# Absorption systems at $z\sim 2$ as a probe of the circumgalactic medium: a probabilistic approach

C. Mongardi,[1]★ M. Viel,[2,3,4]★ V. D'Odorico,[3]★ T.-S. Kim,[3,5] P. Barai,[3,6] G. Murante[2] and P. Monaco[1,3,4]

[1]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[2]*SISSA – International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste, Italy*
[3]*INAF– Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, I-34131 Trieste, Italy*
[4]*INFN/National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*
[5]*Department of Astronomy, University of Wisconsin, 475 N. Charter St., Madison, WI 53706, USA*
[6]*Instituto de Astronomia, Universidade de Sao Paulo (IAG-USP), Rua do Matão 1226, BR-05508-090 Sao Paulo, Brazil*

## ABSTRACT

We characterize the properties of the intergalactic medium (IGM) around a sample of galaxies extracted from state-of-the-art hydrodynamical simulations of structure formation in a cosmological volume of 25 Mpc comoving at $z\sim 2$. The simulations are based on two different subresolution schemes for star formation and supernova feedback: the Multi-Phase Particle Integrator (MUPPI) scheme and the Effective Model. We develop a quantitative and probabilistic analysis based on the apparent optical depth method of the properties of the absorbers as a function of impact parameter from their nearby galaxies: in such a way, we probe different environments from circumgalactic medium to low-density filaments. Absorbers' properties are then compared with a spectroscopic observational data set obtained from high-resolution quasar spectra. Our main focus is on the $N_{\rm C\,IV}$–$N_{\rm H\,I}$ relation around simulated galaxies: the results obtained with MUPPI and the Effective model are remarkably similar, with small differences only confined to regions at impact parameters $b = [1-3] \times r_{\rm vir}$. Using C IV as a tracer of the metallicity, we obtain evidence that the observed metal absorption systems have the highest probability to be confined in a region of 150–400kpc around galaxies. Near-filament environments have instead metallicities too low to be probed by present-day telescopes, but could be probed by future spectroscopical studies. Finally, we compute C IV covering fractions which are in agreement with observational data.

**Key words:** method: numerical – intergalactic medium – quasars: absorption lines.

## 1 INTRODUCTION

The matter that fills the space between galaxies, the so-called intergalactic medium (IGM), is distributed as a cosmic web of filaments and sheets: a component which has too low densities to be probed in emission, but can be observed through absorption lines in the spectra of luminous background sources. There are still significant uncertainties related to galaxy formation and evolution and to the modeling of the baryon's physical processes. In this context, the study of the IGM and of its evolution is of great relevance. First of all, the IGM is the reservoir of baryons from which galaxies form and it is the fuel necessary for sustaining star formation (SF), while at the same time it is being replenished with both newly accreted

intergalactic gas and chemically enriched materials from galaxies, carrying the imprints of galactic feedback. Therefore, intergalactic space provides a critical laboratory for studying the baryon cycle that regulates SF and galaxy growth.

In particular, IGM metal enrichment has gained a lot of interest after the second half of the 1990s (e.g. Nath & Trentham 1997; Murakami & Yamashita 1997; Davé et al. 1998; Gnedin 1998). In fact, it was originally thought that the gas infalling onto galaxies had a primordial composition, as heavy elements can be produced only inside stars and SF is not present in the low-density and high-temperature IGM. Before the era of high-resolution spectroscopy combined with large telescopes, in fact, quasars (QSO) absorption lines were classified into two categories: (1) metal-enriched high HI column density absorbers; and (2) metal-free low HI column density absorbers ($N_{\rm H\,I} < 10^{17}$ cm$^{-2}$). Thus, confirming the above-mentioned idea.

★ E-mail: astro386@hotmail.it (CM) viel@sissa.it (MV) valentina.dodorico@inaf.it (VD)

With the advent of powerful spectrographs, like HIRES (HIgh Resolution Spectrometer) at the Keck telescope and UVES (Ultraviolet and Visual Echelle Spectrograph) at the ESO-VLT (European Southern Observatory-Very Large Telescope), absorption lines of the triply ionized carbon doublet C IV ($\lambda\lambda$ 1548.204, 1550.778), the most common metal transition found in QSO spectra, have been observed at $z \sim 3$ associated with H I lines of the so-called Ly$\alpha$ forest with $N_{\rm H\,I} \geq 10^{14.5}$ cm$^{-2}$ (Cowie et al. 1995; Tytler et al. 1995; Songaila 1998).

Different scenarios have been proposed for the IGM metal-enrichment, such as an early-enrichment by Population III stars at very high redshift ($10 < z < 20$, e.g. Ostriker & Gnedin 1996; Haiman & Loeb 1997), dynamical removal by galaxy mergers (e.g. Gnedin & Ostriker 1997; Aguirre et al. 2001b), or galactic outflows/winds at lower redshift ($z \sim$ 2–3, e.g. Davé et al. 1998; Aguirre et al. 2001a; Schaye et al. 2003; Springel & Hernquist 2003; Murray, Quataert & Thompson 2005; Oppenheimer & Davé 2006).

In the first scenario, outflows from early protogalaxies drive the enrichment: metals can easily escape from the shallow potential wells, polluting large comoving pristine regions of the IGM. The enriched regions evolve following the Hubble flow likely resulting in a uniform distribution of metals far from galaxies at the studied redshift ($z \sim$ 2–3).

The last scenario predicts a metallicity–overdensity relation (Aguirre et al. 2001b; Davé et al. 1998; Oppenheimer & Davé 2006), resulting in a non-uniform metal pollution of the IGM. The dynamical removal of metal-enriched gas by galaxies mergers has attracted less interest, as it cannot account for observed present-day mass metallicity relation (Aguirre et al. 2001b).

Most probably, IGM metal pollution is not due to a single mechanism, although the late enrichment scenario could have a dominant effect, as it seems to have the greatest observational support. In fact, recent observations of local starbursts (e.g. Martin 1999, 2005; Rupke, Veilleux & Sanders 2005; Combes et al. 2006; Georgakakis et al. 2007; Lemaux et al. 2014; Hinojosa-Goñi, Muñoz-Tuñón & Méndez-Abreu 2016) and also the detection of strong outflows around almost all galaxies at high redshift (e.g. Heckman et al. 2000; Pettini et al. 2001; Frye et al. 2002; Shapley et al. 2003; Martin 2005) have provided new insights into the nature and impact of supernova (SN)-driven winds, which can result in very fast and energetic outflows. From an observational point of view, several studies have been carried out, in order to test which enrichment scenario has a dominant effect and to study IGM/galaxies interactions. Efforts have been concentrated on two main approaches: the investigation of the level of pollution in the low-density gas (e.g. Ellison et al. 2000; Schaye et al. 2003; Aracil et al. 2004; D'Odorico et al. 2016) and the study of galaxy/absorbers relations. The latter is based on the detection of metal absorption lines in QSO spectra correlated with the position and redshift of galaxies in the same field of view, both at high (Adelberger et al. 2005; Steidel et al. 2010; Crighton et al. 2011; Turner et al. 2014, 2015) and low redshifts (Prochaska et al. 2011; Tumlinson et al. 2011; Werk et al. 2013; Bordoloi et al. 2014; Liang & Chen 2014; Borthakur et al. 2015, 2016).

These studies, in general, allow only a statistical insight into the problem. The few most ambitious surveys for weak associated metal absorption features are not yet able to reach the shallow densities characteristic of the IGM, while the bulk of observations probes only the denser portions of the Universe, by detecting systems with $\log(N_{\rm C\,IV}/{\rm cm}^{-2}) \gtrsim 12$ associated with $\log(N_{\rm H\,I}/{\rm cm}^{-2}) \gtrsim 14.5$, not providing sufficient evidence to discriminate between the enrichment scenarios. On the other hand, galaxy/absorbers relations have

mainly investigated regions around galaxies limited to a few hundreds of proper kpc. Moreover, galaxy surveys are flux limited, introducing an unavoidable bias in galaxy/absorbers relations, as the faintest galaxies are missed, resulting in a sparse picture of the circumgalactic medium (CGM)/IGM and of the true origin of the absorbers. Numerical hydrodynamical simulations that incorporate the relevant physical processes have emerged as a powerful tool to investigate these issues (Theuns et al. 2002; Cen, Nagamine & Ostriker 2005; Tescari et al. 2009; Tornatore et al. 2010; Tescari et al. 2011; Barnes et al. 2011; Stinson et al. 2012; Dalla Vecchia & Schaye 2012; Barai et al. 2013; Oppenheimer & Schaye 2013; Crain et al. 2013; Hirschmann et al. 2013; Vogelsberger et al. 2014; Ford et al. 2014; Schaye et al. 2015; Davé et al. 2017).

In this work, we aim at characterizing the properties of the CGM/IGM around galaxies at high redshift, using hydrodynamical simulations compared with observational data taken from the literature. We focused on an analysis of a sample of simulated galaxies and the environment around them using two different simulations. The goal was to test different subgrid schemes, in order to better understand the chemical properties of the environment around galaxies and in particular to quantify the physical scales up to which observed state-of-the-art absorption systems can be found. This paper presents a probabilistic approach to the galaxy/IGM interplay that has the goal to provide the first step towards a characterization of the intimate relation between the two.

In this work, we define the CGM as the region surrounding the galactic halo up to a physical distance from galaxy centre of $\sim$300–400kpc, (which at $z \sim 2$ corresponds to few times the virial radius for a galaxy with a halo mass of $\sim 10^{10}$–$10^{11}$ M$_\odot$), while the IGM refers to larger distances and smaller overdensities.

This paper is organized as follows. Section 2 describes the set of used simulations. Section 3 reports a description of the observational samples used to compare with. In Section 4, we describe the selection of the sample of simulated galaxies (Appendix A contains the details for the extraction of mock spectra). In Section 5, the main results of this work are presented, while in Section 6, the summary of our main conclusion is reported. In Appendix B1, we described the method used to compute column densities for simulated absorption systems and we discuss the implications of using different methods in the simulation, when comparing with observational data, whose properties are obtained using Voigt Profile Fitting codes. In Appendix B2, we report the column density distribution functions (CDDFs) of HI and CIV absorption systems used in Appendix B1. In Appendix C, we report the numerous figures, whose discussion is reported in Section 5.3.

## 2 SIMULATIONS

Our simulations were performed using the GADGET-3 code (Springel 2005) which uses a TREEPM (particle mesh) gravity solver algorithm and the gas dynamics is computed with smooth particle hydrodynamics (SPH).

Two cosmological volumes of size $L_{\rm box} = 25$ Mpc comoving are evolved, starting from the same initial conditions, from $z = 199$ up to 0 with periodic boundary conditions. Each volume uses $N_{\rm part} = 2 \times 256^3$ particles in the initial condition, half of which are dark matter (DM) particles and half are gas particles. DM and gas particle masses have the following values: $m_{\rm DM} = 2.68 \times 10^7$ M$_\odot$ and $m_{\rm gas} = 5.36 \times 10^6$ M$_\odot$, respectively.

In both models, we consider energy feedback driven by SNe [while active galactic nucleus (AGN) feedback is not present]; adopt SF models with metal-dependent cooling using the prescriptions of

Wiersma et al. (2009); assume the gas is dust-free, optically thin and in photoionization equilibrium, heated by a uniform photoionizing background [cosmic microwave background plus the Haardt & Madau 2001 model for the ultraviolet (UV)/X-ray]; use the chemical enrichment and stellar evolution model by Tornatore et al. (2007), where each star particle is treated as a simple stellar population and the production of 11 chemical elements (H, He, C, Ca, O, N, Ne, Mg, S, Si, and Fe) is followed. Different stellar yields are considered: from Type Ia SN (Thielemann et al. 2003), Type II SN (Woosley & Weaver 1995), and asymptotic giant branch stars (van den Hoek & Groenewegen 1997), while stellar lifetimes are taken from Padovani & Matteucci (1993). A stellar initial mass function by Kroupa, Tout & Gilmore (1993) in the range [0.1–100] $M_{\odot}$ is used.

The two boxes differ in the baryonic interaction prescriptions governing SF and stellar/SN feedback, which are subresolution phenomena and, thus, they have to be modelled ad hoc. One simulation is described in Barai et al. (2015), from where we used the run E25cw. The other one is very similar to the run M25std from Barai et al. (2015), but with slightly different stellar yields and feedback parameters: $f_{b, out} = 0.2$, $f_{b, kin} = 0.5$, and $P_{kin} = 0.02$. (The parameters $f_{b, out}$, $f_{b, kin}$, and $P_{kin}$ will be described in the next subsection.) Two different subresolution models were adopted: the MUPPI model (MUlti-Phase Particle Integrator; Murante et al. 2010, 2015) in run M25std, and the Effective Model (Springel & Hernquist 2003) in run E25cw, whose main features are described in the following subsections. We refer the reader to the above-mentioned papers for further details on the implementation of the two models.

A flat $\Lambda$ cold dark matter model is used with the following parameters: $\Omega_m = 0.24$, $\Omega_{\Lambda} = 0.76$, $\Omega_b = 0.04$, $H_0 = 72$ km s$^{-1}$ Mpc$^{-1}$, and $\sigma_8 = 0.83$.

## 2.1 MUPPI model

In the MUPPI model, whenever a gas particle's density is higher than a threshold value $n_{thr} = 0.01$ cm$^{-3}$ and its temperature is below the threshold $T_{thr} = 5 \times 10^4$ K, it enters a multiphase regime. The multiphase gas particle is composed of a hot and a cold phase in thermal pressure equilibrium ($n_c T_c = n_h T_h$), plus a virtual stellar component. The temperature of the cold phase is kept fixed at $T_c = 300$ K, while the hot phase $T_h$ is computed from the particle's entropy.

A fraction of the cold gas, $f_{mol}$, is considered to be in the molecular phase. It is the reservoir of material available for SF. It is computed following the observed relation by Blitz & Rosolowsky (2006) between the ratio of molecular to atomic gas surface densities and the external pressure exerted on molecular clouds. The external pressure is the mid-plane pressure of a thin disc composed of gas and stars in vertical hydrostatic equilibrium. In the MUPPI model, the hydrodynamical pressure of gas particles is used in place of the external pressure. This enables us to derive the following simple relation for computing $f_{mol}$:

$$f_{mol} = \frac{1}{1 + P_0/P}. \qquad (1)$$

Here, $P$ is the pressure of the gas particle and $P_0$ is the pressure at which half of the cold gas is molecular and is set to the value $P_0/k_B = 35000$ K cm$^{-3}$ (Blitz & Rosolowsky 2006).

The SF rate is directly proportional to $f_{mol}$ and is given by this equation:

$$\dot{M}_{SF} = f_{\star} \frac{f_{mol} M_c}{t_{dyn}}. \qquad (2)$$

Here, $f_{\star}$ is the SF efficiency, $M_c$ is the gas mass in the cold phase, and $t_{dyn}$ is the dynamical time, $t_{dyn} = \sqrt{3\pi/32 G \rho_c}$, with $\rho_c$ being the cold phase density.

It is important to highlight that in this model, SF is dependent on disc pressure. Thus, no Schmidt–Kennicutt relation is imposed in the MUPPI model. Rather, as demonstrated in Monaco et al. (2012), the Schmidt–Kennicutt relation is naturally recovered from the model.

Mass and energy exchanges between the three gas phases (hot, cold, and stellar) are described by a set of equations. When a gas particle enters the multiphase, all the particle's mass is assigned to the hot phase. Then, matter flows among the three phases in the following way: cooling deposits hot gas into the cold component; evaporation, happening under the action of hot SN bubbles, brings cold gas back to the hot phase; SF deposits mass from the cold phase into the stellar component; and mass restoration from dying stars moves mass from the stellar component back to the hot phase.

A star particle is produced following the stochastic SF algorithm of Springel & Hernquist (2003). A multiphase gas particle undergoing SF is turned into a collisionless star particle whenever a random number drawn uniformly from the interval [0,1] falls below the probability

$$P = \frac{M_p}{m_{\star}} \left[ 1 - \exp\left( -\frac{\Delta M_{\star}}{M_p} \right) \right]. \qquad (3)$$

Here, $M_p$ is the gas particle mass (hot mass + cold mass + stellar mass), and $\Delta M_{\star}$ is the cold gas mass transformed to stellar mass in a single SPH time-step. The mass of the star particle, $m_{\star}$, is defined as $m_{\star} = M_p/N_g$ with $N_g$ being the number of generations of stars per gas particle (with $N_g$ set equal to 4, $m_{\star} = 1.34\,^6 \times 10$ M$_{\odot}$).

Energy feedback from SN is transferred to gas particles both in the form of thermal and kinetic energy. Thermal energy is given to the hot phase, which is lost by cooling and acquired from SN explosions. The thermal energy available for feedback from a single gas particle and returned to the hot phase is a fraction $f_{fb, out}$ of the energy of a single SN. The model assumes to have one SN event per M$_{\star}$ = 120 M$_{\odot}$. A fraction of this energy is given to the local hot phase to sustain the high temperature of the particle itself. The remaining energy is redistributed to the hot phase of neighbouring gas particles within its SPH smoothing length and lying in a bicone of aperture $\theta = 60°$. The bicone axis is aligned along the least resistance direction of the gas density. This is to simulate the explosion of SN bubbles.

