

# Simulating metal and H I absorption lines at the conclusion of Reionization

Luz Ángela García Peñaloza

Presented in fulfillment of the requirements  
of the degree of Doctor of Philosophy

January 2018

Centre for Astrophysics and Supercomputing  
Swinburne University of Technology



---

# Acknowledgements

As a child, I used to look at the sky at night and wonder how many stars were there. To solve the mystery, I received a little telescope and pretended I was a “real” astronomer. Today, the telescope is nothing but an old toy in my parents’ house, but I think I understand a bit better what an astronomer does.

I would like to thank my supervisors Emma, Edoardo and Stuart for offering me the opportunity to develop this project. Their knowledge and guidance have been very important throughout these 3.5 years and each one contributed in a different front to make this work successful. They taught me to think differently while approaching a new problem, a skill that I will find quite useful in Astronomy and research, but also, in other aspects of my life. Living abroad and far away from your beloved ones is tricky, but my supervisors’ advice and patience were extremely important to carry on.

I also want to acknowledge the people who helped me by sharing ideas and discussions to improve the content of my thesis: Jeff Cooke, Klaus Dolag, Valentina D’Odorico and Neil Crighton.

Special thanks to Liz, who has provided me assistance from day 1 till the end. Hard to say what would have happened if Emma and Liz would have not sorted out some situations out of my control.

To the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO) many thanks for their funding and allocating time for the project *Diagnosing Hydrogen Reionization with metal absorption line ratios* during 2015-2017.

I thank my family for their terrific support from the distance. They have always motivated me to accomplish this goal and my mother has been a lighthouse in my life. To Dany, George and Alex, thanks for their insightful comments and hints on my codes, quite messy most of the time! To Anne, many thanks for her discussions about Reionization and her advice as a young researcher. Thanks to those people who shared some ideas and plans in this period in Melbourne.

Last but not least, special thanks to Daniela Carrasco. She reinforced my endurance and patience with the years around. I am very glad for all the meaningful conversations (and the ones that were not) with coffee and biscuits that led me to a new challenge or mood.



---

# Declaration

The work presented in this thesis has been carried out in the Centre for Astrophysics & Supercomputing at Swinburne University of Technology between 2014 and 2017. This thesis contains no material that has been accepted for the award of any other degree or diploma. To the best of my knowledge, this thesis contains no material previously published or written by another author, except where due reference is made in the text of the thesis. The content of the chapters listed below has appeared in refereed journals. Minor alterations have been made to the published papers in order to maintain argument continuity and consistency of spelling and style.

- Chapter 3 has been published as “Simulated metal and HI absorption lines at the conclusion of Reionization” in *Monthly Notices of the Royal Astronomical Society*, 2017, vol. 470, issue 2, pp. 2494-2509. doi: 10.1093/mnras/stx1371. My contribution to this paper was running the simulations, setting the pipeline, analysis and writing the document. My co-authors contributed by setting up the simulations, providing continuous feedback, extended discussions and insightful ideas for the content.
- Chapter 4 has been published as “Theoretical study of an LAE–CIV absorption pair at  $z = 5.7$ ” in *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 469, issue 1, pp. L53-L57. doi: 10.1093/mnrasl/slx053. My contribution to this paper was setting the pipeline, analysis and writing the document. My co-authors contributed with continuous feedback, extended discussions and insightful ideas for the content.



*Luz Ángela García Peñaloza*  
*Melbourne, Victoria, Australia*

2018



*And yet it was only while painting that I noticed how much light  
there still was in that darkness.*

Vincent Van Gogh. The Hague, 1882.



**Dedicado a Eduard Fernando,  
por hacer parte de mi vida.**



---

# Abstract

The study of metal absorption lines in the spectra of high redshift quasars provides information on the state of the gas residing in the intergalactic medium (IGM) and the different processes that occur while the Epoch of Reionization is proceeding and concluding. A small but growing archive of data has been collected with ground based telescopes for different ionic transitions up to  $z \sim 7$ , and these detections constrain the number density, physical and chemical properties of the absorbers.

This thesis investigates the properties of baryonic gas in the redshift range of  $4 < z < 8$  in regions ionized by a uniform HM12 UV radiation field that accounts for the contribution of ionizing photons produced by galaxies and quasars. We use a set of high resolution cosmological hydrodynamical simulations based on a customized version of the GADGET3 code. The numerical runs satisfactorily describe the global star formation rate (SFR) history of the Universe and the chemical enrichment of the IGM from stars and supernovae. Our models include two different feedback prescriptions: Energy and Momentum driven winds (EDW and MDW, respectively). In addition, we implemented a model that accounts for molecular and low-temperature metal cooling.

We provide an accurate description of the evolution of observed ionic transitions that take place in the intergalactic and circumgalactic medium: CIV, CII, SiII, SiIV and OI, as well as some observables associated with HI up to  $z \leq 6$ . All our predictions have been compared with available observations in the literature, and we found that most of the simulated observables show a remarkable agreement with the observations.

In the framework of a uniform HM12 UVB, we reproduce the evolution of the cosmological mass density ( $\Omega$ ) of CII and CIV, with  $\Omega_{\text{CII}}$  exceeding  $\Omega_{\text{CIV}}$  at  $z > 6$ , consistent with the current picture of the tail of the EoR. Moreover, we predict the properties of an observed  $z = 5.72$  Lyman  $\alpha$  emitter galaxy – CIV absorption pair and give a plausible explanation for the metal enrichment of the region where the metal absorber lays. We use our models to identify the properties of faint galaxies close to the absorber, that are responsible for the observed metal signatures and provide an important contribution to the budget of ionizing photons necessary to keep the IGM ionized at  $z = 6$ . At the same redshift, the simulated CII exhibits a bimodal distribution with large absorptions in and around galaxies, and some traces in the lower density IGM. Finally, we calculate the HI

column density distribution function at  $z = 4$  and the HI cosmological mass density  $\Omega_{\text{HI}}$  at  $4 < z < 6$ .

We find some discrepancies between the observed and simulated column density relationships among different ionic species at  $z = 6$ , and propose variations of the assumed HM12 UVB and a lower metallicity content in the photoionization modelling at this redshift. In addition, we probe an alternative method to count single synthetic absorbers with their respective observations in the range defined by the detections.

We conclude that the UVB is crucial to reproduce the observed number of absorbers with our models, and it has larger impact in the observables explored than the feedback prescription or the molecular and low-temperature cooling model. The feedback wind model affects mostly the chemical enrichment history, whereas the inclusion of low-temperature metal and molecular cooling has an effect mostly in neutral Hydrogen statistics, due to the conversion of HI to H<sub>2</sub> at high densities.

In summary, we use metal absorption lines to study the state of the IGM at the tail of Reionization and its pollution with metals released by supernovae and stars. In the future, mock spectra will be improved and constrained with new observational detections of ion absorbers and better measurements of the UVB at high redshift.

**Coordinating Supervisor:** A/ Prof. Emma Ryan-Weber

**Associate Supervisors:** Dr. Edoardo Tescari, Prof. Stuart Wyithe



# Acronyms, Abbreviations and Conventions

- AGN** Active Galactic Nuclei.  
**AMR** Adaptive mesh refinement.  
**BBN** Big Bang nucleosynthesis.  
**BH** Black hole.  
**CDDF** Column density distribution function.  
**CDM** Cold dark matter.  
**CGM** Circumgalactic medium.  
**CMB** Cosmic Microwave Background.  
**DLA** Damped Lyman  $\alpha$  absorber.  
**DM** Dark matter.  
**EoR** Epoch of Reionization.  
**GP** Gunn-Peterson.  
**GRB** Gamma-ray burst.  
**IFU** Integral field unit.  
**IGM** Intergalactic medium.  
**IMF** Initial mass function.  
**ISM** Interstellar medium.  
**JWST** James Webb Space Telescope.  
**LAE** Lyman  $\alpha$  emitter.  
**LLS** Lyman limit system.  
**LSS** Large scale structure.  
**Ly $\alpha$**  Lyman  $\alpha$ .  
**PM** Particle-mesh.  
**POPII** Population II.  
**POPIII** Population III.  
**RT** Radiative transfer.  
**SF** Star formation.  
**SFR** Star formation rate.  
**SMBH** Super-massive black hole.  
**SN** Supernova.  
**SPH** Smoothed particle hydrodynamics.  
**SSh** Self-shielding.

**SSP** Simple stellar population.

**SXB** Soft X-ray background.

**QSO** Quasi-stellar object (quasar).

**UV** Ultraviolet.

**UVB** Ultraviolet background.

---

## List of ion transitions considered

---

**HI** neutral Hydrogen.

**HII** ionized Hydrogen.

**H<sub>2</sub>** molecular Hydrogen.

**CII** singly ionized Carbon.

**CIII** doubly ionized Carbon.

**CIV** triply ionized Carbon.

**CV** four times ionized Carbon.

**SiII** singly ionized Silicon.

**SiIV** triply ionized Silicon.

**OI** neutral Oxygen.

**OVI** Five times ionized Oxygen.

**FeII** singly ionized Iron.

**MgII** singly ionized Magnesium.



# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Declaration</b>	<b>ii</b>
<b>Quote</b>	<b>iv</b>
<b>Dedication</b>	<b>vi</b>
<b>Abstract</b>	<b>viii</b>
<b>Acronyms, Abbreviations and Conventions</b>	<b>xi</b>
<b>List of Figures</b>	<b>xvii</b>
<b>List of Tables</b>	<b>xxi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Early Universe . . . . .	6
1.2 The Dark Ages . . . . .	8
1.3 The first ionizing sources . . . . .	10
1.4 The Epoch of Reionization (EoR) . . . . .	13
1.4.1 Observational evidence for the EoR . . . . .	16
1.4.2 The topology of the HII regions . . . . .	18
1.4.3 UV background variations during the EoR . . . . .	19
1.5 The Universe after Reionization . . . . .	20
1.6 Purpose of the Thesis . . . . .	21
1.7 Thesis outline . . . . .	22
<b>2 Numerical Simulations and Post-process</b>	<b>25</b>
2.1 N-body simulations . . . . .	27
2.2 Galaxy formation . . . . .	29
2.3 GADGET . . . . .	32
2.3.1 Multiphase Star formation . . . . .	37
2.3.2 Chemical enrichment . . . . .	38
2.3.3 Galactic feedback . . . . .	40
2.3.4 AGN feedback . . . . .	41
2.3.5 Molecular and low-temperature metal cooling . . . . .	42

