bubble at the tail of EoR is larger than  $100\text{ cMpc } h^{-1}$ . Therefore, we assume that our boxes represent a region of the Universe already reionized by a uniform HM12 UV ionizing background at  $z \leq 8$ . Although full RT calculations would be more accurate, this is a fair approximation to describe the ionization state of metals (Finlator et al. 2015), since chemical enrichment at these redshifts occurs mostly inside and in close proximity of galaxies. On the other hand, even in small volumes the distribution of H I optical depths cannot be properly described with a uniform UVB or a simple model that assumes galaxies and quasars as ionizing sources and uses a fixed mean free path for the ionizing photons (Becker et al. 2015b). For this reason, we have studied neutral hydrogen statistics only at  $z < 6$ .

A necessary step to complement the analysis of this work will be to vary the spectral hardness and the normalization of the UVB, in order to refine the calculation of the column densities of some ionic species, that are not well represented with the assumed uniform HM12 UVB. We plan to investigate this in a subsequent paper.

In recent years, numerical modelling of absorption features have improved enormously, but the current generation of simulations still struggles to reproduce low-ionization states of metals at high redshift. This is mainly due to insufficient resolution, but it may also be influenced by the lack of a consistent implementation of self-shielding prescriptions for the ions. This work (like many others) just considers the effect of H I self-shielding (Rahmati et al. 2013), but does not introduce any self-shielding of low-ionization absorbers (which lay in clumpy structures) that has been proposed for DLA systems at lower redshift (Bird et al. 2015).

Finally, there is a fundamental problem with the lack of reliable measurements of the mean normalized flux of the Ly  $\alpha$  forest at  $z \geq 5$ , when the QSOs spectra display very large Gunn–Peterson troughs, which leads to the absence of an effective parametrization of the optical depth of H I to re-scale or calibrate the simulated hydrogen and ion spectra. Different approaches have been used in the literature to reproduce the evolution of the metal ions independently from H I. In this work, we calibrate our simulations using the C IV CDDF at  $z = 4.8$  and  $5.6$ .

## 9 CONCLUSIONS

We have introduced high-resolution hydrodynamical simulations to study the physical and chemical state of the IGM at the end of the EoR. We have tested the effect of a low-temperature metal and molecular cooling model and two prescriptions for galactic wind feedback on the evolution of metal absorption lines in the redshift range  $4 < z < 8$ . We have also studied the H I cosmological mass

density at  $z < 6$ . Our theoretical predictions are consistent with the available observations at high redshift.

The evolution of the high- and low-ionization states of carbon is the main focus of our work. The drop in the cosmological mass density of C IV,  $\Omega_{\text{CIV}}$ , from  $z = 4$  to 8 is due to the combined effect of a change in the ionization state of the gas and the decreased metallicity of the IGM (Fig. 4). In fact, the total carbon comoving mass density at  $z = 8$  is more than a factor of 10 lower than at  $z = 4$  (Fig. 6), and the simulations show a consistent transition to more neutral states of metals at high redshift (as shown in Fig. 5). Most notably, the crossover between  $\Omega_{\text{CIV}}$  and  $\Omega_{\text{CII}}$  happens at  $z \sim 6\text{--}6.5$ , in agreement with available observations and the current paradigm of the tail of reionization.

In comparison to other numerical works in the literature, we are able to produce C IV absorbers in the IGM with large column densities. This is mostly due to a combination of efficient galactic winds and a post-processing pipeline that mimics observational methods and, in particular, accounts for individual features to calculate the column densities from Voigt profile fits to the absorption lines.

The simulated C II exhibits a bimodal distribution with large absorptions in and around galaxies, and some traces in the lower density IGM. These latter correspond to systems detected with current observations (Becker et al. 2006; D’Odorico et al. 2013). We predict that, at high redshift, most of the high column density C II, with  $\log N_{\text{CII}} (\text{cm}^{-2}) > 15$ , should lie in proximity of galaxies at impact parameters of the order of  $20 \text{ kpc } h^{-1}$  (Fig. 8) and has not been detected yet due to the low likelihood of reaching these overdense regions with lines of sight towards distant QSOs.