Kinetic feedback is implemented in the following manner. When a particle exits from the multiphase regime, it is assigned a probability $P_{kin}$ to become a wind particle. Then, it can receive kicks from neighbouring gas particles for a given time $t_{wind}$. The kinetic energy available for feedback from a single gas particle is a fraction $f_{b, kin}$ of the energy of a single SN. It is distributed to neighbouring wind particles in a cone within the gas particle's SPH smoothing length in a similar manner as for the thermal feedback. For each wind particle, the total kinetic energy available from all neighbouring kicking gas particles is first computed, as described above. Then, the wind particle's kinetic energy is increased by this total amount. The velocity vector of the wind kick is oriented by energy weighting among all the kicking particles.

Unlike other feedback prescriptions in the literature (Oppenheimer & Davé 2008; Schaye et al. 2010; Springel & Hernquist 2003), the MUPPI model depends only on the local properties of the gas. The mass-loading factor or the velocity of the wind is not input parameters of the feedback model. However, these quantities can be estimated empirically as described in Murante et al. (2015,

estimated mass load factor of ∼1.5 and estimated average wind velocity of ∼600 km s$^{-1}$).

A gas particle stays in the multi-phase regime until its density reaches a value equal to or lower than 1/5 of the density threshold $n_{thr}$. If the energy feedback is incapable to sustain the hot phase resulting in hot phase temperatures that stay lower than $10^5$ K, the particle is forced to escape from the multiphase regime.

The MUPPI model is able to produce realistic disc galaxies even at relatively low resolution such as those used in this paper, as shown in Murante et al (2015). In Goz et al. (2017, figs 1 and 2), a summary of the main galaxy properties found in the MUPPI box used here is given . In particular, their fig. 1 shows that the box contains a fair number of galaxies with relatively small dynamical values of the bulge-over-total stellar mass ratio. In Barai et al. (2015), the SF rate density of a number of different SF and feedback model is shown (models used here are M25std and E25cw, as said before). The effective model has a larger SF at high redshifts and a lower one at low ones. In the same paper, an analysis of the outflows shows that MUPPI is more efficient than the effective model in expelling the gas at high redshift, leaving it available for low-redshift disc formation when it falls back, while the effective model converts it in stars and produces less discs.

### 2.2 Effective model

In the Effective model, a multiphase gas particle is composed of a hot and a cold phase in thermal pressure equilibrium. Gas particles enter a multiphase regime whenever their density is higher than a threshold value $\rho_{SF} = 0.13$ cm$^{-3}$. This threshold is a SF density threshold, as SF prescription is not based on the Blitz & Rosolowsky (2006) relation. It depends only on the cold gas mass, which is directly converted into stars on a characteristic time-scale, given by the formula:

$$\dot{M}_{SF} = \frac{M_c}{t_\star} \tag{4}$$

where $t_\star = t_{\star,0} \sqrt{\rho/\rho_{SF}}$. A value of $t_{\star,0} = 2.1$ Gyr was chosen by Springel & Hernquist (2003) in order to fit the Kennicutt relation.

Mass and energy exchanges are regulated by the same physical processes described in the MUPPI section. Star particles are generated from gas particles using the stochastic scheme introduced by Katz, Weinberg & Hernquist (1996).

Energy feedback is given both in the form of thermal and kinetic energy. Differently from the MUPPI model, the two forms of feedback do not consider any distribution of the energy to neighbouring particles in a cone within the SPH smoothing length simulating the blowout of an SN bubble. Thermal feedback in the form of thermal heating and cloud evaporation is implemented.

Kinetic feedback uses the prescriptions of the *energy-driven wind* scenario with a constant wind velocity. The wind mass-loss rate is given by this formula:

$$\dot{M}_w = \eta \dot{M}_{SF} \tag{5}$$

where $\eta$ is the mass-loading factor, which is set to $\eta = 2$. A fixed fraction $\chi$ of the SN energy is converted into wind kinetic energy:

$$\frac{1}{2} \dot{M}_w v_w^2 = \chi \epsilon_{SN} \dot{M}_{SF}. \tag{6}$$

Here, $v_w$ is the wind velocity and $\epsilon_{SN}$ is the average energy released by SN for each M$_\odot$. So we have:

$$v_w = \sqrt{\frac{2\chi \epsilon_{SN}}{\eta}} \tag{7}$$

In our run, we use a fixed value $v_w = 350$ km s$^{-1}$. Gas particles are given wind kick using a probabilistic approach (see equation 10 in Barai et al. 2013 for details). Their velocity $v$ is incremented according to:

$$v' = v + v_w \hat{n} \tag{8}$$

where $\hat{n}$ is the direction of the wind, preferentially chosen along the rotation axis of spinning objects.

## 3 OBSERVATIONAL DATA SAMPLE

Here, we report a brief description of the observational data used. The comparison with simulated C IV and Si IV absorption systems has been made using the observational sample by Kim et al. (2016, hereafter K16), which are described in Section 3.1, while O VI is compared with the observational sample by Muzahid et al. (2012), described in Section 3.2. Constructed C IV covering fractions are compared with those derived by Landoni et al. (2016), Prochaska, Lau & Hennawi (2014), and Rubin et al. (2015), which are described in Section 3.3.

### 3.1 C IV and Si IV data sample

Data for the comparison with the simulated C IV (and also Si IV) absorption are taken from K16. This data set is chosen because it represents a homogeneous and statistically significant sample of high-resolution absorption lines. The sample consists of 23 QSOs in the redshift range $2 \lesssim z \lesssim 3.5$, which were observed with the VLT/UVES (21 spectra) and Keck/HIRES (2 spectra) instruments. Spectra were first presented in the following papers: Kim et al. (2004, 2007, 2013) and Boksenberg & Sargent (2015).

Spectral resolution is $R \sim 45000$ ($\sim 6.7$ km s$^{-1}$), while typical signal-to-noise ratios (S/N) per pixel are of the order of 30–50 for the Ly$\alpha$ forest and ∼100 for the CIV region. H I and C IV absorption lines detected in all the QSO spectra are in the redshift range $1.5 \lesssim z \lesssim 3.3$.

Data points, used for comparison with simulations in Section 5, are integrated column densities representing H I and C IV *systems*, as defined in K16. First, the entire normalized spectra were inspected to identify the CIV doublets. Then, the absorption lines were fitted with Voigt profiles using the VPFIT code[1] (Carswell & Webb 2014), decomposing the complex absorption features into individual, single component by minimizing the normlized X$^2$ of ∼1.3. The output of the fit consists of the physical parameters, such as the column density in cm$^{-2}$, the redshift and the Doppler parameter in km s$^{-1}$, of the single fitted components.

Then, the C IV systems are defined as all the individually fitted C IV components within the fiducial ±150 km s$^{-1}$ interval centred at the CIV flux minimum of a single or multiple C IV feature. The integrated column density is then calculated by adding up all the column densities of the C IV components within ±150 km s$^{-2}$.

If the C IV absorption extends over ±150 km s$^{-1}$ or when two different absorption profiles are found to lie within ±150 km s$^{-1}$ but they are spread beyond the ±150 km s$^{-1}$, then the velocity range is extended by a step of 100 km s$^{-1}$ in order to include all the C IV absorption.

The *associated* H I system is obtained by applying the same velocity range as the C IV system. While the fiducial integration velocity range of ±150 km s$^{-1}$ could be considered arbitrary, the

---

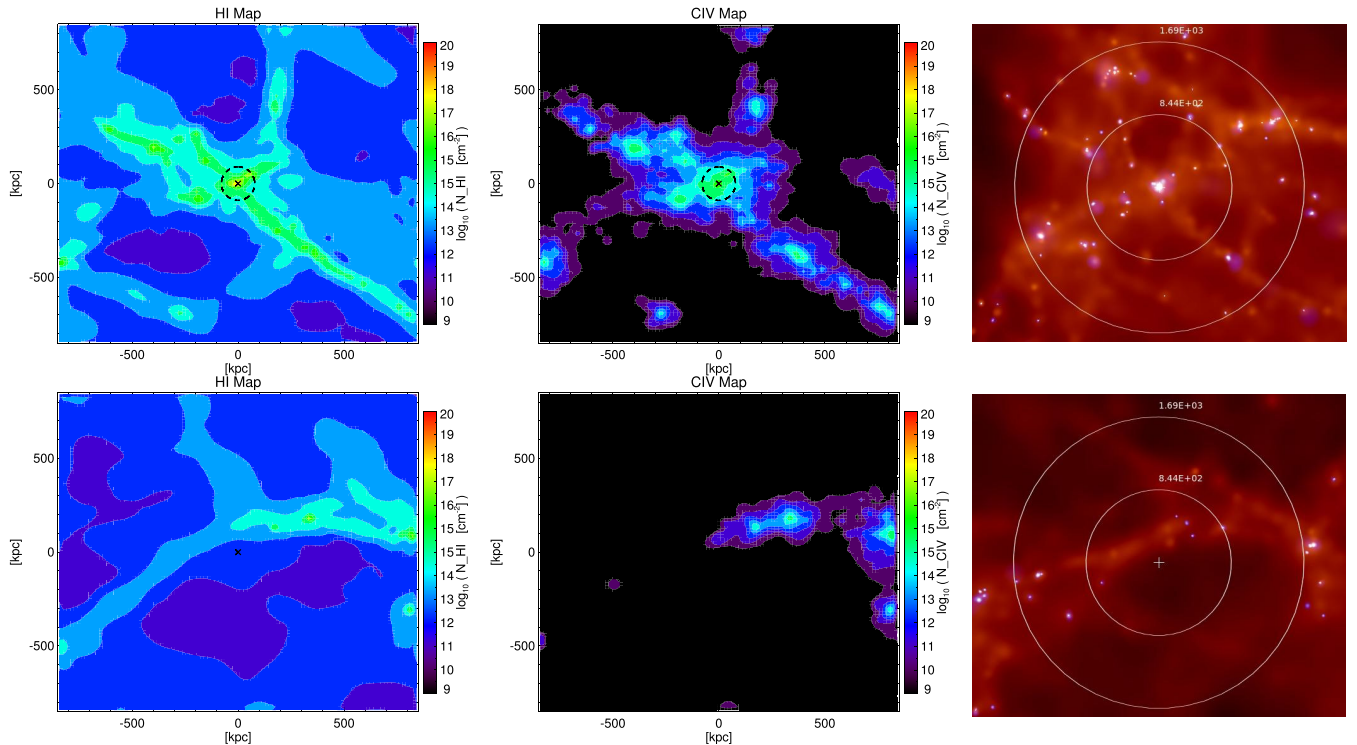[1]http://www.ast.cam.ac.uk/rfc/vpfit.html

**Figure 1.** Top: projected maps of the H I (left) and C IV (centre) column density distribution around a given galaxy of the investigated sample (distances are in proper kpc). Column densities have been computed by integrating the density profile of these sub-boxes along a perpendicular slice of width $\pm 300$ kpc around galaxy's position. Image of the chosen galaxy (right) obtained from GADGETVIEWER (a software for visualization of N-body simulations. The image is colour coded according to particle properties: in red, it is represented as SPH smoothed density plot of gas particles distribution, and in blue, the same but for star particles): the chosen galaxy is in the centre of the image; the external circle has a radius of $1.6 \times 10^3$ comoving kpc, while the internal circle has a radius of $8.4 \times 10^2$ comoving kpc. Bottom: same as the top panel, but for a near-filament environment. The white cross here represents the visually chosen 'centre of the object'.

choice of the velocity range does not affect the observed integrated $N_{\rm H\,I}$ versus $N_{\rm C\,IV}$ relation (see beginning of Secton 5.1 for a brief explanation of this relation) as long as it is larger than $100\,{\rm km\,s^{-1}}$ (K16). With this working definition, the C IV centroid does not necessarily coincide with the H I centroid and there is no evidence that C IV and H I systems are physically connected. This definition is thus more related to volume-averaged quantities, commonly used in numerical simulations.

Finally, it is demonstrated in K16, that the $N_{\rm C\,IV}$ versus $N_{\rm H\,I}$ relation converges when considering velocity integration ranges greater than $70\,{\rm km\,s^{-1}}$.

In order to be consistent with data, we used the same velocity integration range in the analysis of the simulated absorption systems.

We compare the data only with one snapshot of the simulations at $z \sim 1.94$, since we verified that there is no significant evolution in the $N_{\rm C\,IV}$ versus $N_{\rm H\,I}$ relation in the redshift range covered by the K16 sample (see also the discussion in K16).

### 3.2 O VI data sample

Comparison with simulated O VI absorption systems is performed with the observational sample by Muzahid et al. (2012). The sample is formed by 18 QSOs in the redshift range $2.1 \lesssim z \lesssim 3.3$, observed with the VLT/UVES. Typical S/N$\sim$30–40 and 60–70 are achieved in the wavelength range of 3300 and 5500 Å. Spectral resolution is $R \sim 45000$ (FWHM $\sim 6.6\,{\rm km\,s^{-1}}$) over the entire wavelength range.

H I and O VI absorption lines detected in all the QSO spectra are in the redshift range $1.9 \lesssim z \lesssim 2.9$.

They define an absorption system by grouping together all the lines whose separation from the nearest neighbour is less than a linking length $v_{\rm link} = 100\,{\rm km\,s^{-1}}$, following the prescriptions by Scannapieco et al. (2006).

### 3.3 C IV covering fraction data sample

For the comparison with the constructed covering fractions (see Section 5.4), we used the observational samples of Rubin et al. (2015), Landoni et al. (2016), and Prochaska et al. (2014). Landoni et al. (2016) and Prochaska et al. (2014) use QSO projected pairs, that is they use the spectrum of a background QSO to study the gaseous envelope of a foreground QSO host galaxy.

The sample of Prochaska et al. (2014) consists of 427 QSO pairs, whose spectra were taken mainly from the BOSS survey or obtained with the Keck/LRIS instrument, with impact parameters in the range 39 kpc $< b <$ 1 Mpc and redshifts $1.8 < z < 3.5$. The 18 QSO pairs by Landoni et al. (2016) were taken from the SDSS survey and have impact parameters smaller than $b <$ 200 kpc and redshifts $2.0 < z < 2.8$. The main difference between these two samples is that Prochaska et al. (2014) perform a boxcar integration in a $3000\,{\rm km\,s^{-1}}$ velocity window centred on the redshift of the foreground QSO, while Landoni et al. (2016) integrate on a window of $1200\,{\rm km\,s^{-1}}$.

Rubin et al. (2015) use, instead, QSO pairs to study the diffuse gas in the CGM of 40 damped Lyman-$\alpha$ (DLAs, Wolfe et al. 1986) systems, that is the primary sightline probes an intervening DLA system in the redshift range $1.6 < z < 3.6$, while the secondary sightline is used to look for a C IV absorption at the same redshift of the DLA, up to a distance of the order $b < 300$ kpc.

## 4 SIMULATED DATA SAMPLE

In order to interpret the observed data, we have built a set of simulated data, which we have processed in a way as close as possible to that adopted for observations. First of all, we have built simulated mock spectra and then we have constructed absorption systems and computed thier column densities. We give details of those operations in Appendices A and B1 and the following subsection.

### 4.1 Sample selection

The extraction of line of sight (LOS) physical quantities follows the standard procedure of SPH schemes and is summarized in Appendix A.

In order to investigate the IGM/CGM of galaxies, we selected from the snapshot at $z = 1.94$ of both cosmological runs two samples of 20 galaxies with a total halo mass $M_h$ in the range $[\sim 10^{10} – 10^{12}]$ $M_\odot$, identified with the SUBFIND algorithm (Springel et al. 2001) and stellar mass greater than $M_* \gtrsim 2 \times 10^9 M_\odot$.

The selection criteria were the following: the galaxies had not to be subject to major mergers and to be the main halo of their friends-of-friends (FOF) group, so not a substructure, in order to compare similar conditions at a given distance from galaxy centre between the different galaxies of the sample. The first condition puts an upper limit on the chosen halo mass, as only six objects with $M_h > 10^{12}$ $M_\odot$ are present in the two runs, but they are actually all formed by clusters of merging galaxies. The lower limit on the halo mass, instead, is related to the number of particles that sample the galaxy and its various components. We fixed a minimum number of $\sim 100$ gas particles, in order to consider the object reliable.

In Table 1, we report some properties of the two galaxy samples. Virial radii are defined for each FOF halo as the radii of a sphere centred on the main halo of the FOF group and which contains an overdensity of 200 the critical density, so they are calculated with this formula:

$$R_{vir} = \left( \frac{M_h}{\frac{4}{3}\pi \cdot 200 \cdot \rho_{crit}} \right)^{1/3} \tag{9}$$

with $\rho_{crit} = 3 \, H(z)^2 / 8\pi \, G$. We also visually selected from the two boxes, using the software GADGETVIEWER,[2] 10 near-filaments environments, in order to investigate the low-density IGM near galaxies.

In the first row of Fig. 1, a galaxy from the MUPPI simulation is shown, while in the second row, we plot an example of a near-filament environment. In the left-hand panels, the H I column density maps around these two objects are reported; in the centre, the C IV column density maps; and on the right, we show the images of the two environments with GADGETVIEWER.

For near-filament environments, the 'centre' is represented by the position of the white cross, as shown in the bottom left panel of Fig. 1.