2.4	Overview of the simulations . . . . .	45
2.5	Post-process . . . . .	51
2.5.1	The uniform UV background . . . . .	51
2.5.2	Photoionization modelling with CLOUDY . . . . .	54
2.5.3	HI self-shielding . . . . .	55
2.5.4	VPFIT . . . . .	57
<b>3</b>	<b>Metal and HI absorption lines at high redshift</b>	<b>61</b>
3.1	Abstract . . . . .	61
3.2	Introduction . . . . .	61
3.3	Cosmological simulations . . . . .	64
3.3.1	Feedback mechanisms . . . . .	65
3.4	Methodology . . . . .	67
3.5	Evolution of the IGM during Reionization . . . . .	70
3.6	Evolution of Carbon . . . . .	73
3.6.1	CIV column density distribution function . . . . .	73
3.6.2	CIV cosmological mass density . . . . .	78
3.6.3	CII cosmological mass density . . . . .	78
3.6.4	Total Carbon cosmological mass density . . . . .	81
3.7	Column density relationships of metal absorption lines . . . . .	83
3.7.1	Bimodality in CII column densities . . . . .	85
3.8	Cosmological mass density of HI at redshifts lower than 6 . . . . .	88
3.9	Discussion and summary . . . . .	93
3.9.1	Effect of wind feedback . . . . .	94
3.9.2	Influence of the low-temperature metal and molecular cooling . . . . .	94
3.9.3	Limitations . . . . .	95
3.10	Conclusions . . . . .	96
<b>4</b>	<b>Theoretical study of an LAE-CIV absorption pair at <math>z = 5.7</math></b>	<b>99</b>
4.1	Abstract . . . . .	99
4.2	Introduction . . . . .	99
4.3	The LAE – CIV absorption pair of Díaz et al. 2015 . . . . .	101
4.4	Galaxy – absorber connection . . . . .	106
4.5	Conclusions . . . . .	109

---

<b>5</b>	<b>Future directions</b>	<b>111</b>
5.1	Abstract . . . . .	111
5.2	CIII evolution compared with CII and CIV . . . . .	111
5.3	Variations of the UVB in post-process . . . . .	113
5.3.1	Change in the normalization of HM12 . . . . .	116
5.3.2	Influence of the high energy end in the UVB . . . . .	126
5.4	Variation of the metallicity in the photoionization modelling . . . . .	131
5.5	Perspectives and conclusions . . . . .	137
<b>6</b>	<b>Conclusions</b>	<b>141</b>
	<b>Bibliography</b>	<b>166</b>
<b>A</b>	<b>Appendix A</b>	<b>167</b>
	<b>Publications</b>	<b>172</b>



# List of Figures

1.1	Thermal history of the Universe. . . . .	6
2.1	Schematic merger tree for a dark matter halo in a hierarchical clustering model. Taken from Baugh (2006). . . . .	29
2.2	Integration time scheme used in GADGET to evolve the equations of motion in the system. Taken from <a href="http://cvarin.github.io/CSci-Survival-Guide/leapfrog.html">http://cvarin.github.io/CSci-Survival-Guide/leapfrog.html</a> . . . . .	36
2.3	Evolution of the cosmic star formation rate density at $4 < z < 8$ . We sum the SFR in each gas particle in the simulation and divide by the comoving volume. . . . .	46
2.4	Evolution of the cosmic star formation rate density at $4 < z < 8$ . In this case a limit on the absolute magnitude of $M_{UV} < -17$ is imposed to calculate the SFR. . . . .	47
2.5	Simulated galaxy stellar mass function at $z = 8$ . . . . .	48
2.6	Simulated galaxy stellar mass function at $z = 6$ . . . . .	48
2.7	Evolution of the total Oxygen cosmological mass density $\Omega_O$ for our simulations. . . . .	49
2.8	Evolution of the total Silicon cosmological mass density $\Omega_{Si}$ for our simulations. . . . .	50
3.1	Density–temperature diagram for gas particles at redshifts $z = 8, 6$ and $4$ for the fiducial model Ch 18 512 MDW. . . . .	71
3.2	CIV column density distribution function at $z = 4.8$ and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. . . . .	74
3.3	CIV column density distribution function at $z = 5.6$ and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. . . . .	75
3.4	CIV column density distribution function at $z = 6.4$ and comparison with observational data by Bosman et al. (2017) in orange diamonds. . . . .	77
3.5	CIV cosmological mass density at $4 < z < 8$ . . . . .	79
3.6	Evolution of the CII and CIV cosmological mass density. . . . .	80
3.7	Evolution of the total Carbon cosmological mass density for our simulations. . . . .	82
3.8	Column density relationships among metal absorption lines at $z = 6$ . . . . .	84
3.9	CII vs CIV column densities in different environments at $z = 6$ for the reference simulation, Ch 18 512 MDW. . . . .	87
3.10	HI column density distribution function at $z = 4$ in the range $12 < \log N_{HI} (\text{cm}^{-2}) < 22$ for all the simulations. . . . .	90

3.11	Same as Figure 3.10 but zoomed in the DLA range: $20.3 < \log N_{\text{HI}} (\text{cm}^{-2}) < 22$ .	91
3.12	Cosmological mass density of HI.	92
4.1	Theoretical example of the observed LAE – CIV absorption pair.	105
4.2	Distribution of galaxies around the CIV strong absorption shown in Figure 4.1 up to a 3D distance of 1500 ckpc/h.	106
4.3	Mass distribution of the 3D distance from the CIV absorptions to the closest galaxy.	108
4.4	Distribution of the 3D distance from the CIV absorptions to the closest galaxy with respect to the column density of the absorption.	109
5.1	Photoionization modelling of Carbon at $z = 6$ in the reference simulation.	112
5.2	Evolution of the CII, CIII and CIV cosmological mass densities in the reference simulation.	114
5.3	UV emissivity vs. energy for the uniform HM12 background at three different redshifts: $z = 8, 6$ and $4$ (blue, dark red and green, respectively).	115
5.4	UV emissivity vs. wavelength for the uniform HM12 background at three different redshifts: $z = 8, 6$ and $4$ (blue, dark red and green, respectively) in the wavelength range where the ions transitions occur.	116
5.5	UV emissivity vs. energy at $z = 6$ . The dark red curve shows the uniform HM12 and the grey dashed line a variation of - 1 dex in the overall normalization of the spectrum ( $\log \mathbf{J}_\nu - 1$ ).	117
5.6	CIV column density distribution function at $z = 4.8$ and comparison with observational data by D’Odorico et al. (2013) in orange diamonds.	118
5.7	CIV column density distribution function at $z = 5.6$ and comparison with observational data by D’Odorico et al. (2013) in orange diamonds.	118
5.8	CIV column density distribution function at $z = 6.4$ and comparison with observational data by Bosman et al. (2017) in orange diamonds.	119
5.9	CIV cosmological mass density at $4 < z < 8$ for $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ .	119
5.10	Evolution of the CII and CIV cosmological mass density when the normalization of the UVB is varied at 1 Ryd (comparison of molecular cooling content).	120
5.11	Evolution of the CII and CIV cosmological mass density when the normalization of the UVB is varied at 1 Ryd (comparison of feedback prescriptions).	121

5.12	Number of absorbers at $z = 6$ when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd compared with the observations. . . .	125
5.13	UV emissivity vs. energy at $z = 6$ . The dark red curve shows the uniform HM12 and the blue dashed line a very aggressive contribution from quasars ( <b>qua test</b> ). . . . .	127
5.14	Number of absorbers at $z = 6$ with a modification of the high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. The simulation run used is Ch 18 512 MDW. . . .	130
5.15	Number of absorbers at $z = 6$ obtained by varying the metallicity in the CLOUDY modelling compared with the observations. . . . .	135
5.16	Ionization fractions of CII (left panel) and CIV (right panel) as a function of the gas overdensity. Results from CLOUDY modelling at $z = 6$ comparing gas temperatures of $T = 10^{4.6}$ K (solid lines) and $T = 10^{5.0}$ K (dashed lines).136	
5.17	Ionization fractions of CII (left panel) and CIV (right panel) as a function of the gas temperature. Results from CLOUDY modelling at $z = 6$ for gas overdensity of $\Delta \sim 10^{3.6}$ . . . . .	137
A.1	Column density relationships among metal absorption lines at $z = 6$ when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd. The simulation run used is Ch 18 512 MDW. . . . .	168
A.2	Column density relationships among metal absorption lines at $z = 6$ when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd. The simulation run used is Ch 18 512 MDW mol. . . . .	169
A.3	Column density relationships among metal absorption lines at $z = 6$ when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd. The simulation run used is Ch 18 512 EDW. . . . .	170
A.4	Column density relationships among metal absorption lines at $z = 6$ with a modification of the high energy end in the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift ( <b>qua test</b> ). The simulation run used is Ch 18 512 MDW. . . . .	171



# List of Tables

2.1	Reaction network taken from Maio & Tescari (2015). . . . .	44
2.2	Summary of the simulations used in this work. . . . .	45
2.3	The cosmic background photoionization $\Gamma_{\text{HI}}$ and photoheating $\mathcal{H}_{\text{HI}}$ rates in HM12. . . . .	53
2.4	List of the ion lines included in this work and reduced with VPFIT. . . . .	59
3.1	Summary of the simulations used in this work. . . . .	66
3.2	Contribution from DLA systems to $\Omega_{\text{HI}}$ . . . . .	93
4.1	Wind velocity ( $v_w$ ) and travel time ( $t$ ) to enrich a region at 1300 ckpc/ $h$ from our most massive galaxy ( $M_h$ ). . . . .	102
5.1	Number of absorbers at $z = 6$ when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd compared with the observations. . . . .	124
5.2	CII and CIV cosmological mass densities at $z = 6$ with a modification at high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. . . . .	128
5.3	Number of absorbers at at $z = 6$ with a modification of the high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. The simulation run used is Ch 18 512 MDW. . . . .	129
5.4	CII and CIV cosmological mass densities at $z = 6$ obtained by varying the metallicity in the CLOUDY modelling. . . . .	132
5.5	Number of absorbers obtained by varying the metallicity in the CLOUDY modelling compared with the observations. . . . .	134



# 1

## Introduction

*“I get wisdom day and night, turning darkness into light.”*

St. Paul Irish Codex

### **The cosmological standard model**

---

The current knowledge in Physics allow us to look backwards in the history of the Universe right until the Planck time (when the Universe is  $10^{-44}$  s old). However, without a complete theory of Quantum Gravity, any event before that time is beyond our reach. Yet, the astounding observations made in the 20<sup>th</sup> century along with the General Relativity theory have led us to a model that describes well the Universe and its evolution, the  **$\Lambda$ CDM Cosmological Standard model**.