We have studied the column density relationships among different ionic species (Fig. 7). High-ionization states (like  $N_{\text{CIV}}$ ) are reasonably well described by our simulations with the adopted HM12 UVB. On the other hand, a comparison with observations in the literature shows that the low-ionization states are not well represented, regardless of the feedback model implemented.

Finally, our simulations are in good agreement with observations of the H I CDDF at  $z = 4$  (Figs 9 and 10) and the H I cosmological mass density,  $\Omega_{\text{HI}}$ , at  $4 < z < 6$  (Fig. 11). We validate the estimate made in the observational work of Crighton et al. (2015) that DLA systems contribute to  $\sim 80$  per cent of  $\Omega_{\text{HI}}$  (see Table 2).

In a companion paper (García et al. 2017), we have explored the likelihood of reproducing the observed Ly  $\alpha$  emitter galaxy – C IV absorption pair detected by Díaz et al. (2015) and studied the physical processes that produced the metal enrichment in the IGM at  $z \geq 5.6$ . Future work will be focused on analysing more metals and ions to have a broader perspective on the evolution of the IGM at the end of the EoR. Some of the column density relationships are extremely sensitive to the ionizing background, therefore varying the normalization and hardness of the UVB is the next step to obtain a more realistic description of this cosmic era.

## ACKNOWLEDGEMENTS

Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. The numerical simulations were run using the Raijin distributed-memory cluster from the NCI facility, post-processed with the Edward HPC machine from the University of Melbourne and analysed with the gSTAR supercomputer at Swinburne University of Technology. This work was supported by the Flagship Allocation Scheme of the NCI National Facility at the ANU. The authors acknowledge CAASTRO for funding and allocating time for the project *Diagnosing hydro-*

*gen reionization with metal absorption line ratios* (fy6) during 2015 and 2016. LAG thanks Valentina D’Odorico and Neil Crighton for providing their observational data in private communications and Jeff Cooke for the insightful discussions. ERW acknowledges ARC DP 1095600. We also thank the anonymous referee for providing valuable comments that considerably improved our manuscript and Klaus Dolag for his help with the technical details of the SPH code.

## REFERENCES

- Barnes L. A., Haehnelt M. G., 2009, MNRAS, 397, 511  
 Becker G. D., Sargent W. L. W., Rauch M., Simcoe R. A., 2006, ApJ, 640, 69  
 Becker G. D., Sargent W. L. W., Rauch M., Calverley A. P., 2011, ApJ, 735, 93  
 Becker G. D., Bolton J. S., Lidz A., 2015a, Publ. Astron. Soc. Aust., 32, e045  
 Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015b, MNRAS, 447, 3402  
 Bird S., Vogelsberger M., Haehnelt M., Sijacki D., Genel S., Torrey P., Springel V., Hernquist L., 2014, MNRAS, 445, 2313  
 Bird S., Haehnelt M., Neeleman M., Genel S., Vogelsberger M., Hernquist L., 2015, MNRAS, 447, 1834  
 Bird S., Garnett R., Ho S., 2017, MNRAS, 466, 2111  
 Boksenberg A., Sargent W. L. W., 2015, ApJS, 218, 7  
 Bolton J. S., Viel M., 2011, MNRAS, 414, 241  
 Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897  
 Carswell R. F., Webb J. K., 2014, Astrophysics Source Code Library, record ascl:1408.015  
 Cen R., Chisari N. E., 2011, ApJ, 731, 11  
 Chabrier G., 2003, PASP, 115, 763  
 Crighton N. H. M. et al., 2015, MNRAS, 452, 217  
 D’Odorico V., Calura F., Cristiani S., Viel M., 2010, MNRAS, 401, 2715  
 D’Odorico V. et al., 2013, MNRAS, 435, 1198  
 D’Odorico V. et al., 2016, MNRAS, 463, 2690  
 Díaz C. G., Ryan-Weber E. V., Cooke J., Koyama Y., Ouchi M., 2015, MNRAS, 448, 1240  
 Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, MNRAS, 401, 1670  
 Fan X., 2006, New Astron. Rev., 50, 665  
 Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137  
 Finlator K., Muñoz J. A., Oppenheimer B. D., Oh S. P., Özel F., Davé R., 2013, MNRAS, 436, 1818  
 Finlator K., Thompson R., Huang S., Davé R., Zackrisson E., Oppenheimer B. D., 2015, MNRAS, 447, 2526  
 Finlator K., Oppenheimer B. D., Davé R., Zackrisson E., Thompson R., Huang S., 2016, MNRAS, 459, 2299  
 Furlanetto S. R., Mesinger A., 2009, MNRAS, 394, 1667  
 García L. A., Tescari E., Ryan-Weber E. V., Wyithe J. S. B., 2017, MNRAS, 469, L53  
 Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633  
 Haardt F., Madau P., 2012, ApJ, 746, 125  
 Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27  
 Jiang L. et al., 2013, ApJ, 772, 99  
 Katsianis A., Tescari E., Wyithe J. S. B., 2015, MNRAS, 448, 3001  
 Katsianis A., Tescari E., Wyithe J. S. B., 2016, Publ. Astron. Soc. Aust., 33, e029  
 Katsianis A., Tescari E., Blanc G., Sargent M., 2017, MNRAS, 464, 4977  
 Keating L. C., Haehnelt M. G., Becker G. D., Bolton J. S., 2014, MNRAS, 438, 1820  
 Keating L. C., Puchwein E., Haehnelt M. G., Bird S., Bolton J. S., 2016, MNRAS, 461, 606  
 McQuinn M., 2015, ARA&A, 54, 313  
 Madau P., Dickinson M., 2014, ARA&A, 52, 415  
 Maio U., Tescari E., 2015, MNRAS, 453, 3798  
 Maio U., Dolag K., Ciardi B., Tornatore L., 2007, MNRAS, 379, 963