[2] http://astro.dur.ac.uk/jch/gadgetviewer/index.html

**Table 1.** Properties of the sample galaxies of the MUPPI and the Effective Model Boxes. In the first column of the two tables, there is the object ID; in the second column, the halo mass of the galaxy; in the third column, the stellar mass of the galaxy; and in the fourth column, the virial radius of the galaxy in physical units.

| Obj. | $M_h$ ($M_\odot$) | $M_\star$ ($M_\odot$) | $R_{vir}$ (phys. kpc) |
|---|---|---|---|
| | **MUPPI model** | | |
| 1 | $6.87 \times 10^{11}$ | $2.37 \times 10^{10}$ | 94 |
| 2 | $4.03 \times 10^{11}$ | $1.13 \times 10^{10}$ | 78 |
| 3 | $4.02 \times 10^{11}$ | $8.31 \times 10^{9}$ | 78 |
| 4 | $3.63 \times 10^{11}$ | $1.11 \times 10^{10}$ | 76 |
| 5 | $2.74 \times 10^{11}$ | $1.24 \times 10^{9}$ | 69 |
| 6 | $2.60 \times 10^{11}$ | $6.87 \times 10^{9}$ | 68 |
| 7 | $2.08 \times 10^{11}$ | $6.83 \times 10^{9}$ | 63 |
| 8 | $1.66 \times 10^{11}$ | $5.92 \times 10^{9}$ | 58 |
| 9 | $1.43 \times 10^{11}$ | $5.04 \times 10^{9}$ | 55 |
| 10 | $1.06 \times 10^{11}$ | $3.38 \times 10^{9}$ | 50 |
| 11 | $9.49 \times 10^{10}$ | $2.54 \times 10^{9}$ | 48 |
| 12 | $9.22 \times 10^{10}$ | $1.95 \times 10^{9}$ | 48 |
| 13 | $9.12 \times 10^{10}$ | $2.84 \times 10^{9}$ | 48 |
| 14 | $8.91 \times 10^{10}$ | $2.53 \times 10^{9}$ | 47 |
| 15 | $8.21 \times 10^{10}$ | $2.06 \times 10^{9}$ | 46 |
| 16 | $8.14 \times 10^{10}$ | $2.13 \times 10^{9}$ | 46 |
| 17 | $8.10 \times 10^{10}$ | $1.75 \times 10^{9}$ | 46 |
| 18 | $8.05 \times 10^{10}$ | $2.51 \times 10^{9}$ | 46 |
| 19 | $7.79 \times 10^{10}$ | $3.06 \times 10^{9}$ | 45 |
| 20 | $7.36 \times 10^{10}$ | $2.09 \times 10^{9}$ | 44 |
| | **Effective model** | | |
| 1 | $5.04 \times 10^{11}$ | $2.18 \times 10^{10}$ | 84 |
| 2 | $4.91 \times 10^{11}$ | $3.07 \times 10^{10}$ | 84 |
| 3 | $4.90 \times 10^{11}$ | $3.73 \times 10^{10}$ | 84 |
| 4 | $4.11 \times 10^{11}$ | $2.23 \times 10^{10}$ | 81 |
| 5 | $3.66 \times 10^{11}$ | $2.70 \times 10^{10}$ | 74 |
| 6 | $3.04 \times 10^{11}$ | $1.66 \times 10^{10}$ | 71 |
| 7 | $3.00 \times 10^{11}$ | $1.31 \times 10^{10}$ | 71 |
| 8 | $2.39 \times 10^{11}$ | $1.16 \times 10^{10}$ | 66 |
| 9 | $2.32 \times 10^{11}$ | $7.50 \times 10^{9}$ | 65 |
| 10 | $2.12 \times 10^{11}$ | $8.22 \times 10^{9}$ | 63 |
| 11 | $9.91 \times 10^{10}$ | $2.81 \times 10^{9}$ | 49 |
| 12 | $9.66 \times 10^{10}$ | $3.07 \times 10^{9}$ | 48 |
| 13 | $9.40 \times 10^{10}$ | $3.48 \times 10^{9}$ | 48 |
| 14 | $9.01 \times 10^{10}$ | $4.99 \times 10^{9}$ | 47 |
| 15 | $8.95 \times 10^{10}$ | $2.21 \times 10^{9}$ | 47 |
| 16 | $8.93 \times 10^{10}$ | $3.64 \times 10^{9}$ | 47 |
| 17 | $8.25 \cdot 10^{10}$ | $2.21 \times 10^{9}$ | 46 |
| 18 | $7.78 \times 10^{10}$ | $2.30 \times 10^{9}$ | 45 |
| 19 | $7.77 \times 10^{10}$ | $3.48 \times 10^{9}$ | 45 |
| 20 | $3.73 \times 10^{10}$ | $3.64 \times 10^{9}$ | 35 |

## 5 RESULTS

### 5.1 The $N_{C\,IV}$ versus $N_{H\,I}$ relation: piercing around objects

The $N_{C\,IV}$ versus $N_{H\,I}$ relation, constructed here in order to compare with data, is the observational equivalent of the theoretical overdensity–metallicity relation, as the overdensity is related to $N_{H\,I}$ and the metallicity can be traced by the presence of C IV. The C IV doublet, in fact, falls in a region redward of the Ly$\alpha$ forest, mostly free from line blendings, and observable from the ground in the visible band (3000–10000 Å) at $0.9 \lesssim z \lesssim 5.4$. Its ubiquitous presence in QSO spectra also indicates that it traces a gas phase, which is characteristic of many astrophysical gas environments.

We pierced through the cosmological boxes random lines of sight around the 40 galaxies (20 per model) with impact parameters less than 800 kpc, in order to characterize the environment around them.

As the LOS is as long as the box side, we first identified the position of the centre of the galaxy along it. We, then, selected a region along the LOS $\pm 150 \, \text{km s}^{-1}$ from this position and we considered the optical depth profile both for H I and C IV in this region (see Appendix B1). For each pixel of this region, we converted the value of the optical depth into a column density adopting the apparent optical depth (AOD) method (see Appendix B1, equation B1, Savage & Sembach 1991) and then we integrated in the considered velocity range. The same procedure is used for constructing O VI and Si IV systems, discussed in next sections.

In order to test the reliability of our simulation, we have computed the column density distribution functions (CDDFs) for a general sample of C IV and H I lines. The CDDF is defined as the number of absorbers in column density and redshift bins. The simulated CDDFs were then compared with the observed ones computed with the sample of Kim et al. (2013). The comparison is reported in Figs B2 and B3 for the MUPPI and the Effective models, respectively. In particular, the MUPPI model, which is our reference model, reproduced in a reasonable way both the C IV and H I CDDFs in the range of interest.

In Figs 2 and 3, we show the probability distribution functions (PDFs) of the $N_{\text{C IV}}$ versus $N_{\text{H I}}$ relation for the MUPPI and the Effective Models, respectively, compared to the observational sample by K16 (white triangles with error bars). The horizontal dashed lines in all plots represents the observational detection limit $\log N_{\text{C IV}} \simeq 11.8$. The simulated absorption systems are divided according to the distance of the corresponding LOS from the considered galaxy. From top-left to bottom-right, in each plot we show lines of sight with impact parameters $b$ between $[0–1] \, r_{\text{vir}}$, $[1–3] \, r_{\text{vir}}$, $[3–5] \, r_{\text{vir}}$, and $b > 5 \, r_{\text{vir}}$, where $r_{\text{vir}}$ is the virial radius of the considered galaxy reported in Table 1.

Going to larger distances, we can see that the probability distribution gradually shifts to lower values of H I and C IV column densities. Inside the virial radius, all the simulated lines of sight are above the observational limit. In this plot, the bulk of simulated points occupy a slightly different region than the data of K16: it looks like there is a shift toward larger C IV column densities with respect to observational data. This could be due to either local radiation which is neglected in our models or to a wrong metallicity. In fact, we are relying on an ultraviolet background (UVB) which is not accurate, as we are assuming that each gas particle is photoionized by Haardt & Madau UVB, neglecting that near galaxies, there is also the contribution of the ionizing photons from the galaxy itself. It is however reasonable to assume that such effect is not large at least for HI absorbers at $\log(N_{\text{H I}}/\text{cm}^{-2}) < 17$ as suggested by Rahmati et al. (2013) and for CIV could also be a minor effect due to the fact that radiation from galaxies falls sharply above 1 Ryd. A more likely explanation is thus that the simulation produces too many metals, due to inefficient feedback, that in turn causes too much SF.

To investigate this last issue, we constructed the mass–metallicity relation for all the galaxies of both simulations with stellar mass $M_* > 2 \times 10^9 \, \text{M}_\odot$. For simplicity, we show the results only for the MUPPI simulation. The Effective model provides similar results, as it shares the same chemical model and it does not include AGN feedback as in MUPPI. For each gas particle inside one-tenth of the virial radius of a galaxy, we computed the abundance O/H. For each galaxy, we then took the gas particle's mass-weighted average value of O/H. Our result are shown in Fig. 4. The simulation does actually produce too many metals with respect to the observed relation by

Kacprzak et al. (2016) and this could explain the disagreement in the C IV data.

Goz et al. (2017) analysed the mass–metallicity relation of the MUPPI simulation at $z = 0$. They compare the gas metallicity of their sample of MUPPI galaxies, taken from the same box we used, as a function of stellar mass, with the observational relation by Tremonti et al. (2004). They recover the same trend as the observed relation, but with a global offset of $\sim 0.1$ dex, extending up to 0.2 dex in the low-mass end of the relation. They find that massive galaxies show a lack of proper quenching and small galaxies are too massive, passive, metallic and with low atomic gas content. Barai et al. (2015), using the same simulations, find an excess of SF rate density, in particular at high redshift. Both papers claim that these results could be due to the lack of AGN feedback in the simulations, capable of quenching the cooling and SF. In the case of our results, although AGNs are not dominant in the population of galaxies we have considered (the median halo mass is $\sim 10^{11} \, \text{M}_\odot$), we believe the cumulative effect of their feedback during galaxy evolution could alleviate the excess of metallicity we measure at $z \simeq 1.9$, as its contribution at very high redshift prevents the formation at lower redshift of too many small galaxies, formed by too metal-enriched gas particles.

In Figs 2 and 3, we note that there are non-negligible probabilities of finding simulated systems with large values of $N_{\text{H I}}$ and $N_{\text{C IV}}$, even at large distances from the chosen galaxy. This is likely a consequence of the fact that galaxies are not isolated systems.

We investigated this issue by considering all the lines of sight with $b > 5 \, r_{\text{vir}}$ and with absorption systems with $14.0 \leq \log(N_{\text{C IV}}/\text{cm}^{-2}) \leq 18.0$.

We searched for substructures near the lines of sight and we computed the minimum distance from an LOS to a substructure. The distribution of the distances is reported in Fig. 5. The blue solid histogram considers all the DM substructures identified by the SUBFIND search algorithm with no distinction in the mass. That includes also the smallest structures, simply formed by DM and gas clumps without SF. The green dashed histogram has a cut in the stellar mass of the substructure, as discussed in Section 4.1. The blue solid histogram has a median value and a $1\sigma$ error of $19^{+29}_{-11}$ proper kpc, while the green dashed one has a much broader distribution, with a median value and $1\sigma$ error of $225^{+200}_{-140}$ proper kpc. This result suggests that those lines of sight have a higher probability to be related to another nearby substructure. They could pass through a small DM and gas clump or they could be near the halo of another galaxy, closer than the chosen one. In fact, even if the median value of 225 kpc is not a small distance, it is closer than the 5 $r_{\text{vir}}$ of the chosen galaxy.

For completeness, the median values with the $1\sigma$ dispersion of the distance distributions in the ranges $b < 1 \, r_{\text{vir}}$, $1 \, r_{\text{vir}} < b < 3 \, r_{\text{vir}}$, and $3 \, r_{\text{vir}} < b < 5 \, r_{\text{vir}}$ are $15^{+12}_{-8}$, $19^{+18}_{-11}$, and $19^{+21}_{-11}$ kpc, respectively. The total mass of the substructure is in the range $M_{\text{h}} \sim 10^8–10^{12}$ $\text{M}_\odot$.

As imaging surveys are flux-limited, faint galaxies can be missed. For this reason, we can state from Figs 5, 2, and 3 that finding a strong absorption system at a large distance from a galaxy is possible and the two can be related to each other. However, it is more probable that the absorption system is related to a smaller galaxy, which is not detected in present observations.

The plots in Figs 2 and 3 constitute the main result of this work: they give the probability to find an absorption system with a given H I and C IV column density at a certain distance from a galaxy. We can state that the observational data have the highest probability to originate in a region within three virial radii from galaxies. In
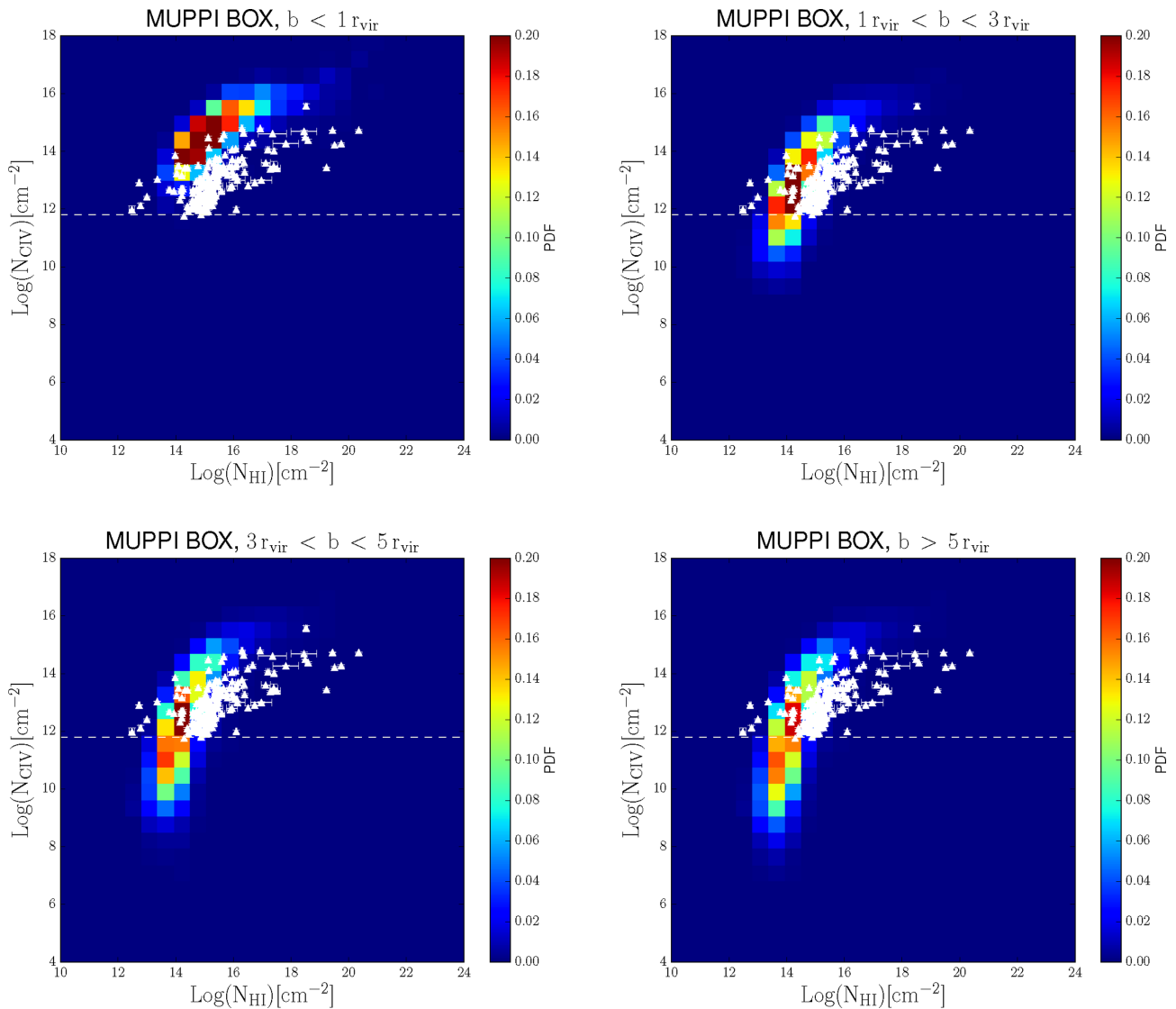
**Figure 2.** PDF of the simulated C IV systems. White triangles: observational data from K16; and the same data are reported in each panel. Dashed horizontal line: observational detection limit.

the regime beyond 3 $r_{vir}$, the most probable systems are below our observational detection limit.

This result is strengthened by the fact that we do not recover any strong difference between the MUPPI and the Effective models. In Fig. 6, we show the comparison between the two models (MUPPI in blue dashed and the Effective model in red solid). The peaks of the probabilities are always coincident even though the Effective model shows a more extended tail for low values of C IV column densities for all the radii above 1 $r_{vir}$. This could be due to the different mechanism of feedback of the Effective model, which is not capable to spread metals with the same efficiency.

We repeated the same procedure for near-filament environments, by piercing through the cosmological box random lines of sight around the chosen centre of near-filament environments with impact parameter less than 800 kpc. Fig. 7 shows the PDFs of all lines of sight around near-filament environments for both simulations. Near-filament environments have the highest probability to have values of the C IV and H I column densities smaller than what it is observed.

The tail of the PDF at higher values of column densities and slightly in correspondence with the observational sample could be due to the presence of haloes near the chosen 'centre' of a near-filament environment.