The discovery by Edwin Hubble in 1929 that distant galaxies are receding from us with a velocity that increases with distance is the first empirical evidence that the Universe is expanding and the matter in the Universe is diluting with time. Hubble proposed an empirical law to describe his observations with the linear relation  $v = HD$ , where  $v$  is the recessional speed,  $H$  the Hubble parameter (that depends on time) and  $D$  the proper distance (that changes with time) from the galaxy to the observer. In this order of ideas, it is natural to ask about the evolution of galaxies in the past, and how everything was formed in the beginning. By reversing Hubble’s main conclusion, it is possible to infer that the Universe was hotter and denser at early times.

The best way to describe an expanding Universe is through a comoving system of coordinates. The expansion is introduced with a scale factor  $a(t)$ , that increases uniformly with time (normalized to  $a_0 \equiv 1$ , today)<sup>1</sup>. Thus, the proper (physical) coordinate  $r(t)$  of an

---

<sup>1</sup>By convention, in Cosmology the sub index 0 is used to describe observables today.

object in the Hubble flow is defined in terms of the scale factor  $a(t)$  and the comoving coordinate  $x$ :

$$r(t) \equiv a(t)x, \quad (1.1)$$

where the comoving distance is invariant in time in the reference system that represents the expanding Universe. The evolution of the scale factor in time can be quantified by the Hubble rate (or Hubble parameter mentioned above):

$$H(t) \equiv \frac{da/dt}{a}, \quad (1.2)$$

The Einstein equations describe the interaction of gravity and matter–energy through geometry. Their solution for an expanding Universe takes into account the **Cosmological principle**, the assumption that the Universe is homogeneous and isotropic on large scales (up to  $\sim 500$  Mpc). As a consequence, the spatial components should evolve with the scale factor  $a(t)$  to describe the expansion given in eq. (1.1). The formal solution is known as the Friedmann-Lemaitre-Robertson-Walker metric:

$$ds^2 = c^2 dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right], \quad (1.3)$$

being  $c$  the speed of light,  $k$  the spatial curvature (-1,0,1) for a hyperbolic, flat or spherical Universe, respectively. Introducing the elements of the metric (1.3) in the Einstein equations for a perfect fluid (with density  $\rho$  and pressure  $p$ ), the evolution of the Hubble rate is given by:

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{a^2(t)}, \quad (1.4)$$

where  $\rho(t)$  is the mass density of all the matter–energy components in the Universe as a function of time. The critical mass density  $\rho_{\text{cr}}$  is the density today for which the spatial geometry of the Universe is flat ( $k = 0$ ),

$$\rho_{\text{cr}} \equiv \frac{3H_0^2}{8\pi G}, \quad (1.5)$$

here,  $H_0$  is the current value of the Hubble parameter. Measures of the present Hubble rate value are usually parametrized in terms of  $h$ , such that:

$$H_0 = 100h \text{ km / s / Mpc}, \quad (1.6)$$

and  $h = 0.6674$  (Planck Collaboration et al., 2015)

The solutions of the Einstein equations describe non-interacting perfect fluids<sup>2</sup>. In the equation (1.4)  $\rho$  includes any component of matter–energy in the Universe and each component can be studied as a perfect fluid with an equation of state:

$$p = w\rho c^2, \quad (1.7)$$

where  $w$  is the parameter from the state equation, while  $p$  and  $\rho$  are the pressure and density of the fluid involved, respectively. The evolution of the density of the fluid is given by (eq. 1.1.34 from Weinberg, 2008):

$$\rho \propto a^{-3-3w}. \quad (1.8)$$

However, it is more convenient to define a quantity to weight the density  $\rho$  in eq. (1.8) with respect to the critical density today, as  $\Omega = \frac{\rho}{\rho_{\text{cr}}}$ , the density parameter.

The main components in the Universe that contribute to the total density in equation (1.4) are:

- Radiation ( $w = \frac{1}{3}$ ): massless particles with positive pressure and density  $\rho \propto a^{-4}$ , as photons and massless neutrinos. The density associated with the radiation is expressed in terms of the density parameter  $\Omega_r$  (the sub index  $r$  stands for radiation), the critical density of the Universe and the dependency with  $a$  determined by (1.8):

$$\rho_r = \Omega_r \rho_{\text{cr}} a^{-4} \quad (1.9)$$

- Matter: particles with mass  $m > 0$ . Hence, the fluid has a positive density ( $\rho \propto a^{-3}$ ) and null-pressure ( $w = 0$ ). The matter density in this case is given by:

$$\rho_m = \Omega_m \rho_{\text{cr}} a^{-3} \quad (1.10)$$

- Baryonic (ordinary) matter: everything in the visible Universe is made by atoms and responds to the electromagnetic interaction. The radiation–matter interaction allows us to collect photons emitted by distant objects with our telescopes.
- Dark matter (DM): There is another form of matter, not detected yet, but modelled with very massive particles. The distinctive characteristic of this

---

<sup>2</sup>If the fluids interchange energy, an additional term should be included in the Einstein equation to account for such energy exchange.

type of matter is that it is completely transparent to light, and indeed any electromagnetic radiation.

Its existence was proposed to solve a problem of the seventies: the rotation curve of disks galaxies. If a galaxy is formed only by baryonic mass, it is expected that the circular velocity of the stars in the disk plane decreases rapidly while moving towards the outskirts of the galaxy. That was not what was observed. Instead, the velocity plateaus at a given radius from the centre of the galaxy. This result indicates that there is an excess of mass in the galaxy, hidden from our instruments. There are now several other indications about the existence of DM. This type of matter can be traced through indirect methods (like gravitational lensing or experiments of dark matter annihilation with gamma-rays production).

The most favoured candidates for dark matter are: sterile neutrinos, axions, supersymmetric particles or WIMPs (weak interacting massive particles), but all of them require energies of the order of GeV – TeV mass to be detected in the large accelerators (not achieved yet).

These matter components satisfy the constraint  $\Omega_m = \Omega_b + \Omega_{\text{DM}}$ .

- **Dark energy:** In 1999, different groups found independently an unexpected result. Using distant type Ia supernovae, they discovered that the Universe is speeding up its expansion in the late times. However, any kind of matter has an attractive gravitational charge, therefore an accelerated expansion has to be caused by a completely different component with negative pressure, such that  $p = -\rho$ .

Although astronomers know the effect of this fluid, there is not a clear idea of how to detect it, mainly because it is a smooth component, dilute throughout all the Universe and the parameter of the equation of state today is most likely  $w_0 = -1$ , even if  $w = w(t)$  in the past. Different models have been proposed in the past 15 years to explain the nature of this component: the cosmological constant  $\Lambda$  that accounts for the quantum vacuum energy, scalar fields with different  $w(t)$  or instead, modified gravity models that effectively impose the accelerated expansion through a geometrical contribution, rather than an energy density.

In the context of the standard model, the cosmological constant  $\Lambda$  is assumed to be responsible for the accelerated expansion, with an associated density parameter  $\Omega_\Lambda = -\frac{\Lambda c^2}{3H_0^2}$ .

Moreover, if the spatial curvature satisfies  $k \neq 0$ , a density parameter can be associated

with the curvature term  $\Omega_k = \frac{3H_0^2}{8\pi G}kc^2$ .

The Hubble parameter as a function of the scale factor  $a(t)$  and the different components discussed above is given by:

$$H^2 = H_0^2 (\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_K a^{-2} + \Omega_\Lambda). \quad (1.11)$$

In the context of the Concordance model,  $\Omega_{r,0} + \Omega_{m,0} + \Omega_{K,0} + \Omega_{\Lambda,0} = 1$  today. However, the contribution from the radiation at this time is completely negligible, since  $\frac{\Omega_{r,0}}{\Omega_{m,0}} \sim 10^{-5}$ . On the other hand, the spatial curvature of the Universe does not play a role as a source of matter–energy from the Einstein equations. The density parameter from the curvature  $\Omega_k \approx 0$  ( $0.0008_{-0.0039}^{+0.0040}$  from the inferred parameters of Planck Collaboration et al. 2015), therefore the spatial curvature of the Universe can be considered flat.

Furthermore, there is great confidence in the measurements of the temperature of the cosmic microwave background (CMB) and its anisotropies with three different missions: COBE (Cosmic Background Explorer), WMAP (Wilkinson Microwave Anisotropy Probe) and PLANCK. The temperature (today) of the CMB photons is  $T_0 = 2.728 \pm 0.004$  K (Fixsen et al., 1996; Noterdaeme et al., 2011).

A photon with energy  $E$  has a corresponding wavelength  $\lambda$ . In the past, as the scale factor was shorter, so was the wavelength of the signal. The energy of the photon evolves as  $E \propto T$ , but also  $E \propto \lambda^{-1} \propto a^{-1}$ , therefore  $T \propto a^{-1}$  (*i.e.* as the Universe expands, the temperature of the plasma where the photons are embedded decreases). This can be expressed as:

$$T(t) = \frac{T_0}{a(t)}. \quad (1.12)$$

In addition, the expansion of the Universe has an impact on light signals sent from receding galaxies. A wave emitted at a certain distance will be stretched out and its measured wavelength will be larger  $\lambda_{\text{obs}} > \lambda_{\text{emi}}$ . Consequently, the farther away the signal is emitted, the lower its detected frequency (or the longer its wavelength). It is a convention to define the stretching factor as  $z$ , the redshift:

$$1 + z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{emi}}} = \frac{1}{a}. \quad (1.13)$$

For  $z \ll 1$ , the classical Doppler effect applies and  $z \sim \frac{v}{c}$ , being  $v$  the velocity of the source and  $c$  the speed of light.

Since the first measurement of the redshift of spiral nebulae in 1910, this quantity has been

extensively used in Astronomy to quantify distances, but also, as an indirect indicator of time. The refinement of telescopes in the past century allows us to measure with high precision the wavelength of signals emitted from very far distances.