- Nagamine K., Springel V., Hernquist L., 2004, MNRAS, 348, 421  
 O’Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666  
 Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265  
 Oppenheimer B. D., Davé R., Finlator K., 2009, MNRAS, 396, 729  
 Pallottini A., Ferrara A., Gallerani S., Salvadori S., D’Odorico V., 2014, MNRAS, 440, 2498  
 Péroux C., McMahon R. G., Storrle-Lombardi L. J., Irwin M. J., 2003, MNRAS, 346, 1103  
 Pettini M., Madau P., Bolte M., Prochaska J. X., Ellison S. L., Fan X., 2003, ApJ, 594, 695  
 Planck Collaboration XIII, 2016, A&A, 594, A13  
 Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487  
 Pontzen A. et al., 2008, MNRAS, 390, 1349  
 Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543  
 Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123  
 Puchwein E., Springel V., 2013, MNRAS, 428, 2966  
 Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, MNRAS, 430, 2427  
 Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034  
 Ryan-Weber E. V., Pettini M., Madau P., 2006, MNRAS, 371, L78  
 Ryan-Weber E. V., Pettini M., Madau P., Zych B. J., 2009, MNRAS, 395, 1476  
 Schmidt M., 1959, ApJ, 129, 243  
 Simcoe R. A., 2006, ApJ, 653, 977  
 Simcoe R. A. et al., 2011, ApJ, 743, 21  
 Simcoe R. A., Sullivan P. W., Cooksey K. L., Kao M. M., Matejek M. S., Burgasser A. J., 2012, Nature, 492, 79  
 Songaila A., 2001, ApJ, 561, L153  
 Songaila A., 2005, AJ, 130, 1996  
 Springel V., 2005, MNRAS, 364, 1105  
 Springel V., Hernquist L., 2003, MNRAS, 339, 289  
 Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776  
 Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411  
 Tescari E., Viel M., D’Odorico V., Cristiani S., Calura F., Borgani S., Tornatore L., 2011, MNRAS, 411, 826  
 Tescari E., Katsianis A., Wyithe J. S. B., Dolag K., Tornatore L., Barai P., Viel M., Borgani S., 2014, MNRAS, 438, 3490  
 Theuns T., Leonard A., Efstathiou G., 1998, MNRAS, 297, L49  
 Thielemann F.-K. et al., 2003, Nucl. Phys. A, 718, 139  
 Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050  
 van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123, 305  
 Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574  
 Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D., 1986, ApJS, 61, 249  
 Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861  
 Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181  
 Zafar T., Péroux C., Popping A., Milliard B., Deharveng J.-M., Frank S., 2013, A&A, 556, A141  
 Zwaan M. A., Prochaska J. X., 2006, ApJ, 643, 675

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