### 5.2 The $N_{CIV}$ versus $N_{HI}$ relation: median values in concentric regions

We performed another analysis providing a more statistical picture of the distribution of metals in the CGM/IGM. We took the lines of sight of 5.1 and for each object, we divided them according to their distance from the centre of the object: going from 0 to 800 kpc, we divided the spatial region considered in radial bins of size ∼ 10 kpc and we grouped lines of sight in each bin according to their impact parameter. For each object, we computed the mean and the 1$\sigma$ dispersion of the integrated $N_{CIV}$ and $N_{HI}$ values of all the lines of sight in each bin. We report this modified $N_{CIV}$ versus $N_{HI}$ relation for the MUPPI and the Effective models in Figs 8 and 9.
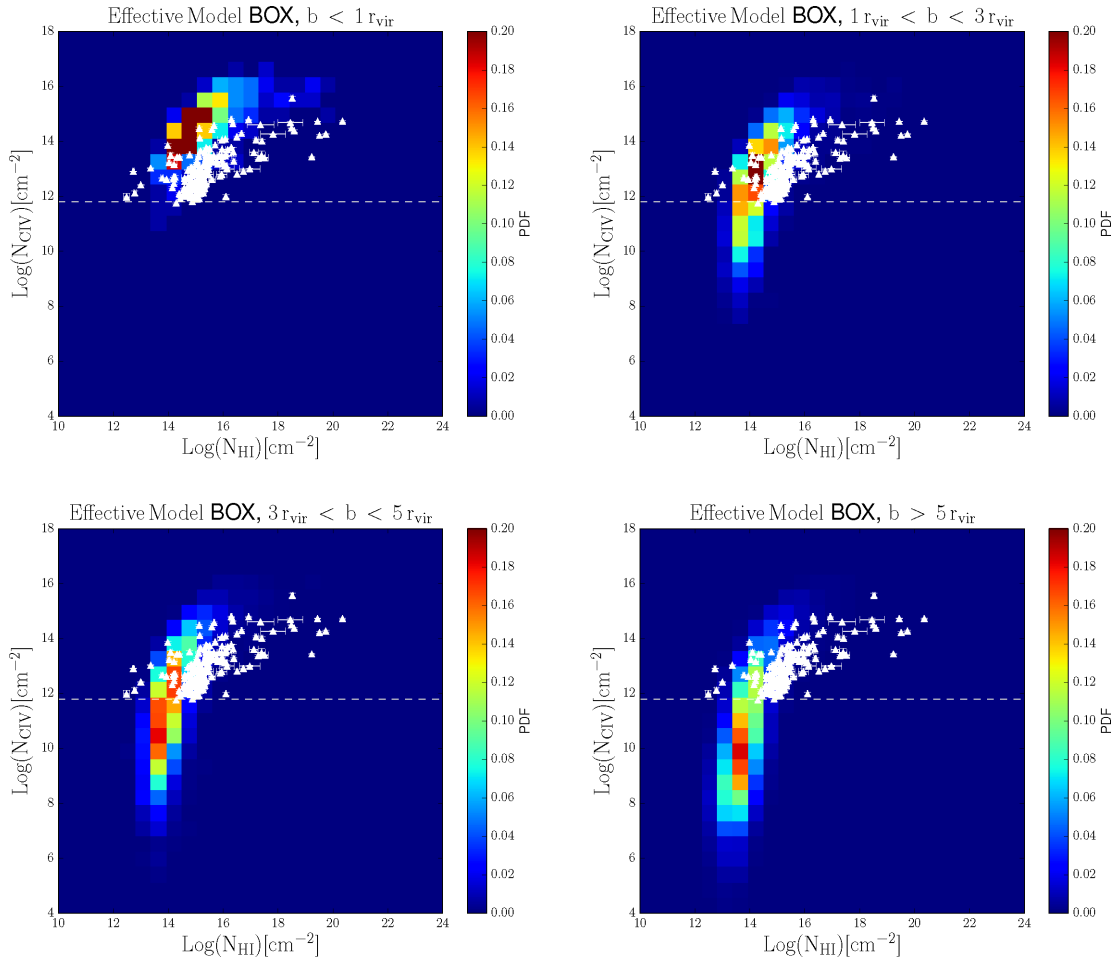
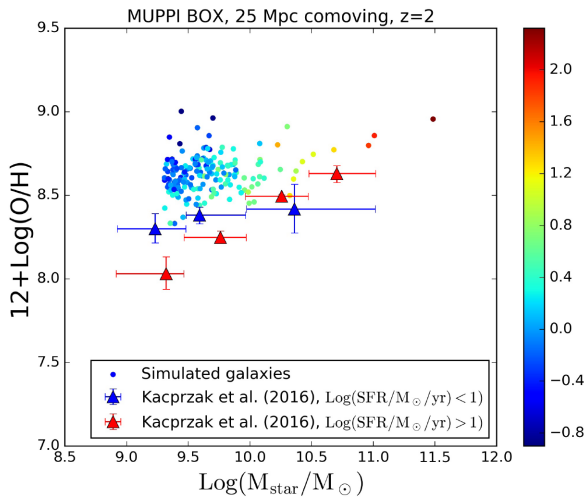**Figure 3.** Same as Fig. 2, but for the Effective model simulation.



**Figure 4.** Mass–metallicity relation for all the MUPPI galaxies with $M_* > 2 \times 10^9$ M$_\odot$ compared to the observed one by Kacprzak et al. (2016). MUPPI galaxies are colour coded according to their SF rate.
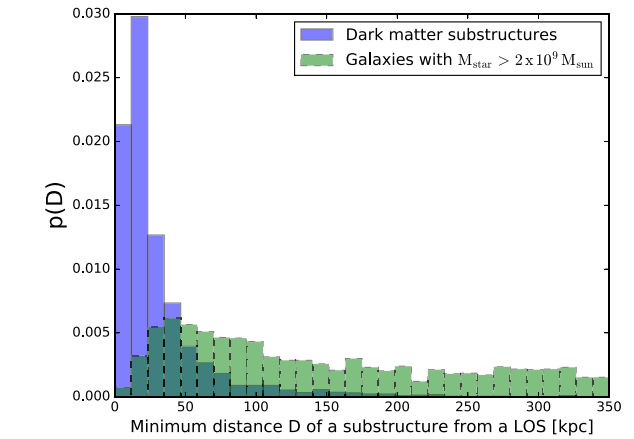


**Figure 5.** PDF of the minimum proper distance between an LOS with $b > 5$ $r_{vir}$ and with an absorption system with $14.0 \leq \log(N_{C\,IV}/cm^{-2}) \leq 18.0$ and a DM substructure identified with SUBFIND (blue solid histogram) or a galaxy with stellar mass $M_* > 2 \times 10^9$ M$_\odot$ (green dashed histogram).

For each galaxy, we divided the mean values according to their distance from galaxy centre, so plotting mean values whose radial bin is at a distance from centre between [0–1] $r_{vir}$, [1–3] $r_{vir}$, [3–5] $r_{vir}$, and r > 5 $r_{vir}$. Each point refers to a single LOS. Blue points

refer to lines of sight around galaxies with a total halo mass between $M_h \sim 10^{11}$ and $10^{12}$ M$_\odot$, while cyan points refer to galaxies with a total halo mass in the range $M_h \sim 10^{10}$–$10^{11}$ M$_\odot$. Yellow represents near-filament environments. Magenta points represent the
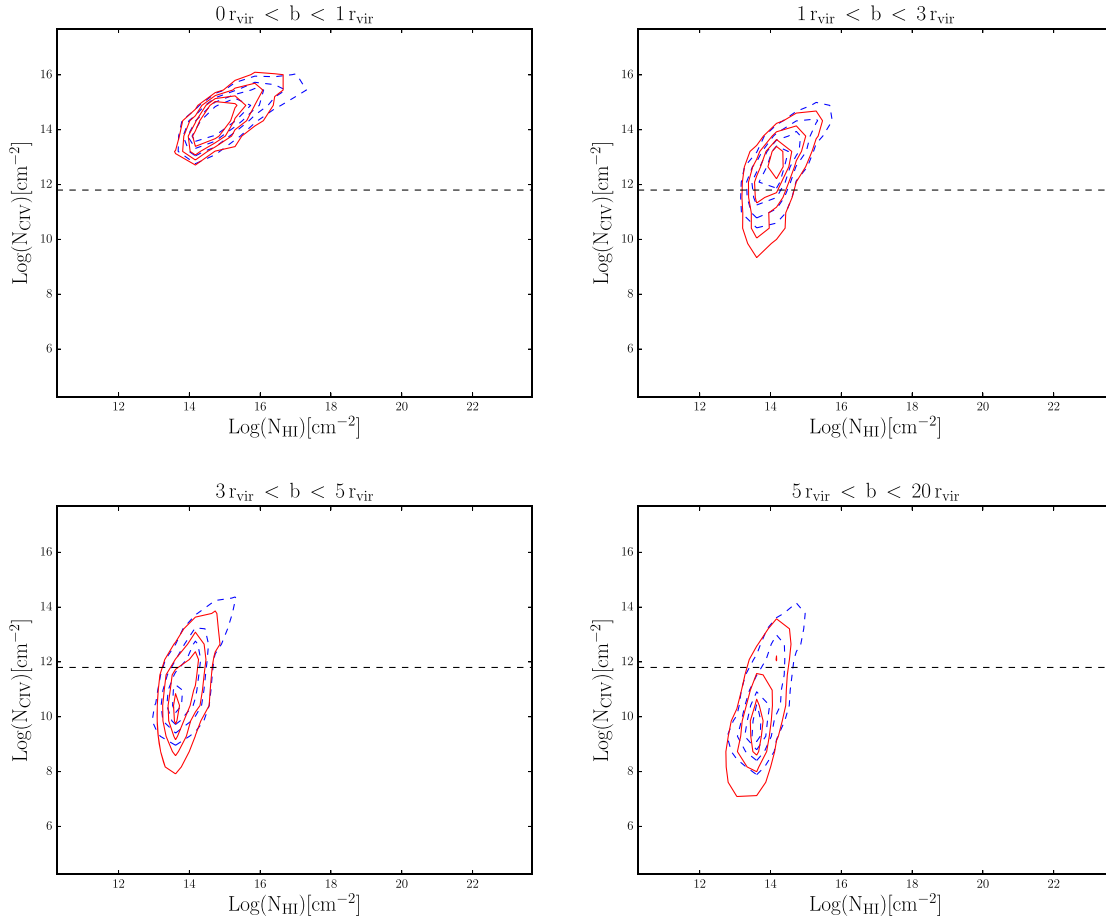
**Figure 6.** Comparison of the PDFs of the MUPPI (blue dashed) and the Effective (red solid) models for the considered distance ranges. Contour levels correspond to PDF values of [0.05,0.1,0.15,0.2]. Dashed horizontal line: observational detection limit.
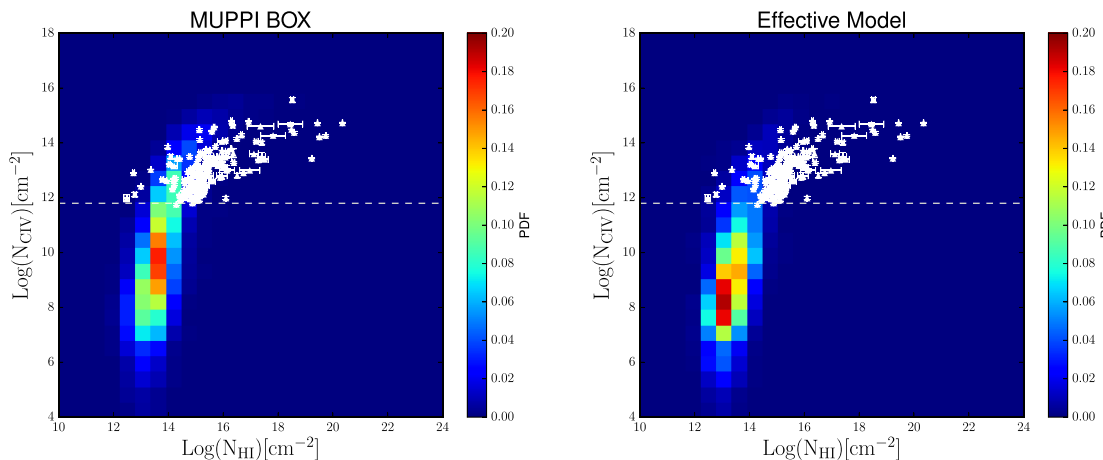


**Figure 7.** PDF for the $N_{C IV}$ versus $N_{H I}$ relation of lines of sight around near-filament environments. All lines of sight with impact parameter less than 800 kpc are considered. White triangles: observational data from K16. Dashed horizontal line: observational detection limit. Left-hand panel: results for the MUPPI model. Right-hand panel: results for the Effective model.

observational sample by Kim et al. (2016) and the green horizontal dashed line is the detection limit of this sample.

For near-filaments environments, it is obviously not possible to define a virial radius, so we plotted in each panel of Figs 8 and 9, the mean values of all the radial bins from 0 to 800 kpc.

Large error bars, representing the $1\sigma$ dispersion, are due to the fact that the medium is not homogeneous, so at same radial distance we could find a low-density environment, having low column densities, or we could hit (or be close to) a substructure, having higher values of column densities.
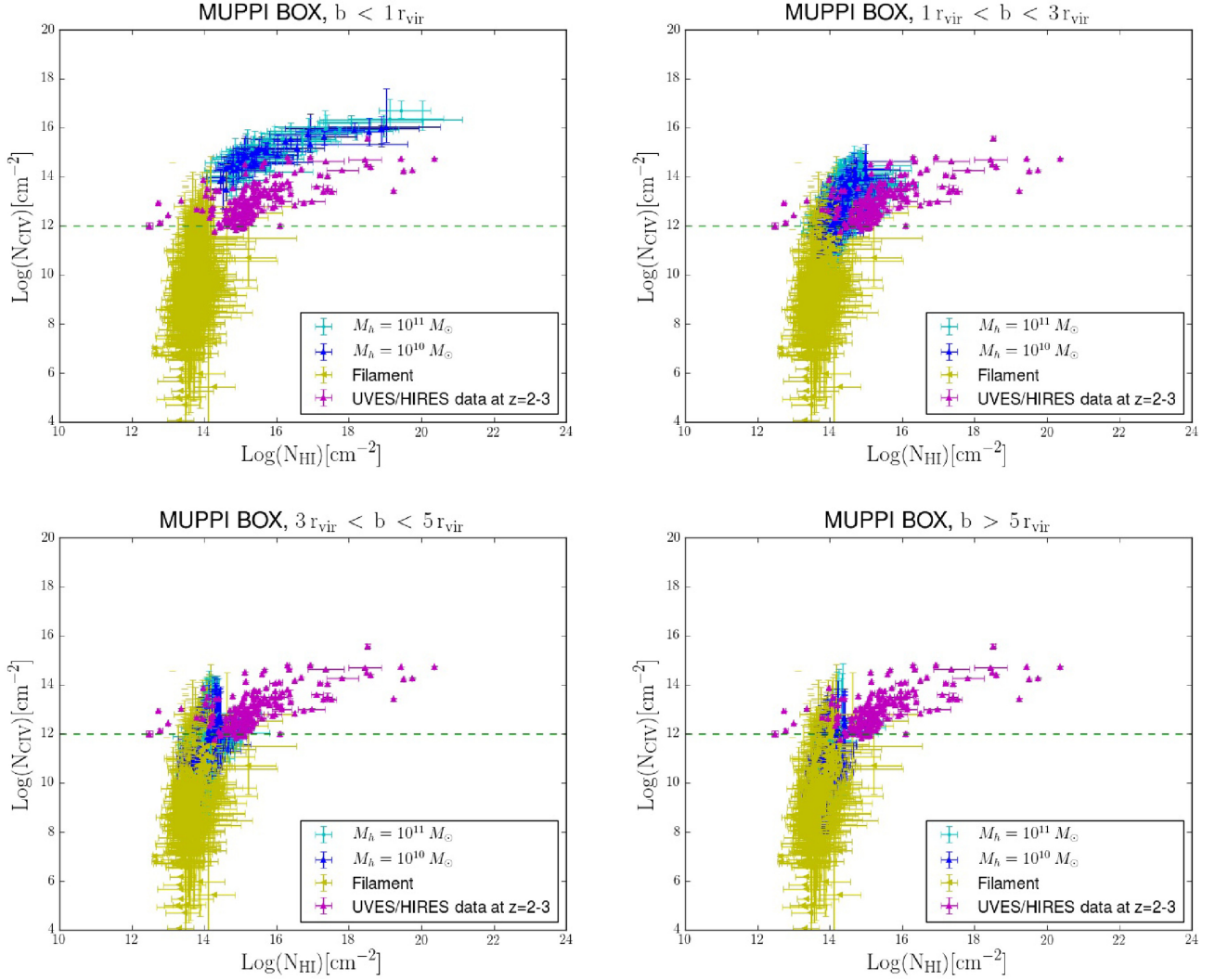
**Figure 8.** Each point is the spherical average value of H I and C IV column densities profiles in radial bins of width of $\sim 10$ kpc around the 30 objects (20 galaxies and 10 near-filament environments) for the MUPPI simulation. Error bars represent the $1\sigma$ dispersion. Cyan points: mean values of lines of sights around haloes with $M_h \sim 10^{11}$ M$_\odot$. Blue points: mean values of lines of sights around haloes with $M_h \sim 10^{10}$ M$_\odot$. Yellow points: mean values of lines of sights around near-filaments points. Magenta points: observational data from K16 as in previous plots (the same data are reported in each panel). Green horizontal dashed line: observational detection limit. Plots have been divided according to the distance of radial galaxy spherical averages from galaxy centre in units of virial radius. This division does not concern near-filaments points, whose all values are reported in each plot.

As in the previous section, we can see that for galaxies median values shift to lower column densities as the distance increases and we do not find any correspondence between simulated and observed data above 3–5 $r_{\rm vir}$ except for the weakest systems of the sample of K16. Near-filaments points are confined to a region under the detection limit, suggesting that they have metallicities too low to be probed by present-day observations. First results in this sense have been obtained with extremely high S/N spectra (D'Odorico et al. 2016; Ellison et al. 2000). However, to collect statistical samples of low column density metal absorbers we will have to wait for next generation high-resolution spectrographs, like ESPRESSO (Echelle SPectrograph for Rocky Exoplanet and Stable Spectroscopic Observations) at the VLT (Pepe et al. 2014) or the HIRES spectrograph at the E-ELT (European-Extremely Large Telescope, Marconi et al. 2016).