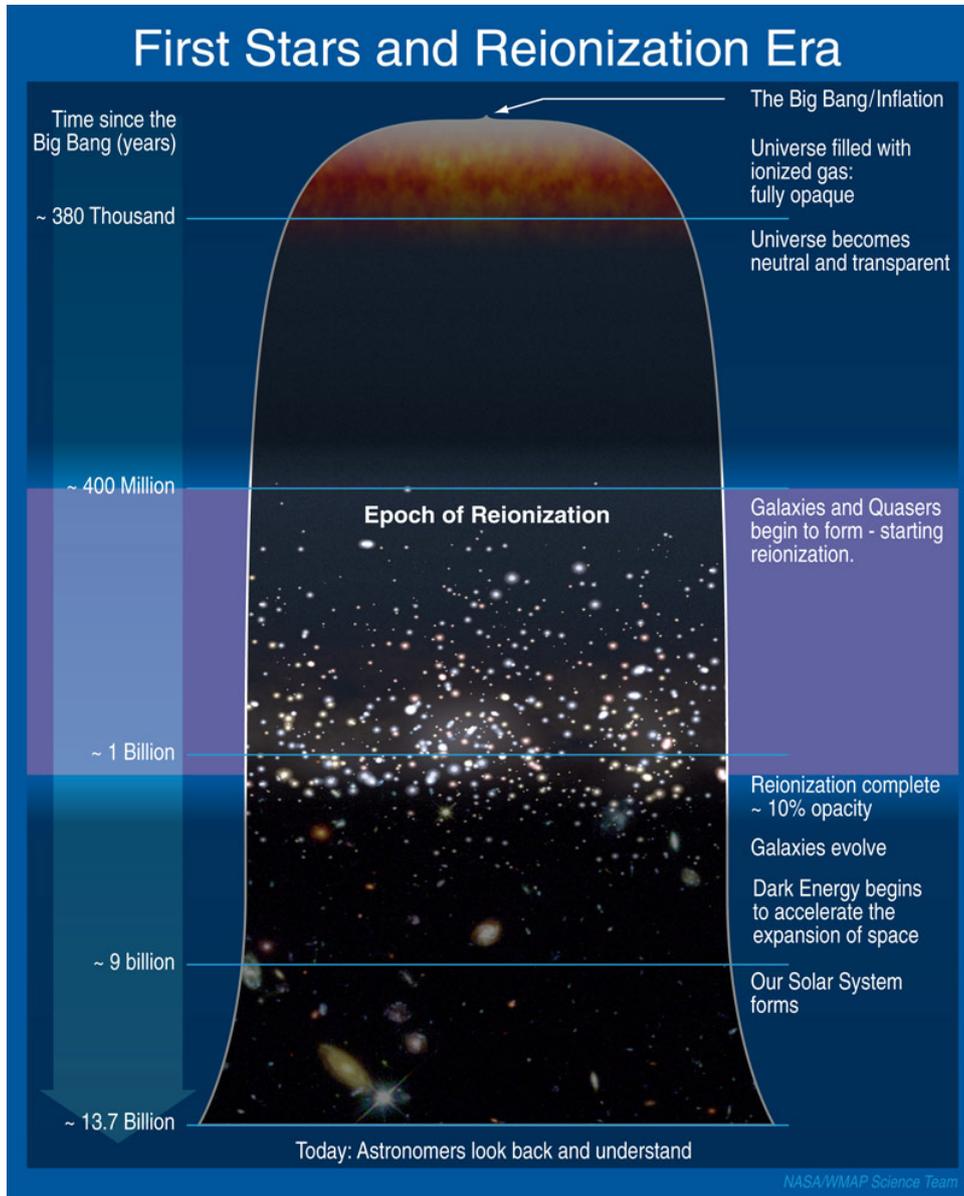


Figure 1.1 Thermal history of the Universe.

## 1.1 The Early Universe

The first minutes of the Universe were very turbulent. At a short  $10^{-35}$  s, the Universe went through an inflationary period and expanded more than  $e^{60}$  times in size. This model,

called *Inflation*, was formulated by Andrei Linde (Linde, 1980a,b,c,d) to solve three fundamental issues in the cosmological standard model in the eighties: the horizon, flatness and magnetic-monopole problems. Inflation also explains how the initial conditions of the matter-energy were set up, placing the seeds for the density anisotropies that would lead to the structure formation later on. Once Inflation concluded, the kinetic energy is transferred to the hot plasma. The matter-antimatter asymmetry left enough matter to allow baryogenesis and leptogenesis processes to form the fundamental particles in nature: nucleons, and, electrons and neutrinos, respectively. With these bricks of matter in the hot plasma, Big Bang Nucleosynthesis started: the nuclei of the light elements (Hydrogen, Deuterium, Helium and some traces of the nuclei of Be, Li, C, O, N, etc.) were formed from the free nucleons. Simultaneously, high energy photons (whose wavelength stretches with the expansion of the Universe) interacted via Compton scattering with free electrons. As the Universe expanded, it cooled, then the thermal equilibrium between the matter and radiation (photons and neutrinos) fell out and the Compton interaction among them was less frequent due to the decreasing energy of the radiation. The decoupling between them occurred and left its imprint in the cosmic microwave background. At this time, the mean free path of the photons was close to the horizon size of the Universe.

At  $z = 1100$  (the Universe is 380000 years old), the electromagnetic interaction between free electrons and protons was favoured and they were tightly coupled via Coulomb scattering. Photons were not energetic enough to photo-dissociate the Hydrogen atoms. This process is known as Recombination and led to the formation of neutral Hydrogen atoms from the bonding of electrons and protons in thermal equilibrium (Dodelson, 2003).

Overdensities in dark and baryonic matter grew from the primordial density fluctuations set down during Inflation. The dark matter began to collapse into halos, followed by gas, which cooled and set up the temperature and pressure to form the first stars by accretion and subsequently galaxies. This period of the Universe is known as the Dark Ages and there is no observational evidence of the process, but recent detections of early galaxies at redshifts  $z \sim 9-11$  show that galaxies exist at this time and we expect to detect more with future telescopes, such as JWST.

The Epoch of Reionization of Hydrogen (referred to “Reionization” here in), as main era of investigation of this thesis, took place from redshift 15 to 6 and left an imprint in the diffuse gas in between galaxies (the intergalactic medium or IGM). The neutral Hydrogen (HI) formed in the Recombination epoch was turned into its ionized state (HII) due to the energetic UV ionizing fronts produced by the first stars, galaxies, quasars and active galactic nuclei (AGNs).

Once the fraction of HI reached one part in ten thousand with respect to the amount of total Hydrogen ( $\frac{n_{\text{HI}}}{n_{\text{H}}} = 10^{-4}$ ), the Epoch of Reionization (EoR) has concluded and the only pockets of neutral Hydrogen in the Universe are found in high density regions, where H is self-shielded. These regions are identified as damped Lyman- $\alpha$  systems (DLAs) and they have been observed up to  $z \leq 5$ . Although the ionization state and density of the HI gas may be the same, the physical similarities of the objects associated with  $z > 5$  and  $z \leq 5$  DLAs are not clear.

## 1.2 The Dark Ages

The stage of the Universe in between the moment when the CMB was released and the first stars formed is known as the Dark Ages (Loeb, 2006, 2010).

The decoupling of the photons from the plasma at  $z = 1100$  leads to a transparent Universe for photons with  $E < 13.6$  eV (below the binding energy of the H atom). Dissociations were less and less frequent due to the loss of energy of the photons, favouring the electromagnetic interaction among electrons and protons and, as a consequence, the formation of the Hydrogen atoms. As the CMB photons are decoupled from the matter, they no longer (directly) trace the distribution of baryons, but they still follow the oscillations of the ordinary matter falling into the dark matter potential wells, as perturbations in the speed of sound in the gas. These baryonic oscillations are imprinted in the CMB map.

The only detectable signature of The Dark Ages is the radiation that was emitted by the 21 cm spin-flip line transition by neutral Hydrogen. The occasional collisions of photons with H atoms cause an alignment of the spin of proton and electron in the atom. When the electron flips back the orientation of the spin, it returns to the configuration of minimal energy and the system releases a photon with 21 cm wavelength.

The intensity of the 21 cm photons depends on the spin temperature,  $T_{\text{spin}}$ <sup>3</sup> as:

$$\frac{n_1}{n_0} = 3e^{\left(-\frac{T_*}{T_{\text{spin}}}\right)}, \quad (1.14)$$

where  $n_1$  and  $n_0$  are the number densities of electrons in the triplet and singlet states

---

<sup>3</sup>The temperature/energy gap between the upper 1 (spins aligned) and lower 0 (opposite spins) configurations in the Hydrogen atom. Although there is not an actual definition for the spin temperature, it measures the ratio of level populations 1/0.

of the hyperfine level, respectively, and  $T_* = 0.0681$  K the temperature corresponding to the 21 cm wavelength (Zaroubi, 2013). This ratio quantifies the intensity of the radiation that emerges from the excited state of two hyperfine levels in the electron of the neutral Hydrogen gas.

There are three physical processes that determine  $T_{\text{spin}}$ : the absorption of CMB photons, collisions with surrounding hydrogen atoms, electrons or free protons, and the (de)excitation due to scattering with Ly $\alpha$  photons, quantified by  $T_{\text{CMB}}$ ,  $T_{\text{kin}}$  and  $T_{\alpha}$ , respectively. These processes are summarized in the following expression (Field, 1958):

$$T_{\text{spin}} = \frac{T_{\text{CMB}} + y_{\text{kin}}T_{\text{kin}} + y_{\alpha}T_{\alpha}}{1 + y_{\text{kin}} + y_{\alpha}}, \quad (1.15)$$

with  $y_{\text{kin}}$  and  $y_{\alpha}$  the kinetic and Ly $\alpha$  coupling factors, respectively.

However, the 21 cm temperature has to be measured with respect to the CMB (background) temperature. The detection of the differential brightness temperature  $\delta T_b \equiv T_b - T_{\text{CMB}}$  in the IGM using radio signal depends on the departure of  $T_b$  (brightness temperature) from  $T_{\text{CMB}}$ , that evolves as  $1+z$ .

$$\delta T_b = 28mK(1 + \delta)x_{\text{HI}} \left(1 - \frac{T_{\text{CMB}}}{T_{\text{spin}}}\right) \left(\frac{\Omega_b h^2}{0.0223}\right) \sqrt{\left(\frac{1+z}{10}\right) \left(\frac{0.24}{\Omega_m}\right) \left[\frac{H(z)/(1+z)}{dv_{\parallel}/dr_{\parallel}}\right]}, \quad (1.16)$$

where  $\delta$  is the density contrast,  $x_{\text{HI}}$  is the fraction of neutral Hydrogen,  $\Omega_b$  and  $\Omega_m$  are the baryon and matter densities with respect to the critical density.  $H(z)$  is the Hubble parameter whereas  $h$  is the Hubble constant in units of 100 km/s/Mpc. Finally, the term  $dv_{\parallel}/dr_{\parallel}$  is the gradient of the proper velocity along the line of sight.

There are several ongoing projects devoted to measure the 21 cm radiation in different radio bands: the Low Frequency Array (LOFAR), the Murchison Widefield Array (MWA), the Precision Array to Probe Epoch of Reionization (PAPER), 21 Centimeter Array (21CMA) and the Giant Metrewave Radio Telescope (GMRT). In the future, the Square Kilometer Array (SKA) will expand and complement the observations with a larger collective area. The main goal with all these arrays is to build a 3D map of the HI distribution on the sky at different redshifts (or times) and compare this map with the location of the first galaxies to have a complete picture of the Reionization process and the sources that ionized the diffuse Hydrogen.