### 5.3 Comparison between MUPPI and the Effective model

As already stated, the subresolution models do not show strong differences in the results for C IV.

To better investigate this last issue, we performed the same overall analysis with two other ions: O VI and Si IV. The results for the O VI are discussed in Appendix C, since the observational data were taken from the literature and the procedure to compute column densities is slightly different from our approach. In general, O VI shows a similar behavior as C IV without any strong differences between the two models, apart from the tail of the Effective model, which extends more to low values of O VI column densities in the range above 1 $r_{\rm vir}$.

The results for Si IV are instead shown in Figs 10 and 11. For the Si IV, we compared always with the observational sample by K16.
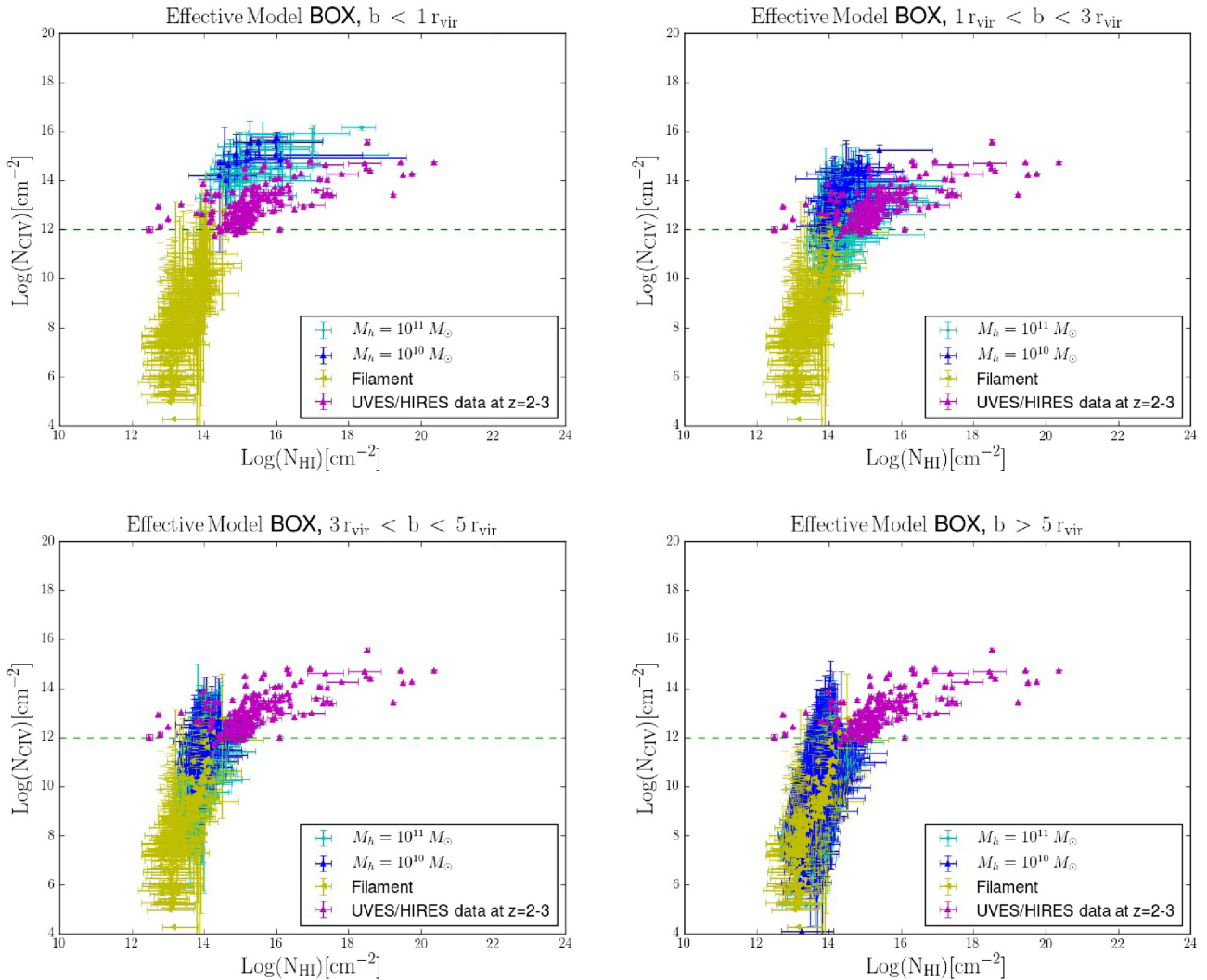
**Figure 9.** Same as Fig. 8, but for the Effective model simulation.

Unlike C IV and O VI, we do not see any difference between the models in every radial bin. Si IV traces a gas phase at higher $N_{\mathrm{H\,I}}$, closer to galaxies. At these close distances to galaxies, the effects of the feedback could not be traced.

One explanation of these results could be that the differences in the feedback prescriptions of the two models could be more related to the thermal feedback than to the kinetic one. Fig. 12 shows a temperature map around the same galaxy for both models, in which the mean value of the temperature profile along the LOS in a slice ±300 proper kpc is computed. It is clear that the MUPPI model is more capable of heating the surrounding gas with respect to the Effective model, while the differences related to kinetic feedback, that is how metals are distributed outside galaxies, are less visible.

Barai et al. (2015) already studied different feedback schemes, among which the ones used in this work. In their paper, they constructed radial profiles of the total gas metallicity around galaxy centres at $z \sim 2$ and they inferred that the MUPPI model distributes metals more adequately than the Effective model. We do also recover a slight difference between the two models in the range above $1\,r_{\mathrm{vir}}$, but when comparing with observational data, these differences seem not to significantly impact on the IGM properties investigated

here. It is important to highlight, as already said in Section 2, that the MUPPI simulation, that we used, was run with slightly different model parameters, due to small changes in the chemical sector with respect to the one by Barai et al. (2015).

In Fig. 13, we show a comparison for a particular galaxy between column density maps for H I, C IV, O VI, and Si IV calculated with the AOD method as previously done, so taking into account gas peculiar motions. Also by looking at the distributions of the gas around the galaxy for each element, we do not see any strong difference between the two models.

## 5.4 Covering fraction

The covering fraction of a given ion is by definition the ratio between the number of lines of sight showing an absorption system due to that ion with an equivalent width (EW) greater than a threshold value and the total number of lines of sight. In practice, it is a measure of the clumpiness of the medium (which depends also on the ionization of the medium itself). If the medium is perfectly homogeneous, the covering fraction is equal to 1. The clumpier the medium, the lower the covering fraction is. In Fig. 14, we report the C IV covering fractions for our sample of MUPPI data,
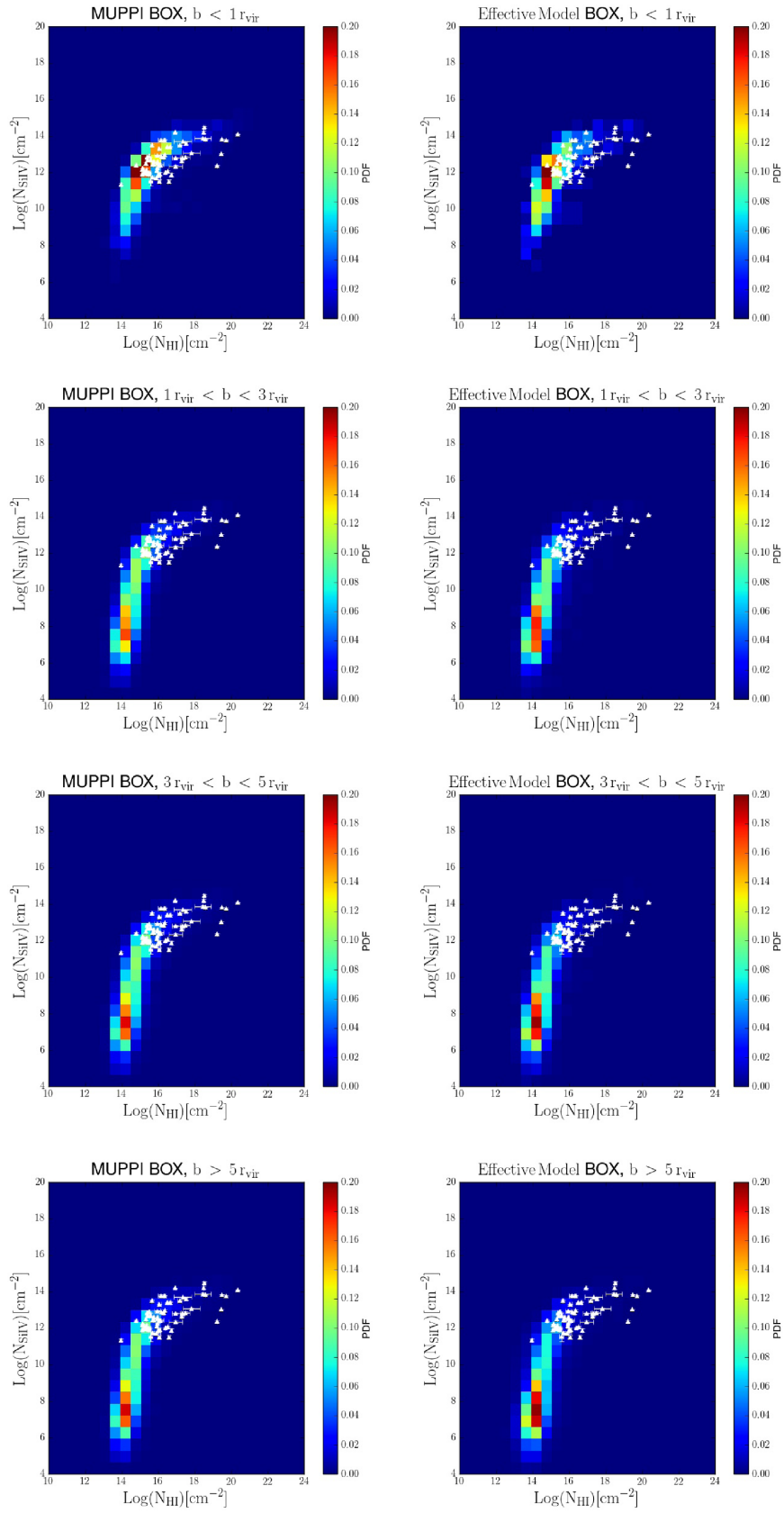
**Figure 10.** PDF of $N_{\mathrm{SiIV}}$ versus $N_{\mathrm{HI}}$ relation. White points: observational data from K16 (the same data are reported in each panel).
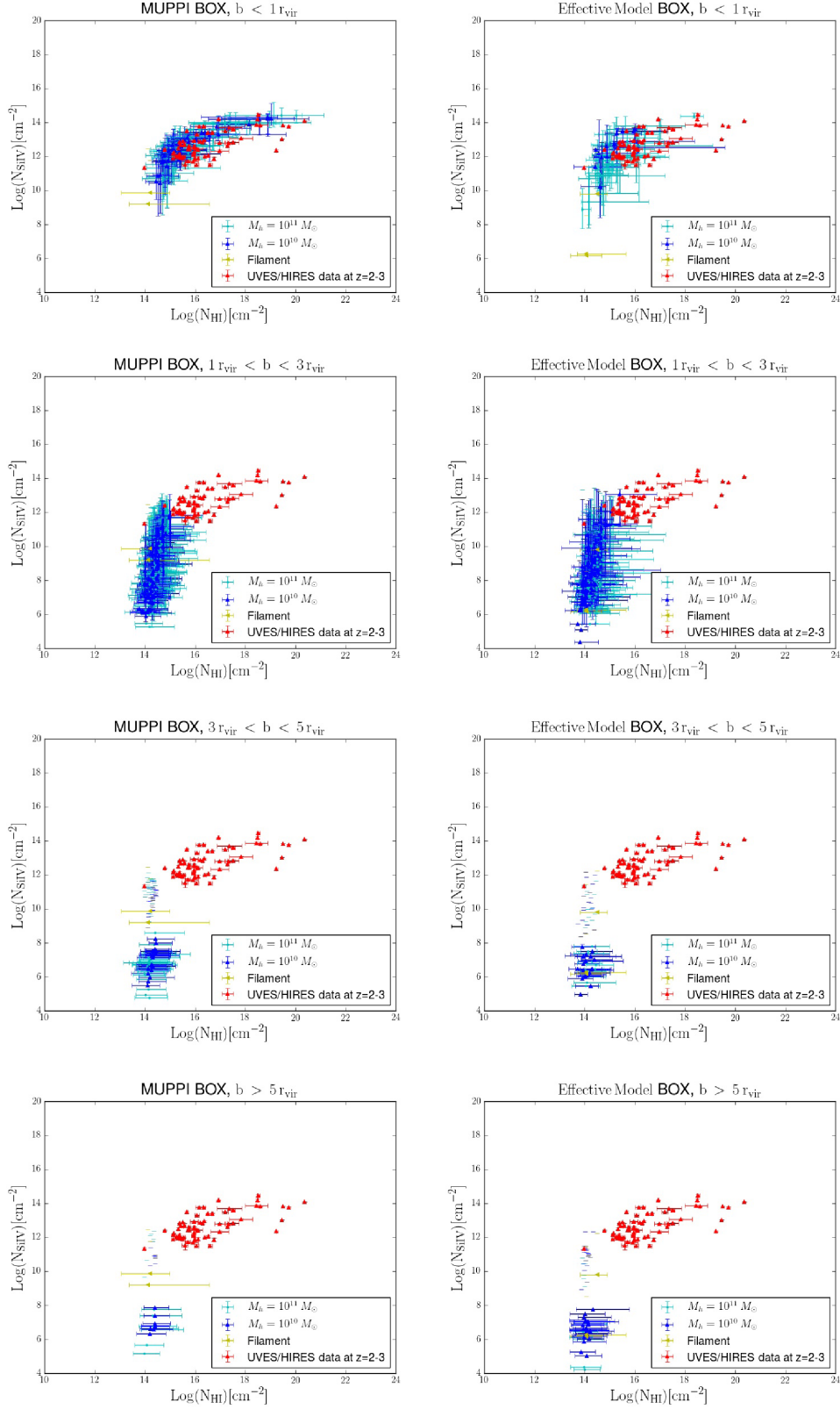
**Figure 11.** Each point is the spherical average value of H I and Si IV column densities profiles in radial bins of width of $\sim 10$ kpc around the 30 objects (20 galaxies and 10 near-filament environments) for the MUPPI simulation. Error bars represent the $1\sigma$ dispersion. Cyan points: mean values of lines of sights around haloes with $M_{\rm h} \sim 10^{11}$–$10^{12}$ M$_\odot$. Blue points: mean values of lines of sights around haloes with $M_{\rm h} \sim 10^{10}$–$10^{11}$ M$_\odot$. Yellow points: mean values of lines of sights around near-filaments points. Red points: observational data from table 4 of K16 (red points are equal to white points in Fig. 10); the same data are reported in each panel. Plots have been divided according to the distance of radial galaxy spherical averages from galaxy centre in units of virial radius. Near-filament points are the same in each plot.
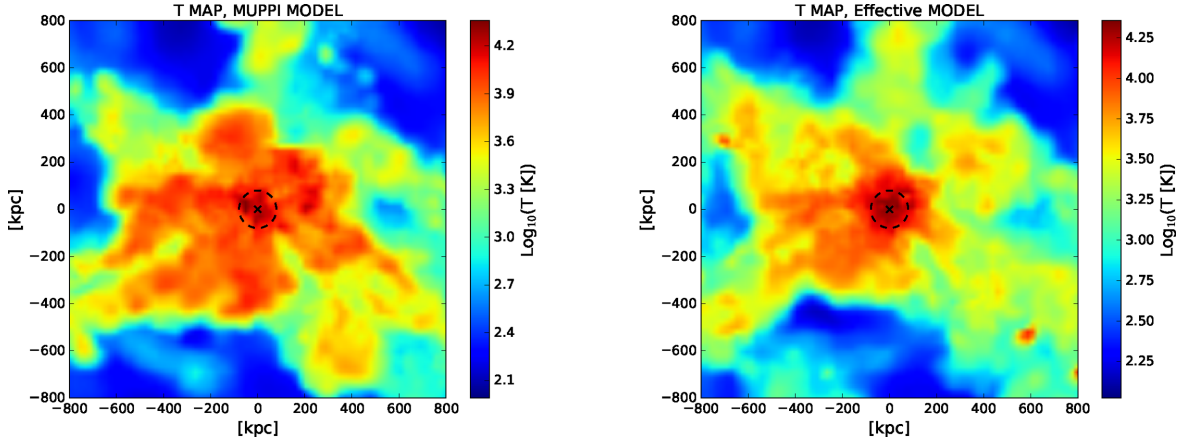
**Figure 12.** Comparison between temperature maps of the same galaxy between the MUPPI and the Effective models. Temperatures are computed by taking the mean value of the temperature profile along the LOS in a slice $\pm$ 300 kpc from galaxy's position. (Distances are in proper kpc.)

calculated in bins of 100 kpc, in which we computed EWs by always integrating $\pm$ 150 km s$^{-1}$ from galaxy position along the LOS. Dashed and dot–dashed lines are the median values of the C IV covering fractions of our total sample of lines of sight for two different threshold values of the EW. Shaded areas are the 1$\sigma$ confidence intervals.

In Fig. 15, instead, we show the covering fractions obtained by integrating $\pm$ 600 km s$^{-1}$, in order to be consistent with the sample of Landoni et al. (2016) and Rubin et al. (2015).

Our calculated covering fraction (orange line with EW > 0.3 Å) is in good agreement with the sample of Landoni et al. (2016). The fact that the data by Landoni et al. (2016) have a slightly higher normalization than our prediction could be explained by larger masses of the galaxies they are probing (since they consider QSO host galaxies). For this reason, it should be more correct to report the relations as a function of the distance in units of virial radius. Another possible bias is related to the fact that they are observing all face-on galaxies. In other directions, their covering fraction could be lower, due to inhomogeneities of the medium.

The sample of Rubin et al. (2015) is actually formed by 40 DLAs, so a little bit different from our work. The masses of the structures forming the DLAs are not known and this could explain why the sample of Rubin et al. (2015) has slightly higher values.

The sample of Prochaska et al. (2014) has a window of integration of $\pm$ 1500 km s$^{-1}$, due to uncertainties in the redshifts of their quasar sample. Unfortunately, our box size is not big enough, as our maximum range of integration is $\pm$ 800 km s$^{-1}$. We tried to integrate along all the box size and we did recover higher values with respect to those shown in Fig. 15, but still not in agreement with the Prochaska et al. (2014) sample. The comparison with a bigger cosmological box can be done in a future work.

Computing the covering fraction as a function of the EW is what it has been done so far in the literature, as measuring the EW of an absorption system is the simplest thing that can be done, especially with low-resolution data. With the advent of high-resolution spectroscopy, it has become possible to measure column densities of absorption systems, using Voigt Profile Fitting codes, which give a more reliable and complete information on the physical state of the gas that produced the observed absorption.

For comparison with future samples of high-resolution data, we decided to compute the covering fraction as a function of the column density, as it can be seen in Fig. 16. We divided the covering fractions in different column densities ranges: 0 $\leq$ log(N$_{CIV}$/cm$^{-2}$) $\leq$ 10.0, 10.0 $\leq$ log(N$_{CIV}$/cm$^{-2}$) $\leq$ 12.0, 12.0 $\leq$ log(N$_{CIV}$/cm$^{-2}$) $\leq$ 14.0, 14.0 $\leq$ log(N$_{CIV}$/cm$^{-2}$) $\leq$ 16.0, and 16.0 $\leq$ log(N$_{CIV}$/cm$^{-2}$) $\leq$ 18.0. We can see that C IV absorption systems with log(N$_{CIV}$/cm$^{-2}$) > 14.0 dominate in the haloes of galaxies ($d$ < 100 kpc), even if their covering fraction does not drop to zero at greater distances, due to the presence of substructures. At higher distances, weaker systems start to dominate, with their covering fraction increasing.

## 6 CONCLUSIONS

We analysed the output of high-resolution hydrodynamical simulations with SN feedback implemented both in the thermal and kinetic forms. In particular, two different subresolution models were considered: the MUPPI model (Murante et al. 2010, 2015) and the Effective model (Springel & Hernquist 2003). These two sets share the same large-scale structure evolution, but they are decoupled in the hydrodynamical part, as they have different SF and feedback prescriptions.

The main findings of this work can be summarized as follows.

(i) We performed a state-of-the-art post-processing analysis and produced a set of mock galaxies and mock spectra. Spectra are constructed using an SPH formulation and profiles of different physical quantities can be reproduced along the LOS. A sample of 40 galaxies with halo mass in the range $M_{\rm h} \sim 10^{10}$–$10^{12}$ M$_\odot$, half of which has been selected from the MUPPI simulation and the other half from the Effective model box. They have to satisfy the following criteria: galaxies do not have to be on a major merger, they have to be the main halo of the FOF group, in order to analyse the same conditions at the same distance and be sampled by a sufficient number of particles. We also considered 20 near-filaments environments from the two runs, by visually centring our volume as close as possible to a filamentary structure.

(ii) We identified an optimal analysis of mock QSO spectra based on the AOD method and compared to the results obtained with Voigt
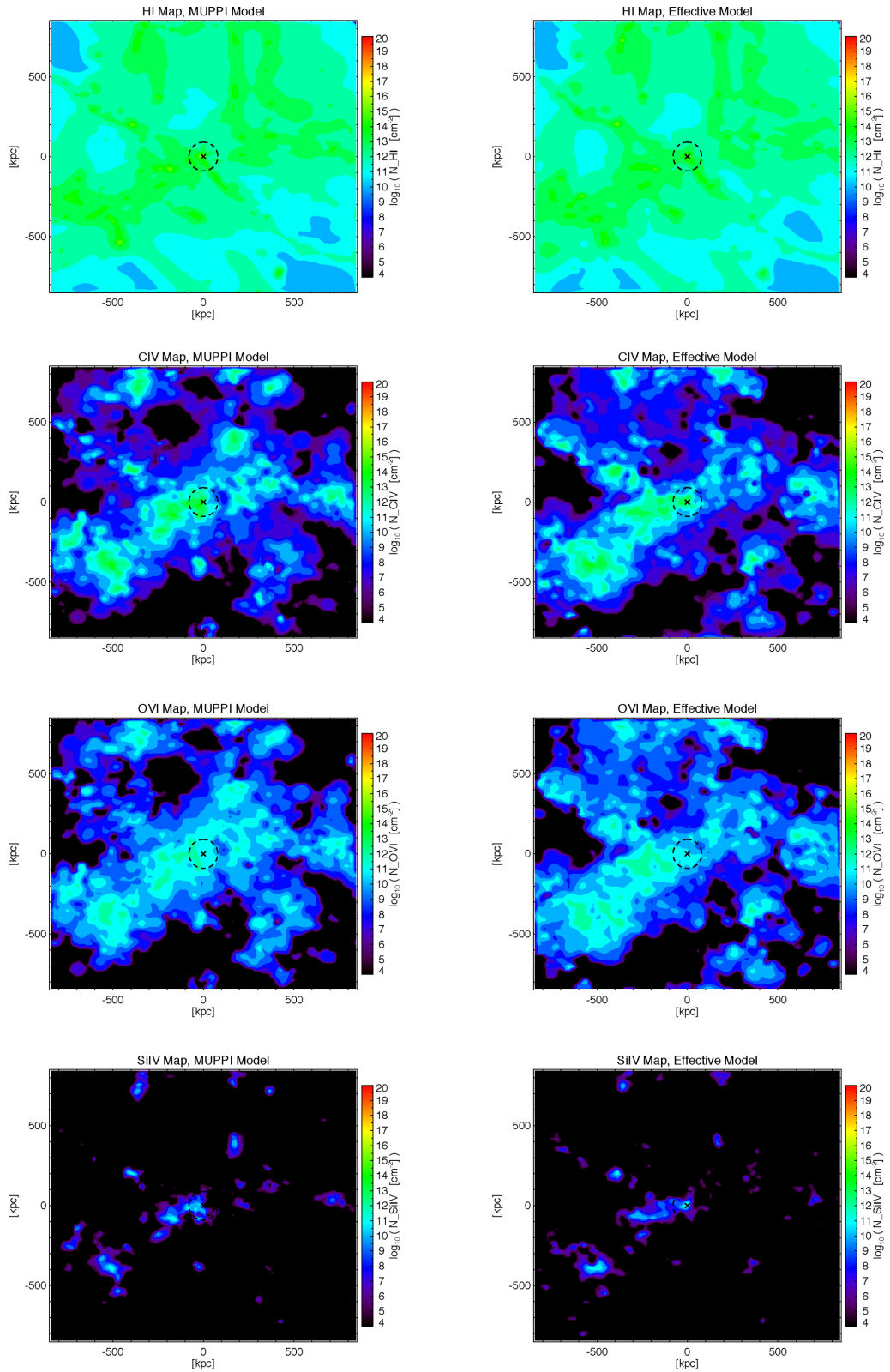
**Figure 13.** Comparison between column density maps of different chemical elements between the MUPPI and the Effective models. Column densities are computed using the AOD method $\pm$ 150 km s$^{-1}$ from galaxy's position. (Distances are in proper kpc.)

Profile Fitting methods, as described in Appendix B1. We performed a comparison between column densities derived with the VPFIT code

and the ones directly derived from the simulation. We found the best and faster method to use on mock QSO spectra when comparing
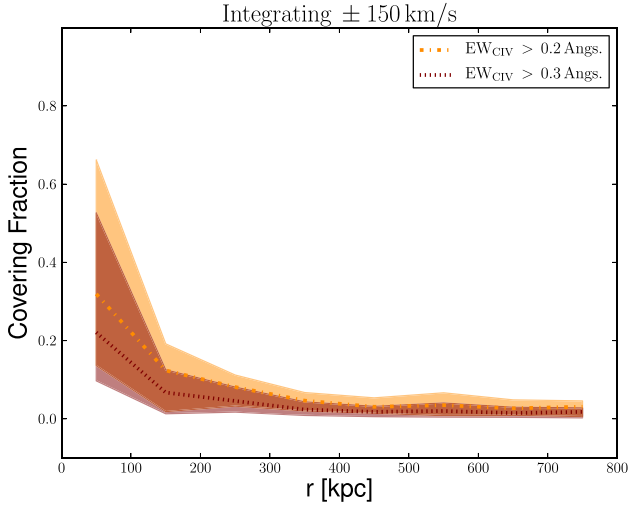
**Figure 14.** C IV covering fraction of our sample of lines of sight, in which rest-frame EW have been integrated ± 150 km s$^{-1}$ from galaxy position. Dotted and dot–dashed lines are the median values corresponding to two different EW thresholds (orange: EW$_{\rm CIV}$ >0.2 Å; and red: EW$_{\rm CIV}$ >0.3 Å), while shaded areas are the 1$\sigma$ confidence intervals. (Distances are in proper kpc.)
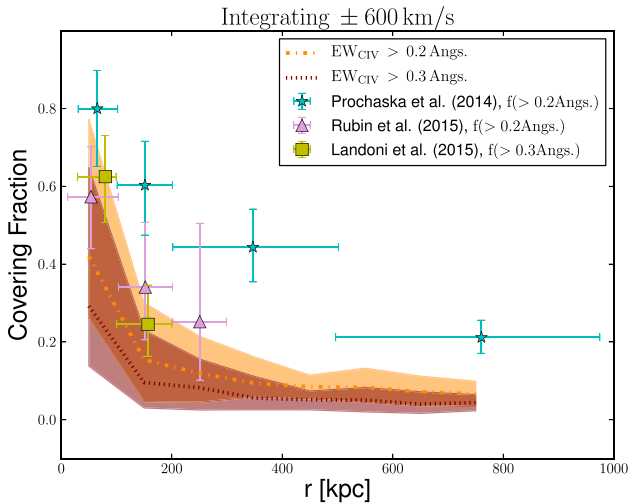


**Figure 15.** C IV covering fraction of our sample of lines of sight, in which rest-frame EW have been integrated ± 600 km s$^{-1}$, to be consistent with the observational sample of Rubin et al. (2015) and Landoni et al. (2016). The integration window of Prochaska et al. (2014) is ± 1500 km s$^{-1}$, a value larger than our box size. This explains the disagreement with the sample of Prochaska et al. (2014). Dotted and dot–dashed lines are the median values corresponding to two different EW thresholds (orange: EW$_{\rm CIV}$ >0.2 Å; and red: EW$_{\rm CIV}$ >0.3 Å), while shaded areas are the 1$\sigma$ confidence intervals. (Distances are in proper kpc.)

with observational data is the AOD method. With this method, gas peculiar motions are taken into account and, at the time, a reliable measure of the column density is obtained.

(iii) We performed a full environmental characterization of the CGM and IGM for absorption systems. In particular, we pierced 4000 lines of sight around each selected object in the simulation with impact parameters less than 800 kpc. Using the C IV as a tracer of the metallicity, we constructed the $N_{\rm CIV}$ versus $N_{\rm HI}$ relation in different radial bins from the object's centre, which we compared

with observational data by K16. We found that observational data have the highest probability to be confined in a region up to 3–5 virial radii from galaxies, which correspond (at this redshift) to a physical distance of ∼ 150–400kpc. Near-filament points are instead confined to a region under the detection limit, suggesting that they have metallicities too low to be probed by present-day observations. These results are validated by the fact that we do not recover any strong difference between the MUPPI and the Effective models, which display similar trends in the results.

(iv) We presented a probabilistic approach to the galaxy/IGM interplay which carefully quantifies the probability to find an absorption system with a given column density at a certain distance from a galaxy. For example, if a system with a column density $N_{\rm CIV}$ of ∼$10^{15}$ cm$^{-2}$ and $N_{\rm HI}$ of ∼$10^{16}$ cm$^{-2}$ is found, it has a probability five times higher to be located inside the virial radius of a galaxy of mass ∼$10^{10-11}$ M$_\odot$ than at a distance of 3–5 virial radii from such a galaxy. If such a system is observed at this great distance from a luminous galaxy, it could be actually related to another smaller and less luminous galaxy.

(v) We quantified the distribution of ionization species around galaxies in terms of covering factors and compared to data taken from Landoni et al. (2016), Rubin et al. (2015), and Prochaska et al. (2014). Considering also the 1$\sigma$ dispersion, our data are in good agreement with the observed one, even if they show a somewhat lower normalization. This could be due to the fact that the observational samples probe galaxies with slightly higher masses than those in our simulation. Covering fractions are computed as a function of the EW of the absorption system. Furthermore, we constructed covering fractions depending on the column density of absorption systems, which can be used to compare with future data. Higher column density systems (log($N_{\rm CIV}$/cm$^{-2}$) ≥ 14) dominate at impact parameters ≲ 100 kpc, while lower column density systems become prevalent at larger distances. The contribution of higher column density systems does not drop to zero with increasing separation due to the fact that galaxies are not isolated systems, but they are surrounded by many other structures.

(vi) We extended the work also to other two chemical elements, O VI and Si IV. Simulated O VI absorption systems show a very similar behaviour as C IV. This is an indication that these two elements trace a similar gas phase, having both the highest probability to be confined up to a few virial radii from galaxies, characterizing a region typical of the CGM. The comparison with the observational sample suffers from the limitations of the simulation, which is not capable to properly treat ionization. Si IV traces, instead, more internal regions, as it has the highest probability to be observed inside the virial radius, as the comparison with data shows.

(vii) The two simulations which we analysed in this work give similar results apart from the comparison in the radial bins $b > 1$ $r_{\rm vir}$, for C IV and O VI. The Effective Model has a higher probability to have absorption systems with smaller values of H I–C IV–O VI column density and this could be due to a less efficient feedback with respect to the MUPPI model, which is not capable to spread metals with the same strength. For Si IV no difference is observed, but as it is confined in a smaller region, the effects of the feedback could be less visible. This is a quite surprising and unexpected result from the simulation point of view, as the two models are significantly different, both in the SF and feedback prescriptions. It suggests that galactic feedback and different SF processes are not strongly impacting on the IGM properties investigated here.

In conclusion, we have shown that observed integrated $N_{\rm CIV}$ versus $N_{\rm HI}$ relation likely arises in a region around the galactic halo,
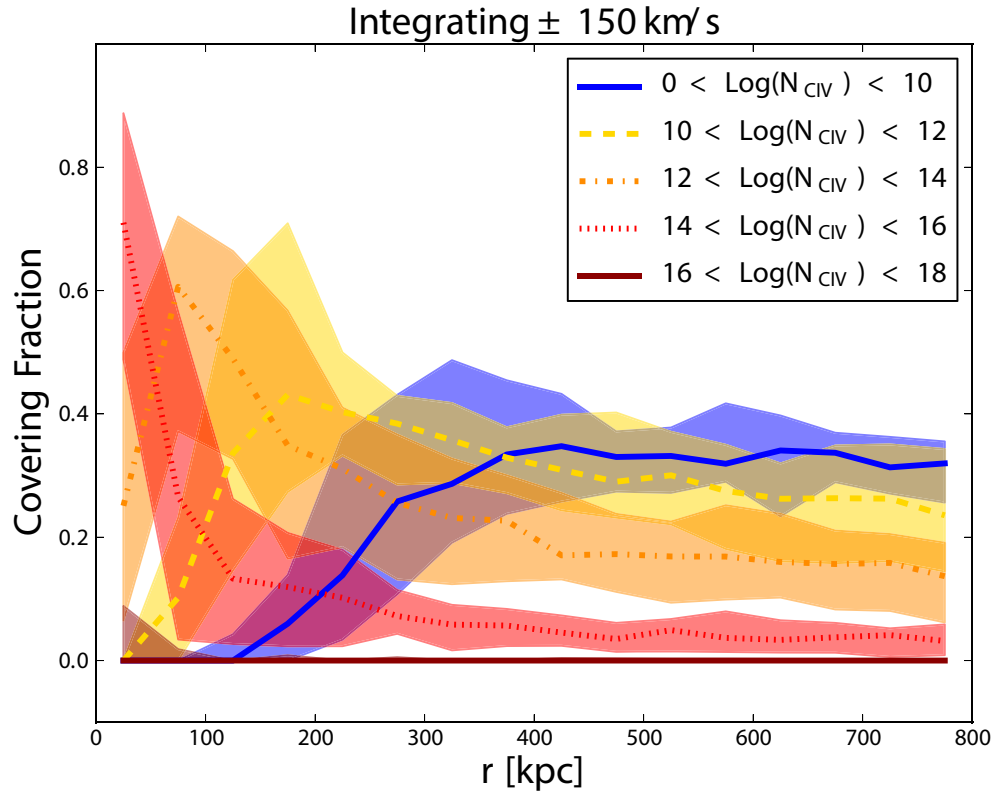
**Figure 16.** C IV covering fraction of our sample of lines of sight as a function of the column density. Column densities are integrated $\pm\,150\,\mathrm{km\,s^{-1}}$ from galaxy position. Dotted and dot–dashed lines are the median values, while shaded areas are the $1\sigma$ confidence intervals. (Distances are in proper kpc.)

more typical of a CGM environment (meaning the region surrounding the galactic halo up to $\sim 300$–$400$kpc) than of the IGM. This could be a hint that physical processes that are responsible for the metal enrichment did not affect significantly regions at very low density.