However, all these radio telescopes have to deal with a very faint cosmological signal, large systematic errors and contamination from the Galactic foreground, ionospheric perturba-

tions, instrument response and noise.

### 1.3 The first ionizing sources

Density fluctuations were formed during the inflationary period and grew during the matter domination epoch (Liddle & Lyth, 2000). They are the nursery of the large scale structure (LSS) that has been observed today by 2dFGRS (Cole et al., 2005), SDSS (Tegmark et al., 2006), WiggleZ (Blake et al., 2011) and BOSS (Alam et al., 2015) and reproduced in simulations up to several hundreds comoving Mpc in size as the Millennium project (Springel et al., 2005b). The distribution of baryonic matter in the early Universe produced very dense knots in the intersections of filaments that form the cosmic web, and it seems to mimic the dark matter distribution and its evolution<sup>4</sup>. These knots correspond to overdensities of ordinary matter.

Dark matter perturbations are characterized by their wavelength (mass) and they have evolved independently since Inflation. As DM started to collapse, it produced a complex network of dark matter halos. However, baryonic collapse in the DM halos took much longer because the gravitational Jeans instability is counterbalanced by the radiative pressure of the gas. The collapse is triggered in the matter domination epoch, when the threshold over-density for gravitational collapse is non-negligible (Yoshida, 2010):

$$\delta_{\text{crit}} = \frac{1.686}{D(z)}, \quad (1.17)$$

where  $D(z)$  is the linear growth factor of perturbations to  $z$ . The star formation began when a sufficient amount of cold dense gas was accumulated in a DM halo. At  $z \sim 150$ , gas starts to fall into the growing conglomerates of cold dark matter, the gas is then reheated by the energy released by two simultaneous processes: the gravitational collapse and the kinetic feedback from the first stars (Shapiro et al., 1994). The only coolant available in the gas at high redshift is molecular Hydrogen ( $\text{H}_2$ ), that once formed, changes its quantum rotational and vibrational levels, releasing a photon. At this time there are enough free electrons that catalyze molecular Hydrogen and cool off the gas to  $\sim$  hundreds of K in the interstellar medium (ISM) and eventually condense to form gas clouds (Osterbrock & Ferland, 2006).

In a gas cloud in hydrostatic equilibrium, the kinetic energy of the gas is counter-

---

<sup>4</sup>DM distribution is well described by N-body simulations.

balanced by the potential energy of the gravitation interaction, according to the Virial theorem (Dodelson, 2003). With the continuous accretion of mass, the cloud grows in size and the pressure of the gas can no longer hold the equilibrium, because the gravitational potential grows proportionally to the mass in the system, leading to the gravitational collapse of the cloud. The critical mass at which the collapse occurs is known as the Jeans mass (Jeans, 1902; Weinberg, 2008):

$$M_J = \left( \frac{5\pi K_B T_{\gamma,0}}{3\mu m G} \right)^{3/2} \rho_{m,0}^{-1/2} = 2.02 \times 10^5 (\Omega_m h^2)^{-1/2} M_\odot \quad (1.18)$$

where  $K_B$  is the Boltzmann constant,  $T_{\gamma,0}$  the temperature of the CMB photons today,  $\mu$  is the mean molecular weight of gas composed of H and He atoms,  $m$  the mass of the gas cloud,  $G$  the gravitational constant.  $\rho_{m,0}$  and  $\Omega_m$  are the mass density and density parameter of matter today and  $M_\odot$ , the mass of the Sun.

Observations of galaxies across time together with cosmological simulations have shown that structures in the Universe form hierarchically, from small structures, that grow, evolve and merge to larger systems. The first stars formed in dark matter halos that accreted metal-free baryonic matter, with just little traces of light primordial elements. This first generation of stars (or POPIII) is predicted to be formed in DM halos of mass  $\sim 10^6 M_\odot$  that collapsed at  $z \sim 20-30$  (Tegmark et al., 1997; Yoshida et al., 2003; Bromm & Loeb, 2003; Bromm, 2003). The absence of a vigorous mechanism for the gas fragmentation, the large gas mass accretion rate, as well as the lack of dust, provided favourable conditions for the formation of very massive stars in the first stages of cosmic star formation. Since POPIII were quite massive stars, their lives were very short, leading to an early explosion into the first supernovae. This cycle, that happened quickly in the story of the Universe, released enriched gas into the surroundings and triggered new stellar generations. The primordial stellar generation has not been detected yet, thus, there is large uncertainty in the initial mass function (IMF) that describes its distribution in stellar mass. On the other hand, due to the chemical composition of the subsequent generations (they were not exclusively formed out of H and He, but also, other metals built in nuclear reactions that took place in the POPIII), POPII stars are typically less massive than their predecessors.

POPIII stars seemed to form relatively low-mass stellar clusters and galaxies, due to their short life-time, comparable with the time-scale of the galaxy assembly. During the

same era, with the evolution of the first massive stars, quasars and active galactic nuclei (AGN) were formed by gravitational collapse. First quasars were likely to form in massive halo environments, at redshifts  $z \geq 10$  (e.g., Umemura et al., 1993; Loeb & Rasio, 1994; Eisenstein & Loeb, 1995; Haiman & Rees, 2001).

With the formation of the first collapsed objects, the physical scenario turns very complex, due to the interaction of stars and proto-galaxies with their environment. The feedback produced by the first stars is crucial, because the surrounding gas is heated by ionizing radiation and kinetic energy ejected via winds and supernova explosions (with an estimated maximum rate of one supernova per year at high redshift, Rees, 2000). Supernovae produce and disperse elements to the surrounding gas, and their location impacts how effective the chemical enrichment will be at a given redshift. The maximum radius of influence from a galactic outflow at high redshift is given by  $r_{\max} \sim 50 \text{ kpc} \left( \frac{E}{10^{56} \text{ erg}} \right) \left( \frac{1+z}{10} \right)^{-6/5}$  (Rees, 2000). By estimating the number of supernovae that contribute per galaxy ( $E_{\text{SN}} \sim 10^{51} \text{ erg}$ ) and accounting for the energy lost by cooling in the circumgalactic medium (CGM), it is possible to quantify the maximum distance reached by the galactic winds and thus, the efficiency of the chemical pollution in the IGM (Loeb, 2010).

Atomic and molecular cooling become more efficient when the heavy elements are introduced to the IGM by supernovae and galactic feedback. In addition, the UV ionizing photons produced by the new born stars, easily dissociate the H molecules and reheat the gas, suppressing the formation of new stars while the gas is above the temperature necessary to keep the cloud virialized.

With halos at higher temperatures than  $10^4 \text{ K}$ <sup>5</sup>, and the subsequent appearance of POPII stars, atomic Hydrogen becomes the main coolant. Metal transitions that contribute to the cooling processes at low temperature are the main channel to the cooling when the H molecules are suppressed in the ISM (Gnat & Ferland, 2012; Ferland & Williams, 2016).

At  $z \sim 15$ , there is a large number of stable self-gravitating stars distributed around the centre of dark matter halos. Gas is pulled into potential wells defined by the mass of the halos and forms galaxies (collections of stars and cooled gas with a luminous core and a surrounding dark matter halo). The assembly of these galaxies is regulated by the dynamics of the gas and the collisionless nature of DM and stars, which tend to conglomerate in spheroids. On the other hand, collisions of the gas particles transfer energy to the outskirts of the system in form of radiation and galactic winds. This feedback with

---

<sup>5</sup>The characteristic temperature for a proto-star to reach the thermodynamical equilibrium.

the CGM and IGM, cools the gas and reduces its kinetic/thermal energy. However, the angular momentum is (mostly) conserved during the gravitational collapse, leading to a distribution of gas with high angular momentum but low kinetic energy, *a.k.a* a thin disk (Mestel, 1963; Freeman, 1970; Efstathiou & Jones, 1980; Gunn, 1982; van der Kruit, 1987; van der Kruit & Freeman, 2011). The angular momentum of the mass distributed in the galaxy along a disk prevents the stars from collapsing to the centre to form a super-massive black hole.

As the Universe evolves, gas continues to flow in -via accretion- and out -via galactic winds- of the CGM. Baryonic matter just beyond the reach of accretion is located in the IGM. The latter does not remain neutral. It is in fact highly ionized by the background UV radiation emitted by galaxies and quasars.

## 1.4 The Epoch of Reionization (EoR)

---

The epoch of Reionization is a cosmological process where neutral Hydrogen was transformed into ionized Hydrogen. The EoR commenced when the first stars, quasars and AGN switched on producing UV radiation that transformed the surrounding HI gas into HII.

The complete process of the Reionization of Hydrogen involves several stages. The initial stage consists of many individual ionizing sources turning on and emitting Ly $\alpha$  continuum photons at  $\lambda \leq 912\text{\AA}$  (in order to ionize H, photons with  $\geq 13.6$  eV are required) that ionize their surroundings. These photons travel through the gas inside the host halo until they emerge generating bubbles. The ionizing fronts propagate easily into the low-density regions (voids) leaving behind neutral gas in the high-density regions. During this period, the IGM is a medium characterized by highly-ionized regions and neutral zones separated by ionizing fronts. This stage was followed by a rapid overlap of the HII bubbles (*a.k.a* **Overlap epoch**). The gas in the intersection of the bubbles is exposed to a strong radiation field, coming from many sources. As a consequence, the HII regions grow rapidly until they reach the low-density gas. The process leads to a phase where the ionizing background fills all the IGM and there are only a few self-shielded regions with neutral Hydrogen (Rees, 2000; Barkana & Loeb, 2001; Barkana, 2002; Wyithe & Loeb, 2004).

The candidates that effectively produce high energy photons include: the first stars (POPIII, Abel et al., 2002; Bromm et al., 2002; Yoshida et al., 2003), the second generation of stars (POPII, Ciardi & Ferrara, 2005; Mellema et al., 2006) and quasars (with a black

hole seed of  $10^6 M_\odot$  (Haiman & Loeb, 1998; Dijkstra et al., 2004; Hassan et al., 2017). Other candidates have been proposed, as miniquasars, with masses around  $10^{3-6} M_\odot$  (Mortlock et al., 2011; Bolton et al., 2011; Smith et al., 2017), decaying or self-annihilating dark matter particles or decaying cosmic strings. Nonetheless, the latter objects seem to be unlikely to ionize the Universe by themselves (Zaroubi, 2013).

The influence of POPIII objects depends on how much of their radiation escapes into the IGM. Much of the Ly $\alpha$  continuum emitted within a galaxy could be absorbed. In this case, the number of massive stars and black holes required to build up the UV budget close to the Lyman limit to complete the Hydrogen Reionization would be greater than the current detections (Rees, 2000). Moreover, miniquasars would not ionize the Universe (according to the current observational signatures) for two reasons: they would heat the IGM beyond the observational limits for the 21 cm radiation (Nusser, 2005; Chuzhoy et al., 2006; Zaroubi et al., 2007; Thomas & Zaroubi, 2008) and their black holes (BH) powered sources produce very hard photons in the X-rays range ( $\geq 10$  keV), that penetrate much into the IGM and heat it up further than the UV radiation. In addition, the production of hard X-rays at  $z \sim 6$  would be observable today as a soft X-ray background (SXB Haiman, 2016), but such background has not been detected. On the other hand, cosmological simulations have shown that BH are quite rare at this stage of the Universe, and their number density would have been too low to provide the UV photons required to ionize the Universe (Dijkstra et al., 2004; Salvaterra et al., 2005).

The maximum ionized region induced by a spherical source of radiation is defined as the Strömgen radius. Under the assumptions that the recombinations are counterbalanced by the ionizations, the source is steadily ionizing its surroundings and it is not affected by the Hubble flow:

$$r_{strom} = \left( \frac{3 \frac{dN_\gamma}{dt}}{4\pi\alpha n_H^2} \right)^{1/3}, \quad (1.19)$$

where  $\alpha$  is the recombination coefficient and  $n_H$  is the number density of Hydrogen. For O or B stars, this radius is typically  $\sim 1$  pc.

Nonetheless, a more realistic approach takes into account the instability of the ionizing object and the cosmological expansion (Loeb, 2010):

$$\bar{n}_H \left( \frac{dV}{dt} - 3H(t)V \right) = \frac{dN_\gamma}{dt} - \alpha \langle n_H^2 \rangle V, \quad (1.20)$$

where  $V$  is the volume of the spherical HII region,  $H(t)$  the Hubble parameter and  $\langle n_H^2 \rangle$ , the volume-averaged H density. The mean density  $\bar{n}_H$  scales as  $a^{-3}(t)$ . At some point, the ionizing fronts encounter each other, merge and keep growing in size.

If the ionized gas is not uniform, but instead is rather clumpy, the recombination time is short. This is the case during the progression of Reionization. There is a parameter used to define the distribution of gas in non-uniform environments, the clumping factor:

$$C = \frac{\langle n_H^2 \rangle}{\bar{n}_H^2}. \quad (1.21)$$

Basically,  $C$  quantifies how clumpy the distribution of matter is in a medium.

Nowadays, there is a consensus in the community to keep the numerical value of  $C$  as being 2-3 in low-density regions  $\Delta = \delta + 1 = \frac{\rho_{\text{gas}}}{\langle \rho \rangle} < 100$  (i.e. CGM and IGM) according to radiative transfer (RT) results from Pawlik et al. (2009).

One way to study the photon budget required to complete Reionization is by counting the number of ionizing photons released from a given source, such that:

$$N_{\text{ion}} = N_{\gamma} f_* f_{\text{esc}}, \quad (1.22)$$

where  $N_{\gamma}$  is the number of photons released by a galaxy,  $f_*$ , the conversion efficiency from baryons to stars and  $f_{\text{esc}}$  is the escape fraction. If recombinations are negligible and all the photons could escape to the IGM,  $N_{\text{ion}} = N_{\gamma}$ . However,  $f_*$  is quite small (around 10%) and the adopted value for  $f_{\text{esc}}$  varies in the literature from 1 (measured up to  $z \sim 4$  by Grazian et al., 2017; Vanzella et al., 2015) to 50 percent (Vanzella et al., 2016). Conservative constraints indicate that  $f_{\text{esc}} \geq 10\text{-}20\%$ . In other words, the number of ionizing photons is significantly lower than the number of photons released by a given source. At  $z$  higher than 5, it is impossible to estimate  $f_{\text{esc}}$  from galaxies due to the increased opacity of the IGM, however, theoretical works showed that the efficiency is quite low at high redshift (Hutter et al., 2014).

It is possible to calculate the number of released photons  $N_{\gamma}$  by type of object (POPIII, POPII, quasars, AGNs, etc.), but without an accurate estimate of  $f_{\text{esc}}$  and  $f_*$ , quantifying the relative contribution of each type of ionizing source to the total budget of photons required to complete the EoR is extremely challenging. In addition, this calculation implies a prior knowledge on the density of ionizing sources per unit volume at any stage. This

degeneracy in the parameter space makes very expensive a self-consistent model.

However, if the main contributor to the ionizing background in the EoR are star-forming galaxies (Robertson et al., 2013, 2015; Fontanot et al., 2014), equation (1.22) accounts for the total rate of UV ionizing photons per unit volume produced for galaxies (Lidz, 2016). In such case:

$$\dot{n}_\gamma = f_{\text{esc}} \zeta_{\text{ion}} \rho_{\text{UV}}, \quad (1.23)$$

where  $\dot{n}_\gamma$  is the number of ionizing photons per comoving volume per unit time,  $\zeta_{\text{ion}}$ , the number of ionizing photons per erg per second per Hz emitted at a restframe wavelength of  $1500 \text{ \AA}$ <sup>6</sup> and  $\rho_{\text{UV}}$  the UV luminosity density in  $\text{erg/s/Hz/Mpc}^{-3}$ . Robertson et al. (2013) explored the different stellar population synthesis models proposed by Bruzual & Charlot (2003) and adopted the best values for  $\zeta_{\text{ion}} = 10^{25.2} \text{ erg/s/Hz}$  and  $f_{\text{esc}} = 0.2$  (this value assumes that galaxies with escape fractions in this order can reionize the Universe and sustain the ionization later on, Fan, 2012; Grazian et al., 2017). Under these assumptions,  $\rho_{\text{UV}}$ :

$$\rho_{\text{UV}} = 1.6 \times 10^{26} \text{ ergs/Hz/Mpc}^{-3} \left( \frac{0.2}{f_{\text{esc}}} \right) \left( \frac{\dot{n}_\gamma}{3\gamma/\text{H atom/Gyr}} \right) \left( \frac{10^{25.2} \text{ erg/s/Hz}}{\zeta_{\text{ion}}} \right). \quad (1.24)$$

The choice of  $\dot{n}_\gamma$  corresponds to the common assumption that 2 photons per atom per Gyr are required to maintain the IGM reionized at  $z = 6$  (Miralda-Escudé, 2003).

At the end of the EoR ( $z \sim 6$ ), the Universe goes through a **post-overlap** stage and the ionized bubbles are typically  $\sim 10$  physical Mpc in size (about 70 Mpc at the present time). The IGM is mostly an ionized medium and there are some pockets of neutral gas left behind in high density regions that self-shield Hydrogen in very thick environments. By this time, the global distribution of the UV sources is highly uniform.

#### 1.4.1 Observational evidence for the EoR

Observations of the EoR are extremely challenging. Nevertheless, progress is being made with different techniques, that suggest that the IGM is highly ionized at its end. Gunn & Peterson (GP, Gunn & Peterson, 1965) predicted “an appreciable optical depth in Thomson scattering for very distant objects” with the form  $\tau_{GP} = 2.2 \times 10^4 (1+z)^{3/2} (n_{HI}/n_H)$ .

<sup>6</sup>The luminosity function of ultraviolet selected galaxies at rest-frame  $1500 \text{ \AA}$  has been measured at  $z > 6$  using the break technique in a high- $z$  spectrum that covers a rest-frame Ly $\alpha$ , and therefore, it has signatures of IGM Ly $\alpha$  absorption (McQuinn, 2016).

As can be seen from this equation, a small increase in the neutral H fraction results in a large boost in the optical depth. With the discovery of quasars at  $z > 6$ , the predicted Gunn–Peterson effect was finally observed by Becker et al. (2001) and confirmed with subsequent detections from Fan (2006). Additionally, Fan et al. (2006) showed that the effective GP optical depth -the observed transmitted flux ratio- changes abruptly from  $(1+z)^{4.3}$  to  $(1+z)^{\geq 11}$  at redshifts greater than 5.7 heralding the end of the Reionization era. This sudden change cannot be explained by random fluctuations in the background density field (Becker et al., 2015b).

On the other hand, the study of the CMB is remarkably important for different reasons: WMAP confirmed that the CMB is polarized, however, neutral Hydrogen cannot polarize the CMB radiation, but ionized H can. This is indirect evidence of a process posterior to the decoupling of matter–radiation that polarized the primordial CMB spectrum. Polarization measurements of the CMB anisotropies quantify the optical depth due to Thompson scattering of electrons freed in the EoR (Zaldarriaga, 1997; Furlanetto et al., 2006; Keating & Miller, 2006; Larson et al., 2011). With the new measurements by the Planck Collaboration et al. (2015), the Thompson optical depth inferred is  $\hat{\tau} = 0.066 \pm 0.012$ , significantly lower than the value reported with the WMAP-9 data (Bennett et al., 2013), leading to  $z_{EoR} = 8.8 \pm 1.2$  (mid–point in a model of instantaneous Reionization, with  $\Delta z = 0.5$ ). These results alleviate the difficulty of explaining Reionization in the IGM from starlight produced by very early and faint galaxies, that have not been detected yet (Robertson et al., 2013; Fontanot et al., 2014; Madau & Haardt, 2015; Robertson et al., 2015) and reduce the tension with the inferred evolution of the ionizing background at  $z \sim 7$  deduced from the UV luminosity function of Ly $\alpha$  emitters (LAE) (Bouwens et al., 2015; Mitra et al., 2015).

Moreover, the temperature anisotropies measured by the WMAP and Planck missions (one part in 100000 in the mean temperature of the CMB) correlate with the density inhomogeneities and give a first approximation of the location of the overdensities where massive galaxies are preferentially found.

CMB results can be also folded in with the kinetic Sunyaev-Zeldovich effect (Sunyaev & Zeldovich, 1980)<sup>7</sup>. For example, Zahn et al. (2012) found a lower limit for EoR of

---

<sup>7</sup>The Sunyaev-Zeldovich effect is a spectral distortion of the CMB by the scattering of the CMB photons off a distribution of high energy electrons (this imprint is proportional to the Thompson scattering cross section  $\tau_e$ ). If there is a cluster in the region where the anisotropy is found, its properties can be inferred from the departure of the temperature detected with respect to the  $T_{\text{CMB}}$ . For instance, the velocity of such cluster with respect to the CMB rest frame leaves an additional spectral distortion due to the

$z > 5.8$  and a duration of  $z \leq 7.9$ . Further constraints on the likely end redshift of Reionization come from the evolution in the Ly $\alpha$  luminosity function from  $z \sim 5.7$  to  $z \sim 6.5$  (Kashikawa et al., 2011). A complete overview of the observational constraints on the end of the EoR to date is shown in Bouwens et al. (2015). Future telescopes such as the Square Kilometer Array (SKA) and the James Webb Space Telescope (JWST) will give us additional information about the duration and evolution of the EoR.