The fact that we obtain similar results in two different numerical models used to interpret the data sets suggests that these results are qualitatively correct.

Similar studies on the importance of absorption lines systems in the proximity of galaxies as derived from hydrodynamical simulations have been presented in several works (e.g. Hummels et al. 2013; Suresh et al. 2015; Bordoloi et al. 2014; Turner et al. 2016). However, it is quite difficult to compare with these findings given the different problems addressed and also the physical implementations of feedback models and the setup of the cosmological simulations (in terms of *considered redshift*, box sizes, objects studied and numerical resolution). Our work relies on a subresolution model (the MUPPI one), which succeeded in reproducing realistic disc galaxies with new implementations of the SF and feedback subresolution recipes.

In the future, it would be interesting to better investigate the comparison between the different subresolution models, in order to understand which physical processes affect more deeply the IGM properties.

From the observational point of view, the comprehension of the enrichment mechanism could be enlarged thanks to the new facilities that will come online in the next decade. The very weak metal absorptions associated with the IGM will be revealed, if present, by the next generation of spectrographs like ESPRESSO at the VLT, which will see the first light in 2017, and even more with the high-resolution spectrograph for the 39 m European ELT telescope. On the other hand, the relation between the level of enrichment of the IGM studied with metal absorbers and the distribution of galaxies in the same field will be better quantified when also the very faint objects will be identified, in particular with the *JWST* (James Webb Space Telescope).

## REFERENCES

Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005, ApJ, 629, 636
Aguirre A., Hernquist L., Schaye J., Weinberg D. H., Katz N., Gardner J., 2001a, ApJ, 560, 599
Aguirre A., Hernquist L., Schaye J., Katz N., Weinberg D. H., Gardner J., 2001b, ApJ, 561, 521
Aracil B., Petitjean P., Pichon C., Bergeron J., 2004, A&A, 419, 811
Barai P. et al., 2013, MNRAS, 430, 3213
Barai P., Monaco P., Murante G., Ragagnin A., Viel M., 2015, MNRAS, 447, 266
Barnes L. A., Haehnelt M. G., Tescari E., Viel M., 2011, MNRAS, 416, 1723

Blitz L., Rosolowsky E., 2006, ApJ, 650, 933
Boksenberg A., Sargent W. L. W., 2015, ApJS, 218, 7
Bordoloi R. et al., 2014, ApJ, 796, 136
Borthakur S. et al., 2015, ApJ, 813, 46
Borthakur S. et al., 2016, ApJ, 833, 259
Carswell R. F., Webb J. K., 2014, Astrophysics Source Code Library, record ascl:1408.015
Cen R., Nagamine K., Ostriker J. P., 2005, ApJ, 635, 86
Combes F., García-Burillo S., Braine J., Schinnerer E., Walter F., Colina L., Gerin M., 2006, A&A, 460, L49
Cowie L. L., Songaila A., Kim T.-S., Hu E. M., 1995, AJ, 109, 1522
Crain R. A., McCarthy I. G., Schaye J., Theuns T., Frenk C. S., 2013, MNRAS, 432, 3005
Crighton N. H. M. et al., 2011, MNRAS, 414, 28
D'Odorico V. et al., 2016, MNRAS, 463, 2690
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Davé R., Hellsten U., Hernquist L., Katz N., Weinberg D. H., 1998, ApJ, 509, 661
Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, MNRAS, 467, 115
Ellison S. L., Songaila A., Schaye J., Pettini M., 2000, AJ, 120, 1175
Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137
Ford A. B., Davé R., Oppenheimer B. D., Katz N., Kollmeier J. A., Thompson R., Weinberg D. H., 2014, MNRAS, 444, 1260
Frye B., Benitez N., Broadhurst T., 2002, in American Astronomical Society Meeting Abstracts #200. p. 705
Georgakakis A., Rowan-Robinson M., Babbedge T. S. R., Georgantopoulos I., 2007, MNRAS, 377, 203
Gnedin N. Y., 1998, MNRAS, 294, 407
Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581
Goz D., Monaco P., Granato G. L., Murante G., Domínguez-Tenreiro R., Obreja A., Annunziatella M., Tescari E., 2017, MNRAS, 469, 3775
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, p. 64
Haiman Z., Loeb A., 1997, ApJ, 483, 21
Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, ApJS, 129, 493
Hinojosa-Goñi R., Muñoz-Tuñón C., Méndez-Abreu J., 2016, A&A, 592, A122
Hirschmann M. et al., 2013, MNRAS, 436, 2929
Hummels C. B., Bryan G. L., Smith B. D., Turk M. J., 2013, MNRAS, 430, 1548
Kacprzak G. G. et al., 2016, ApJ, 826, L11
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kim T.-S., Viel M., Haehnelt M. G., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Kim T.-S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, MNRAS, 382, 1657
Kim T.-S., Partl A. M., Carswell R. F., Müller V., 2013, A&A, 552, A77
Kim T.-S., Carswell R. F., Mongardi C., Partl A. M., Muecket J. P., Barai P., Cristiani S., 2016, MNRAS, 457, 2005
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Landoni M., Falomo R., Treves A., Scarpa R., Farina E. P., 2016, MNRAS, 457, 267
Lemaux B. C. et al., 2014, A&A, 572, A90
Liang C. J., Chen H.-W., 2014, MNRAS, 445, 2061
Marconi A. et al., 2016, Proc. SPIE, 9908, 990823
Martin C. L., 1999, ApJ, 513, 156
Martin C. L., 2005, ApJ, 621, 227
Monaco P., Murante G., Borgani S., Dolag K., 2012, MNRAS, 421, 2485
Murakami I., Yamashita K., 1997, in Petitjean P., Charlot S., eds, Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System. Editions Frontieres, Paris, p. 77
Murante G., Monaco P., Giovalli M., Borgani S., Diaferio A., 2010, MNRAS, 405, 1491
Murante G., Monaco P., Borgani S., Tornatore L., Dolag K., Goz D., 2015, MNRAS, 447, 178
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569

Muzahid S., Srianand R., Bergeron J., Petitjean P., 2012, MNRAS, 421, 446
Nath B. B., Trentham N., 1997, MNRAS, 291, 505
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Oppenheimer B. D., Schaye J., 2013, MNRAS, 434, 1043
Ostriker J. P., Gnedin N. Y., 1996, ApJ, 472, L63
Padovani P., Matteucci F., 1993, ApJ, 416, 26
Pepe F. et al., 2014, Astron. Nachr., 335, 10,
Pettini M., Shapley A. E., Steidel C. C., Cuby J.-G., Dickinson M., Moorwood A. F. M., Adelberger K. L., Giavalisco M., 2001, ApJ, 554, 981
Prochaska J. X., Weiner B., Chen H.-W., Mulchaey J., Cooksey K., 2011, ApJ, 740, 91
Prochaska J. X., Lau M. W., Hennawi J. F., 2014, ApJ, 796, 140
Rahmati A., Schaye J., Pawlik A. H., RaicŽevic M., 2013, MNRAS, 431, 2261
Rubin K. H. R., Hennawi J. F., Prochaska J. X., Simcoe R. A., Myers A., Lau M. W., 2015, ApJ, 808, 38
Rupke D. S., Veilleux S., Sanders D. B., 2005, ApJS, 160, 87
Savage B. D., Sembach K. R., 1991, ApJ, 379, 245
Scannapieco E., Pichon C., Aracil B., Petitjean P., Thacker R. J., Pogosyan D., Bergeron J., Couchman H. M. P., 2006, MNRAS, 365, 615
Schaye J. et al., 2010, MNRAS, 402, 1536
Schaye J. et al., 2015, MNRAS, 446, 521
Schaye J., Theuns T., Leonard A., Efstathiou G., 1999, MNRAS, 310, 57
Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Songaila A., 1998, AJ, 115, 2184
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289
Stinson G. S. et al., 2012, MNRAS, 425, 1270
Suresh J., Bird S., Vogelsberger M., Genel S., Torrey P., Sijacki D., Springel V., Hernquist L., 2015, MNRAS, 448, 895
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
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Theuns T., Viel M., Kay S., Schaye J., Carswell R. F., Tzanavaris P., 2002, ApJ, 578, L5
Thielemann F.-K. et al., 2003, Nucl. Phys. A, 718, 139
Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050
Tornatore L., Borgani S., Viel M., Springel V., 2010, MNRAS, 402, 1911
Tremonti C. A. et al., 2004, ApJ, 613, 898
Tumlinson J. et al., 2011, Science, 334, 948
Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, MNRAS, 445, 794
Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2015, MNRAS, 450, 2067
Turner M. L., Schaye J., Crain R. A., Theuns T., Wendt M., 2016, MNRAS, 462, 2440
Tytler D., Fan X.-M., Burles S., Cottrell L., Davis C., Kirkman D., Zuo L., 1995, in Meylan G., eds, QSO Absorption Lines, Springer-Verlag, Berlin Heidelberg, p. 289
van den Hoek L. B., Groenewegen M. A. T., 1997, A&AS, 123, 305
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O'Meara J. M., Peeples M. S., 2013, ApJS, 204, 17
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
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181

## APPENDIX A: CONSTRUCTING LINES OF SIGHT

Simulated data are obtained by piercing through the cosmological box with lines of sight, either randomly or forcing them to have impact parameters smaller than a certain value around a given galaxy or position in the simulation.

Lines of sight are as long as the box side and are always parallel to it along one of the three directions $x - y - z$.

Each LOS is divided in 2048 pixels of $\Delta L \sim 10$ proper kpc each. Along each LOS, we can compute different quantities, such as the density profile of the total gas or of a particular chemical element, the temperature profile, the gas peculiar velocity, or the optical depth profile of a given chemical element.

Profiles are computed following the prescriptions by Theuns et al. (1998): the total gas density $\rho_i$ of each pixel $i$ is the sum of the densities of all gas particles $\rho_{part}$ (multiphase or single phase) convolved with their SPH kernel functions $W$, that means that all the gas particles $j$ contribute to the pixel's density; these particles have a smoothing length $h$ which intersect the LOS. The density of a gas particle is defined as $\rho_{part} = M_p/h^3$. The same procedure is applied to compute the other physical quantities, such as the gas peculiar velocity or the temperature.

$$\rho_i = \sum_j W_{ij}$$

$$v_i = \sum_j v_{part,j} W_{ij}$$

$$T_i = \sum_j T_{part,j} W_{ij} \qquad (A1)$$

Here, $W_{ij}$ is:

$$
\begin{aligned}
W_{ij} &= \rho_{part} \cdot W(q_{ij}) \\
&= \rho_{part} \cdot \left[ \frac{1}{4\pi}(4 - 6q_{ij}^2 + 3q_{ij}^3) \right] \quad , \text{if} \quad q_{ij} \le 1 \\
&= \rho_{part} \cdot \left[ \frac{1}{4\pi}(2 - q_{ij})^3 \right] \quad , \text{if} \quad 1 \le q_{ij} \le 2 \\
&= 0 \quad \text{otherwise} \qquad (A2)
\end{aligned}
$$

with

$$q_{ij} = \frac{|x(i) - x(j)|}{h} \qquad (A3)$$

$x(i)$ is the pixel's position and $x(j)$ is the particle's position.

We can also compute density profiles of a given chemical species in a given ionization state.

The metallicity of a specific chemical element in a single gas particle is defined as the ratio between the mass of that element in the gas particle and the total mass of the particle, such as:

$$Z_X = m_X/M_p \qquad (A4)$$

with X = [H, He, C, Ca, O, N, Ne, Mg, S, Si, Fe]. In order to have the fraction of that element in a particular ionization state $k$ for a specific gas particle, we used the formula:

$$f_{X,k} = f_{X,k,\odot} \cdot Z_X \qquad (A5)$$

$f_{X,k,\odot}$ is the fraction of an element X in the ionization state $k$ per solar metallicity. Given the density, the temperature and the redshift

of the particle, $f_{X,k,\odot}$ is computed using the CLOUDY code[3] (Ferland et al. 2013) with the UVB by Haardt & Madau (2005)[4]. For the H I, $f_{H I}$ is an output of the simulation and, as described in Section 2, it is computed considering a UVB by Haardt & Madau (2001), which is not significantly different from Haardt & Madau (2005) at $z \sim 2$. Defining $\rho_{X,k} = f_{X,k} \cdot \rho_{part}$ and using this quantity in place of $\rho_{part}$ in equation (A2), we can compute all the profiles of equation (A1) but for specific ions.

Optical depths along lines of sight were computed by considering their density, temperature, and velocity in each pixel. We consider a pixel $i$ and first calculate the central optical depth of the line, which falls in that pixel, using its density and temperature with this formula:

$$\tau_{0,i} = \frac{\sigma_\alpha c}{\sqrt{\pi}} \frac{\rho_i \Delta L}{b_i} \qquad (A6)$$

with $b_i = \sqrt{\frac{2K_B T_i}{m_{atom}}}$ and $\sigma_\alpha = (3\pi\sigma_T/8)^{1/2} f\lambda_0$, where $\sigma_T = 6.625 \times 10^{25}$ cm$^{-2}$ is the Thomson cross-section, $f$ is the transition oscillator strength, and $\lambda_0$ is the rest-frame transition wavelength.

According to the Voigt profile line shape, neighbouring pixels $q$ will suffer absorption from pixel $i$ by an amount $e^{-\tau_q}$, where:

$$\tau_q = \tau_{0,i} \cdot \exp\left\{ -\left[ \frac{v_q - v_i}{b_i} \right]^2 \right\} \qquad (A7)$$

The number of pixels $q$ that will suffer absorption from pixel $i$ is set to be equal to $3b_i \Delta v_L$, where $\Delta v_L$ is the length of a pixel in velocity, so depending on temperature of pixel $i$.

Shifting from pixel to pixel, we can repeat the same operation along the LOS, summing each new contribution to the optical depth in a pixel to what was previously computed a number $l$ of times, depending on the temperature of neighbouring pixels. In this way, the optical depth in one pixel is the sum of its own optical depth plus the contribution from $l$ pixels, where $l$ depends on their temperature and density.

We can also construct $\tau$-weighted density or temperature profiles. Focusing only on density, we define:

$$\rho_{\tau,q} = \frac{\sum_l \rho_i(l) \cdot \tau_i(l)}{\sum_l \tau_i(l)} \qquad (A8)$$

where $\rho_{\tau,q}$ is the $\tau$-weighted density in a given pixel $q$, $\rho_i(l)$ is the density of an adjacent pixel $i$, considering only its contribution and not by its own adjacent pixels, and $\tau_i(l)$ the optical depth contribution to pixel $q$ by pixel $i$ given by equation (A7), that is the contribution to pixel $q$ is given by the wings of the Gaussian profile whose central contribution is given by pixel $i$. The sum is over a number $l$ of pixels, as defined as before.