Finally, evidence indicates that the IGM underwent a patchy reionization process with a non-uniform temperature and an inhomogeneous ionizing background that filled the Universe at that time (Becker et al., 2006). There are alternative ways to infer the redshift of the end of the EoR such as the shape of the red damping wing of H saturated absorption features in the quasar spectra or from the transmitted flux between the Lyman series lines. Nonetheless, these features are complicated to detect in the vicinity of quasars, which tend to ionize their surrounding environment (Wood & Loeb, 2000).

#### 1.4.2 The topology of the HII regions

Understanding the spatial distribution of the ionizing sources and its impact on the evolution of the HII regions in the IGM during the progression of the EoR is fundamental. This requires to investigate the transition of the IGM from a neutral Universe prior Reionization to the *swiss cheese* configuration at  $z \sim 6$ . Observations of the IGM at the end of the EoR show that the Hydrogen in the intergalactic medium is 99% ionized, but there are still some large isolated pockets of neutral gas (Becker et al., 2015a). By following the distribution and luminosity of the ionizing sources and the sinks (IGM and voids) of these UV photons at high redshift with hydrodynamical simulations, it is possible to predict the topology of the IGM during and after the Hydrogen Reionization.

Numerical results from Furlanetto et al. (2004), McQuinn et al. (2007), Mesinger & Furlanetto (2007), Wyithe & Morales (2007), Bauer et al. (2015) favour an inside-out Reionization, from high to low density regions. In this scenario the ionizing fronts leave the sources (at high density) and progress slowly as the UV bubbles expand with time. However, some authors (Choudhury et al., 2009) claim the inside-out progression is not realistic, because it ignores the impact of the recombination and self-shielding in high density regions. (Miralda-Escudé et al., 2000; Choudhury & Ferrara, 2005).

---

Doppler effect of the cluster bulk velocity on the scattered CMB photons. The peculiar velocity of the cluster  $v_{\text{pec}}$ , perpendicular to the line of vision, leads to an observable thermal distortion known as kinetic Sunyaev-Zeldovich effect, proportional to  $\tau_e$  and  $v_{\text{pec}}$ .

Choudhury et al. (2009) and Choudhury (2009) showed that ignoring the inhomogeneous spatial distribution of the high-density regions -where recombinations play an important role- leads to ambiguous conclusions on the topology of the IGM, because the distribution of these regions reduces the mean-free path of the ionizing photons and the size of the HII bubbles. In this case an outside-in topology is favoured (the low density regions, as voids and IGM, far from the ionizing sources, are ionized first rather than the high density gas that is producing the ionizing fronts). In particular, Choudhury et al. (2009) test how clumpy the IGM looked at  $6 < z < 12$ , following the distribution of the ionizing sources with a friends-of-friends algorithm. The main conclusion from that work indicates that Reionization initially proceeds in an inside-out fashion: the high density regions were ionized first, because they host the sources of the UV photons, but later on, when Reionization is reaching its end, the low density regions are highly ionized, in an outside-in topology. The recombination rate around the high-density regions -where the UV sources are located- is so strong, that these regions remain mostly neutral until  $z \sim 5$ . This complete scenario seems to be more physically realistic and it gives a more insightful idea of the overall progression of the EoR.

### 1.4.3 UV background variations during the EoR

Understanding the Epoch of Reionization is intimately tied to the evolution of the UV background: the grand sum of all photons that have escaped from quasars and galaxies. Its spectral energy distribution is measured reasonably well at  $z < 5$  (Bolton et al., 2005) and modelled (Haardt & Madau, 2001). Haardt & Madau (2012) used a cosmological 1D radiative transfer model that follows the propagation of H and He Lyman continuum radiation in a clumpy ionized IGM, and found that the mean free path and the hydrogen photoionization rate decrease with redshift. However, as  $z \geq 6$  is approached, the population of UV sources is not well determined (Haardt, 1999). The uncertainty on estimating the UV photon emissivity from each type of object is caused by the lack of knowledge on the star formation rate, clumping factor and UV escape fraction (Cooke et al., 2014) at the redshift of interest, which are strongly model-dependent. Measuring Lyman series absorption or UV emissivity in a spectrum blueward of  $\text{Ly}\alpha$  at  $1216 \text{ \AA}$  is rendered almost impossible by the increasing density of matter and neutral hydrogen fraction at redshifts greater than 5.5.

On the other hand, the assumption of a uniform UV radiation field approximation

breaks down close and during the EoR, when the interaction of the ionizing sources with the IGM requires a very accurate description (Lidz et al., 2006). A real-time Reionization simulation should ionize first high density regions and fill some regions before than others, leading to a multiphase IGM with spatial fluctuations (Lidz, 2016).

## 1.5 The Universe after Reionization

---

Once Reionization has concluded ( $z \leq 6$  or  $t > 1$  billion years after the Big Bang), some gas remains neutral in cold high-density environments. These regions are gradually ionized while galaxy formation is proceeding. As this occurs, the mean absorption path is growing and the ionizing background is more smooth and uniform in the IGM. Thus, the **post-overlap** period of Reionization continues and some isolated pockets of neutral Hydrogen are left inside collapsed structures, that have been detected and studied at redshift  $z < 5$  (Barkana & Loeb, 2001).

Later on, at  $z \sim 3$ , the Universe faced another phase transition, the Reionization of Helium. Since the ionization potential of HeII is 54.4 eV, the vast majority of photons produced by hot stars are unable to induce this transition. Therefore, Helium remains singly ionized until a large population of quasars with a typical luminosity  $L_\nu \approx \nu^{-1.6}$  (Telfer et al., 2002, found that quasars described by this power-law represent the entire ionizing continuum from  $\sim 2 - 10$  eV) can produce enough *hard* photons to ionise Helium at redshift 3 (Sokasian et al., 2002; Wyithe & Loeb, 2003; Furlanetto & Oh, 2008). Observations of the low density IGM show the existence of a HeII Ly $\alpha$  forest (Croft et al., 1997; Miralda-Escudé et al., 1996; Fardal et al., 1998) simultaneous to the peak of the stars and quasars formation in the Universe.

Today, the Universe is experiencing a completely different stage: its expansion is speeding up, induced by an unknown component. Either if it is a smooth field, the energy of the vacuum or just a purely geometrical effect, astronomers know that at some point in the future matter will be so isolated and far apart that merging of galaxies will be less likely, and so too the formation of new population of stars.

---

## 1.6 Purpose of the Thesis

---

Understanding when the EoR ended, its progression and the identification of the main sources that contributed to the bulk of ionizing photons are still open questions in Cosmology. In particular, this thesis will address: can we use metal ions to determine the ionization state of the IGM at  $4 < z < 8$ ? and what types of galaxies are responsible for the IGM enrichment at  $4 < z < 8$ ?

This work studies the physical conditions of the IGM and CGM in the high redshift Universe, taking advantage of state-of-the-art SPH (Smoothed-particle hydrodynamics) simulations based on the constraints on the star formation rate density at high redshift, wind models that describe galactic feedback and a uniform ionizing UV background. By following the evolution of metal ions at high redshift (when Hydrogen is opaque), it is possible to determine the state of ionization of the IGM.

In order to make predictions on the evolution of the metal and HI content of the IGM and CGM at high redshift, this thesis has a major aim to study and analyze the evolution of the metal–line absorptions present in the spectra of the very distant quasars by generating many mock observations via high-resolution hydrodynamical simulations. This work is carried out in three stages: first, a reliable comparison between observations of metals ( $\lambda > 1216 \text{ \AA}$ ) and HI absorbers with synthetic spectra. Secondly, the simulations are used to provide insight into the underlying physical conditions for the observations. Finally, theoretical predictions are made about the galaxies that are responsible for the metal and HI line systems.

The theoretical approach offers many advantages: it is possible to calibrate the simulations to match observations at low redshift and compare with different observables at high redshift. Besides, the numerical runs reproduce the physical conditions of the gas during the Epoch of Reionization and later on and, as a consequence, they allow us to infer the ionization state of the gas with metal absorption lines and HI (up to  $z = 6$ ). Finally, the sample of synthetic absorbers is at least an order of magnitude larger than the observations of metals available at high redshift. Thus, there is more freedom to study different column density ranges and the distribution of the absorptions in the regions of interest. Ultimately, this method sheds light on the chemical enrichment that is proceeding parallel (but not independently) to the evolution of the ionizing state of the IGM during the EoR.

The numerical simulations also allow us to probe regimes where the observations are not achievable, either because the Universe is opaque to the photons collected by current telescopes or, instead, missions planned to study the Reionization and galaxies at high redshift are not launched yet.

Additionally, this thesis attempts to study to first approximation variations of the assumed UVB (Haardt-Madau 2012) and metallicity in the photoionization modelling (set to solar in most of the studies) in post-process. The work of Becker et al. (2015b) concluded that the distribution of the effective GP optical depth cannot be reproduced with a uniform ionizing UVB or a simple background model that assumes galaxies are the only sources for the ionizing photons and uses a fixed mean free path. Therefore, a relevant component of this research is focused on describing and imposing some theoretical constraints on the UV ionizing background, from the observations of metals at  $z \sim 6$ . In summary, we use mock spectra of metal absorption lines at a redshift where HI is opaque (and can no longer probe the ionization state of the IGM) to describe the intergalactic medium and its evolution.

## 1.7 Thesis outline

---

The structure of this thesis is divided as follows: Chapter 2 describes the role of simulations in the study of Cosmology, explains the numerical simulations used in this work, the different feedback models implemented and the molecular and low-temperature metal cooling module. Subsequently, the properties of the galaxies at high redshift are reproduced. Finally, a discussion of the UV background implemented and the codes used to post-process the simulations is presented.