## APPENDIX B: COLUMN DENSITIES OF SIMULATED ABSORPTION SYSTEMS

### B1 Comparison among different methods to derive the column density

In the regime of high-resolution spectroscopy, observed parameters of absorption systems, such as the $b$-parameter or the column

---

[3]http://www.nublado.org/
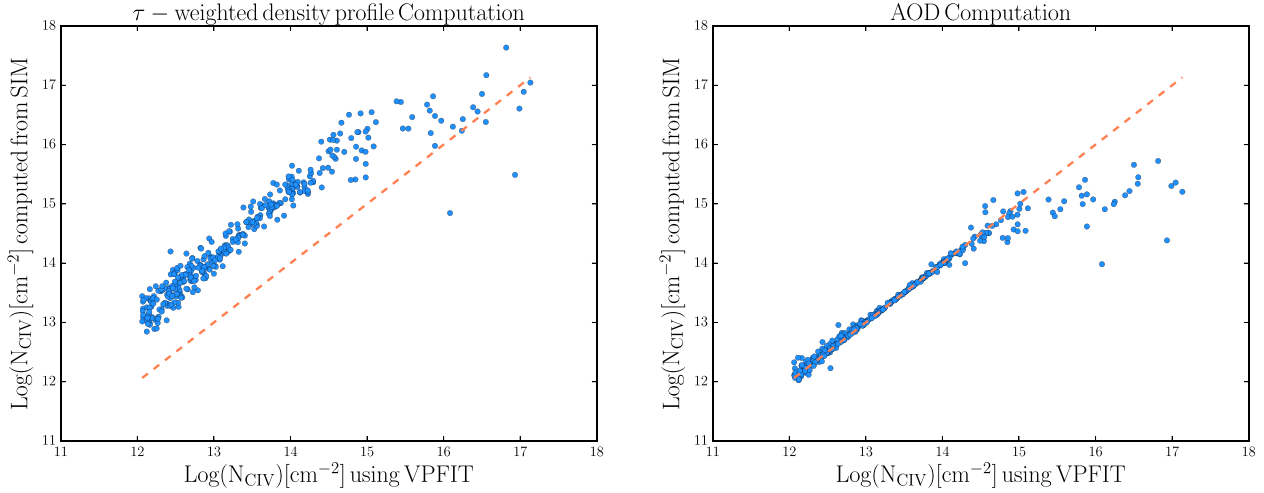[4]The UV backgrounds could be obtained from http://www.ucolick.org/pm adau/CUBA/HOME.html

**Figure B1.** Left: comparison between column densities of simulated absorption systems derived with the VPFIT code and those derived by integrating the τ-weighted density profile of the same spectra and in the velocity range ± 150 km s$^{-1}$ from VPFIT mean group redshift. The orange line is the 1:1 ratio. A disagreement of one order of magnitude is observed. Right: comparison between column densities of simulated absorption systems derived with the VPFIT code and those of the same spectra derived using the AOD method. The two values are the same. In both panels, the disagreement at high column densities (log(N$_{C IV}$/cm$^{-2}$) ≳ 15) is due to the fact that when lines become saturated, VPFIT is no more reliable in calculating column densities without higher order transitions.
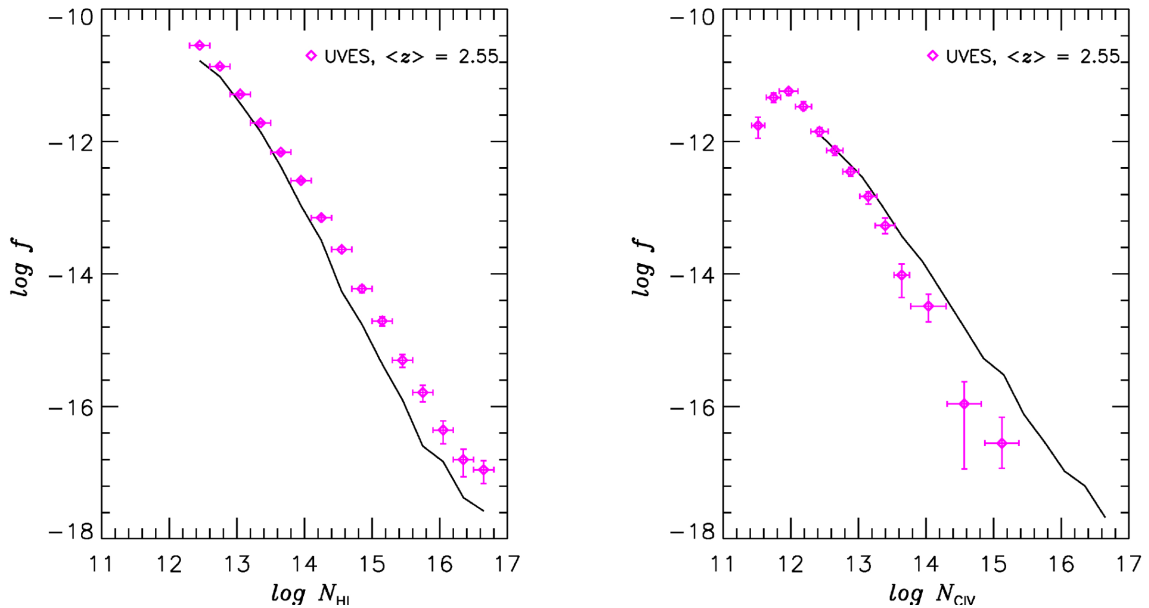


**Figure B2.** Left-hand panel: HI CDDF of the absorption systems in the lines of sight of our MUPPI simulation. Right-hand panel: C IV CDDF of absorption systems in the lines of sight of our MUPPI simulation. Observed H I and C IV CDDFs (magenta points with error bars) are both taken from Kim et al. (2013).

density, can be derived using Voigt Profile Fitting codes, such as VPFIT[5] (Carswell & Webb 2014). Observed absorption lines contain the information of gas peculiar motions and thermal broadening, so, in order to compare with them, it is important to take into account redshift space distortions also in the simulation. Using these fitting codes when dealing with simulated spectra is a delicate procedure and difficult to implement in an automatic way for thousands of spectra. For this reason, we searched for alternative methods to compute column densities of simulated absorption systems, suitable for comparison with our sample of observational data.

Since we cannot just integrate the real gas density profile along the simulated lines of sight, because in this way we neglect gas peculiar motions, we considered two different methods that we compared with the VPFIT output:

(i) the integration along the LOS of of sight LOSofthe optical depth weighted density profile;
(ii) the AOD the Apparent Optical Depth (AOD) AODmethod 1991)method method.

We first computed our sample of simulated VPFIT outputs to compare with the two methods. We therefore constructed simulated C IV spectra by piercing through the cosmological box 1000 random lines of sight, which we post-processed by adding instrumental noise (to
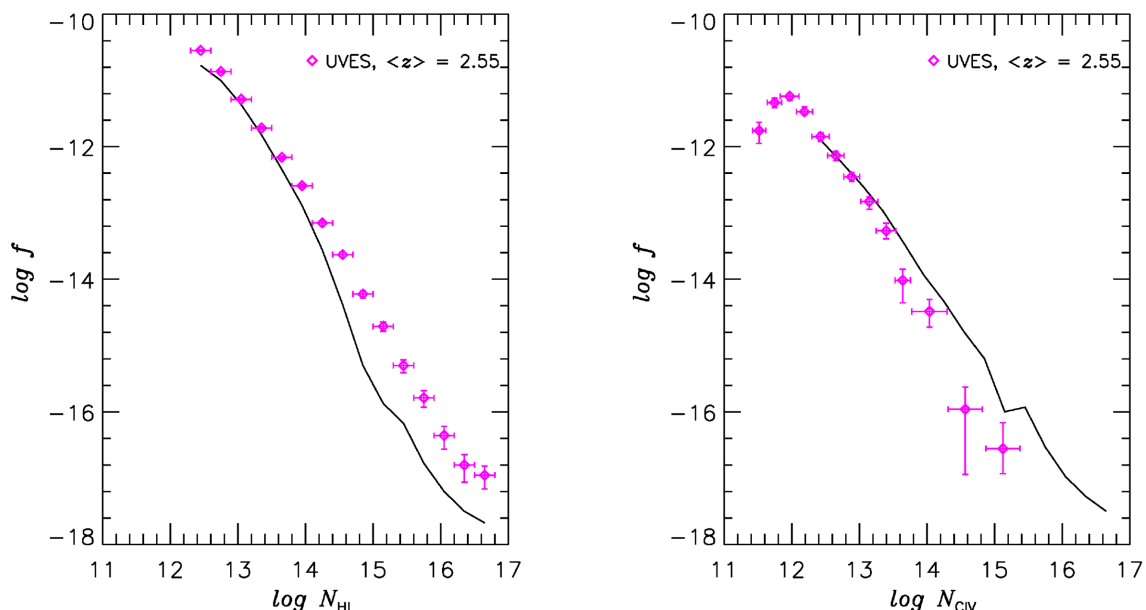
**Figure B3.** Same as Fig. B2, but for the Effective model.

reach S/N = 100) and rebinning to the spectral resolution of the observational sample of K16 ($\Delta v = 6.7\,\mathrm{km\,s^{-1}}$). We also rescaled all the fluxes by a factor $\approx 0.44$ in order to match the effective optical depth by Kim et al. (2007). We chose to perform the comparison with the C IV, as H I absorption lines are more easily saturated and VPFIT is no longer reliable in estimating column densities without higher order transitions than Ly$\alpha$, which are not present in our simulated lines of sight.

The produced spectra were then fitted by automatized VPFIT. We discarded all the spectra having fitted single components with column densities smaller than $\log(\mathrm{N/cm^{-2}}) = 12.06$ and errors on $b$ and $N$ greater than 50 per cent. The first condition is due to the fact that many spectra show no C IV absorption, but VPFIT attempts to fit the spectra anyway by putting a small component. The threshold of 12.06 is our chosen value to discard all false detections. 369 spectra satisfy these conditions.

We then constructed absorption systems trying to be as consistent as possible with the procedure by K16. With an algorithm similar to the 'FOF', we identify an absorption system as a group of single VPFIT components that lie within $\pm$ 150 km s of each other along the $^{-1}$LOS. The condition 'any friend of a friend is a friend' is valid, that is, starting from one component, we identified all the components lying within $\pm$ 150 km s$^{-1}$ of this first one. Then, we moved to the second closest component and we repeated the operation. If new friends (new with respect to the first considered component) are identified, they are added to form a unique system with the first one, and so forth with all other components satisfying this algorithm. This scheme was chosen in order to translate in an automatic way the procedure adopted by K16.

The column density of the group is the sum of the column densities of all the single components that form the system. We then determined a mean redshift of the group by weighting the redshifts of single components with their column density. Then, for every C IV system, we considered its mean redshift and we defined the velocity range of integration by extending $\pm$ 150 km s$^{-1}$ from this mean redshift.

At this point, we performed a comparison between column densities of simulated absorption systems derived with the VPFIT code and the ones derived with the other two methods.

Schaye et al. (1999) showed that absorption line minima correspond to peaks in the $\tau$-weighted density profile and the $\tau$-weighted density measured in the line centre is proportional to the column density measured by Voigt Profile Fitting. However, this does not imply that integrating this quantity along the LOS, it is possible to recover the correct column density of the absorption features, as we demonstrate in Fig. B1 (left-hand panel).

In this comparison, we took the 369 simulated lines of sight in our sample, not post-processed for VPFIT, and we computed the column density by integrating in the previously defined velocity range the $\tau$-weighted density profile. These values are shown in the $y$-axis of Fig. B1 (left-hand panel), while the VPFIT total column density of simulated absorption systems, as defined above, is shown on the $x$-axis.

The orange line is the 1:1 ratio. The plot shows clearly that there is a disagreement of one order of magnitude between the two determinations. This is due to the fact that the process of $\tau$-weighting does not preserve the integral of the density function along the LOS.

We then tested the AOD method. This method is often used dealing with observational data, because it provides a quick and reliable measure of the column density of unsaturated lines in a regime of high-resolution spectroscopy. According to this method, an apparent optical depth can be converted into an apparent column density using the formula:

$$\log[N_a(v)] = \log \tau_a(v) - \log(f\lambda) + 14.576 \qquad (B1)$$

Here, $\tau_a(v)$ is the AOD of a velocity pixel, that is the true optical depth $\tau_{\mathrm{true}}(v)$ convolved with the instrumental response, while $f$ is the transition oscillator strength and $\lambda$ the transition wavelength. In a regime of high-resolution spectroscopy, $\tau_a(v) \approx \tau_{\mathrm{true}}(v)$, so for this reason, we used the simulated true optical depth profile along the LOS not convolved with the instrumental response.

We performed the same comparison with the simulated outputs of the VPFIT code: for every spectrum, we integrated column densities derived with equation (B1) in the velocity range defined by the mean redshift of the VPFIT system of the same spectra, as explained before. In this case, before the computation, we rescaled all the optical depth values in each pixel by the same factor $\approx 0.44$, in order to be consistent with VPFIT. The comparison is shown in the right-hand panel of Fig. B1. The AOD method predicts the same column densities as VPFIT.

From an observational point of view, the AOD method does not give reliable column densities for saturated lines, because, in this case it is impossible to recover the real optical depth that produced a flux equal to zero. On the other hand, in the simulations, we know the value of the true optical depth in each pixel, thus this method always gives the correct result. The disagreement at high column densities ($\log(N_{C_{IV}}/cm^{-2}) \gtrsim 15$) shown in Fig. B1, is due to the difficulty of VPFIT to derive reliable column densities for saturated lines as already described above. On the base of the carried out tests, in our study we have adopted the AOD method for the computation of column densities in the simulated spectra.

### B2 Comparison of simulated and observed cddf

We report here the results of our simulated H and C IV CDDFs for both the two models, compared with the observational ones by Kim et al. (2013).

Simulated spectra were treated as described in Appendix B1 with the same post-processing and absorption lines were fitted with the VPFIT code. The cut at $10^{12}$ cm$^{-2}$ is artificially introduced as explained in Appendix B1. In particular, the MUPPI model reproduced quite well both the C IV and H I CDDFs (up to $10^{14}$ cm$^{-2}$). The Effective model shows slight worse agreement, in particular for H I lines at column densities larger than $10^{15}$ cm$^{-2}$. In general, under-production of H I could be ascribed to several reasons, among which the large uncertainties in the column densities determinations with the VPFIT code without higher order transitions. The overproduction of simulated C IV can be instead ascribed to the fact discussed in Section 5.1 about metal production in the simulation.

## APPENDIX C: LOOKING AT O VI IN THE CGM

We report here the comparison between simulated and observed O VI absorption systems. The observational sample is taken from Muzahid et al. (2012) and they use a slightly smaller window of integration ($v_{link} = 100$ km s$^{-1}$) than this work for the construction of absorption systems.

In Fig. C1, we show the PDFs of the $N_{OVI}$ versus $N_{H_I}$ relation for the MUPPI and the Effective models, respectively, compared to the observational sample by Muzahid et al. (2012, white triangles with error bars). Simulated absorption systems are divided according to the same distance division as previous ions.

The comparison between the two models is already discussed in Section 5.3, so here we examine only the comparison with the observational sample.

Going to larger distances, we can see again that the probability distribution gradually shifts to lower values of H I and O VI column densities, as the C VI, but in this case, the simulated sample occupies a region shifted towards lower values of O VI column densities with respect to observational data, unlike C VI.

This behavior seems not to be in agreement with the evidence that the simulation produces too many metals, but these results can be ascribed to the fact that O VI could be due both by photoionization and collisional ionization, which is not treated in our simulation. In this case, the problem of ionization has a stronger importance than the metal production. Also other works find the same effect (see e.g. Hummels et al. 2013).
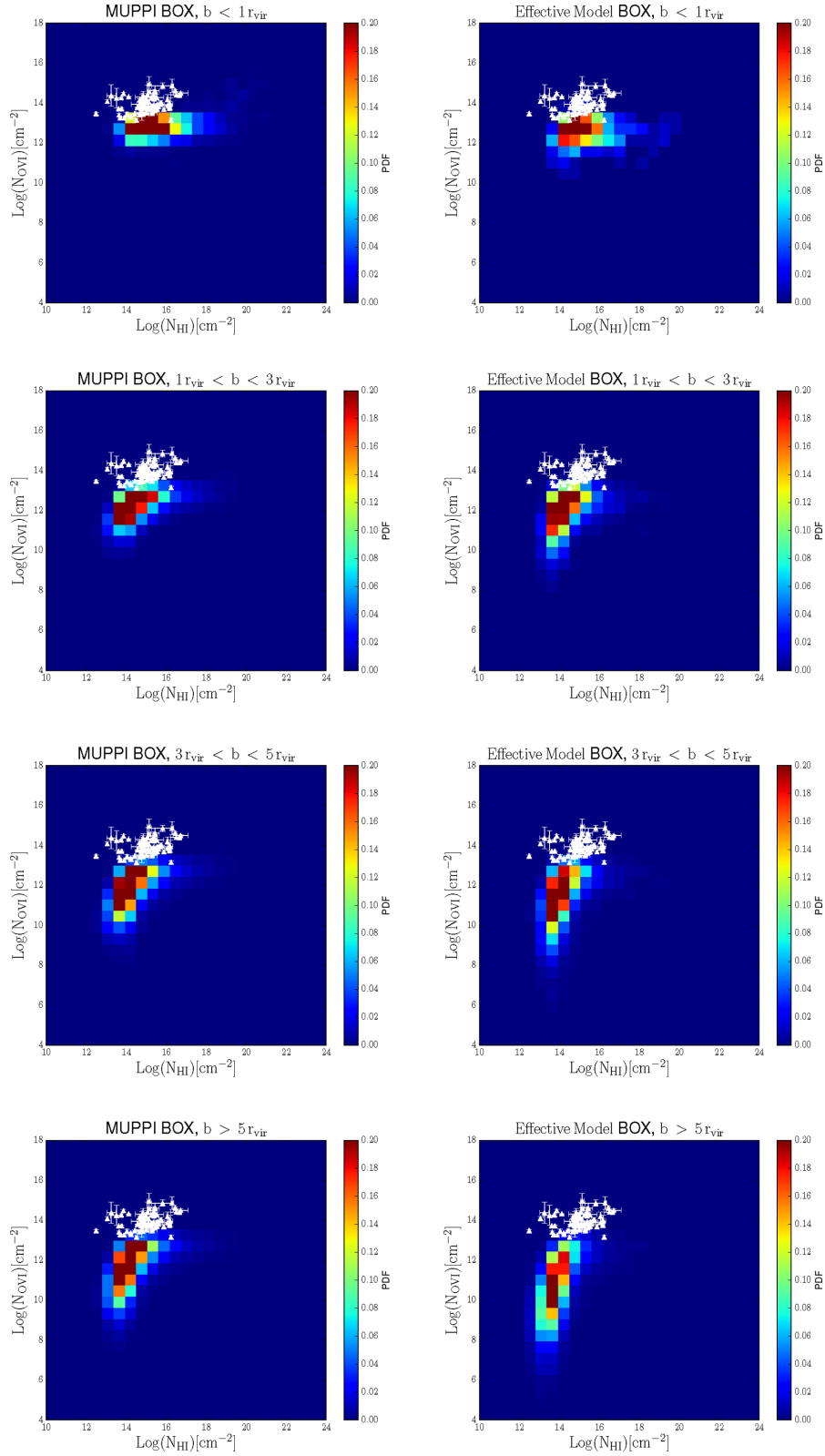
**Figure C1.** PDF of $N_{OVI}$ versus $N_{H_I}$ relation around galaxies. White triangles: observational data from Muzahid et al. (2012).

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