Chapter 3 presents results from García et al. (2017a): a discussion of the method used to include the uniform UV background Haardt & Madau 2012 (hereafter HM12), the photoionization modelling and the column density calculation procedure. Moreover, an analysis of the IGM evolution and its physical features (such as metallicity, number density and ionization state) with redshift are displayed. Finally, it presents the analysis of the evolution of the ionic species in the simulations (under a uniform HM12 UVB) using: 1) the cosmological mass density ( $\Omega$ ), 2) relationships among different species to compare with observations at  $4 < z < 8$  and 3) predictions on the ionization state of the gas during the EoR. The discussion mainly focuses on CIV, but additional results with CII, OI, SiII and SiIV are presented. Complementary, the evolution of the cosmological mass density

of neutral Hydrogen, the importance of the introduction of the molecular cooling module in the simulations and the limitations of the method used to compute  $\Omega$  from the DLA definition are discussed. The chapter closes with the main conclusions from García et al. (2017a) and the discussion of the caveats of the procedure adopted in this work.

Chapter 4 is devoted to present a theoretical model to explain the properties of an observed  $z = 5.72$  Lyman- $\alpha$  emitter (LAE) galaxy – CIV absorption pair separated by 1384 comoving kpc/ $h$  (Díaz et al., 2015). Additionally, it includes a prediction of the origin of the metals in the IGM, made from the best geometrical configuration found in the simulations. Finally, this chapter explores the physical connection of these galaxy–absorption systems, using the properties of the galaxies (mass, distance from the absorption) and the column densities of the absorbers in the simulated box. All the results of this chapter are published in García et al. (2017b).

Chapter 5 introduces modifications of the HM12 UVB (normalization, hardness of the quasar part of the spectrum) and variations of the metallicity content in the photoionization modelling and presents results on the observables described in Chapter 3 (column density distribution function, cosmological mass density and column density relationships among metal ions).

Finally, Chapter 6 summarizes the conclusions and implications derived from the projects involved in this thesis.

Throughout the thesis, the prefix c for comoving and p for physical distances is used.



# 2

## Numerical Simulations and Post-process

*What is unconceivable about the universe is that it is at all  
conceivable.*

Albert Einstein

### **Cosmological simulations**

---

In the last decades, there has been a huge development of numerical simulations devoted to reproduce the structure formation in the Universe. The numerical approach has proved to be a very powerful tool to describe the observations and, in some cases, predict future discoveries when the observational facilities in hand could not survey such a large area of the sky. Theoretical models have also guided the design of telescopes/instruments, providing sensitivity limits and constrains on the conditions necessary to improve ongoing observations. In addition, numerical runs have been used to calibrate cosmological sets of parameters and/or test different cosmological models by predicting several observables, as the growth of cosmological density perturbations (Peebles, 1993; Padmanabhan, 1993; Peacock, 1999) or the large scale structure evolution (Springel et al., 2006). Lastly, simulations allow astronomers to build a bridge between theory and observations.

Cosmological simulations have also given a deep insight in the galaxy formation process. Different approaches can be considered to describe the formation of galaxies at different times, each formulation with its shortcomings and effective solutions to improve the physical modelling. The introduction of cold dark matter (CDM) in the standard cosmological model, and therefore, in the simulations, proved how effective the numerical schemes can be when describing the large scale structure (LSS). The ultimate goal is to

evaluate the evolution of the structures observed today given the initial conditions for the density perturbations, a large range of physical processes and a cosmological model: the density associated with ordinary matter, dark matter, radiation and dark energy (cosmological constant  $\Lambda$ , scalar fields or modified gravity model), on top of the spatial curvature of the Universe.

The cold dark matter model has tremendous power in describing the growth of perturbations and, so far, its theoretical predictions are highly compatible with observations of the temperature anisotropies of the cosmic microwave background and the baryonic acoustic oscillations (Efstathiou et al., 1992; Aubourg et al., 2015). Moreover, the evolution of dark matter halos from the primordial matter fluctuations is well represented by current numerical simulations, which are a well established method to study the galaxy formation and the LLS. The treatment of dark matter is relatively simple to implement in the simulations due to its collisionless nature: DM particles only interact via gravity, they are extremely massive and non-relativistic. Dark matter is expected to decouple from the plasma very early in the Universe.

Important theoretical results have been established with numerical algorithms that include CDM (Springel, 2005): the density profile of dark matter halos (Navarro et al., 1996), the existence of DM internal structure (Tormen et al., 1997), clustering properties of DM in a non-linear regime (Jenkins et al., 1998), the halo abundance (Jenkins et al., 2001), the temperature/gas profile of galaxy clusters (Evrard, 1990), the formation of structures that shape the Ly $\alpha$  forest (Theuns et al., 1998) and the properties of the Ly $\alpha$  absorption lines in the IGM (Hernquist et al., 1996), among others.

On the other hand, a proper description of baryonic matter is extremely complicated. In order to incorporate the gas dynamics to the treatment, several approximations and effective modelling are required, due to the diversity of interactions of the baryons, in addition to gravity.

Since the first attempts in the early seventies, cosmological simulations have grown in size and spatial/mass resolution. Every year, more sophisticated physics modules are included. Galaxy formation simulations have introduced the growth and evolution of DM halos and many elements of gas accretion and feedback of the flowing gas, in addition to the stellar evolution in galaxies and the formation of supernova, BH, etc. These theoretical efforts have been continuously compared and complemented with observations.

Nonetheless, cosmological simulations are still constrained by a limited dynamical range, numerical errors and missing physics. The main endeavor in the field is to increase the computational power to improve the resolution and enable more rigorous examination of the current models and assumptions introduced numerically, especially when treating baryonic physics. Although this fine-tuning is quite delicate, current models have vastly improved over the last 20 years, with for example the inclusion of parallelization algorithms or better integration methods.

## 2.1 N-body simulations

---

N-body simulations were introduced to numerically compute the gravitational interaction of  $N$  bodies and then evolve their equation of motion. Historically, the N-body problem has been one of the main struggles for astronomers. There is an analytical solution to predict the motion of two bodies under their mutual gravitational field, a restricted solution for a subset of 3 bodies and for  $N > 3$  only numerical methods can be used to follow the dynamics of bodies under their mutual gravitational interaction. The problem becomes more complicated if the objects not only interact through gravity, but their physical response depends on temperature and density of the surrounding medium.

There are different N-body methods to test self-gravitational mass configurations. In all of them, mass is treated as point-like systems (herein, particles). The most widely used, the **Direct N-body** technique, is the most appropriate for collisionless systems, with large masses and negligible interactions except the gravitational one. In this method, the collection of particles interact with a softening function that regulates the gravitational field from  $N(N-1)$  particles (and prevents the self-interaction). In such way, the Newtonian equations of motion are calculated for  $N$  bodies moving at any time  $t$  with the acceleration computed from the contribution from the residual mass of the system (Aarseth, 1985). The main advantage of imposing the softening is that it prevents divergences of the field. In addition, this treatment does not require prior assumptions for the geometry of the system, hence, it is easily adaptive for different astrophysical applications, unlike grid methods (Sellwood, 1987).

A more efficient method than the direct N-body scheme is the use of *Trees* (Appel, 1985; Barnes, 1986). In this case, the gravitational force due to a distant system composed of a large collection of particles is modelled as being approximately an effective contribu-

tion from a single particle in the center of mass of such far system. As a consequence, if the systems are reasonably distant, their physical extension is less important in the calculation of the force (according to the far-field approximation).

An alternative to the techniques above is the implementation of grids (or particle-mesh, PM). These are the most computationally efficient techniques to calculate the interaction for large N-bodies (first used in an astrophysical context by Efstathiou & Eastwood, 1981; Klypin & Shandarin, 1983; White, 1983). The gravitational field is computed in a cell that contains  $x$  number of particles inside a fixed volume, and the acceleration of each particle is an interpolation of the overall contribution of the particles in the cell. By reducing the number of calculations of the gravitational field from N-particles to M-grids (where  $M < N$ ), the evaluation is much efficient and the timesteps needed to compute the force acting on the particles decrease significantly. However, due to the interpolation, the method is not the best option for small scales, because there is a loss in resolution in nearby cells.

More sophisticated versions of PM use a grid to calculate the distant field and the direct N-body method to compute the interaction among neighbours. In order to alleviate the low resolution of grids in short-range interactions, Fourier expansions can be placed in those regions (Couchman, 1991).

N-body simulations have been used in different astrophysical problems, from solar system dynamics to the LSS of the Universe, as well as formation, internal dynamics, merging and clustering of galaxies (with and without bulges or halos) in an expanding Universe. Yet, one of the greatest victories of the N-body runs is the dark matter distribution, mainly because DM is collisionless. Dark matter halos growth is represented with merger trees. These diagrams show how halos grow in size/mass through merger events and by accretion of objects below the halo mass resolution. A schematic merger tree of DM halos is shown in Figure 2.1.

In the cosmological context, N-body routines are an alternative to describe the galaxy assemblage, due to the multiple interactions of a large conglomerate of stars in presence of DM halos and gas (as discussed by Gott, 1977). Recent observations of galaxy morphology also introduced a new level of difficulty, since each galaxy shape seems to be supported by different gas accretion models.

N-body simulations present some limitations in terms of the spatial resolution. On one hand, structures described by just a few particles are not well-defined (for instance, small scales in dense regions or large scales in low density regimes). On the other hand, if the gravitational potential is expressed as a multipolar expansion, the numerical resolution is limited to the highest order included. Finally, the physical size of the particles imposes a lower limit in the resolution of the structures that can be studied with a given simulation.

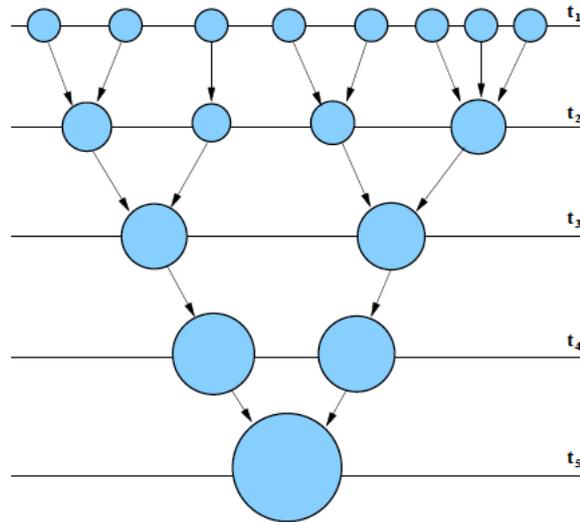


Figure 2.1 Schematic merger tree for a dark matter halo in a hierarchical clustering model. The horizontal lines represent different times in the history of the halo  $t_i < t_{i+1}$  (timesteps in N-body simulation/Monte Carlo realization of the merger tree). The size of the circles is an indicator of the mass of the halo. The halos grow in size through merger events and by accretion of objects below the halo mass resolution. The resulting halo is shown in  $t_5$ . Taken from Baugh (2006).

## 2.2 Galaxy formation

The big picture of the galaxy formation process is still a delicate puzzle. Although there is not a plausible candidate detected to date for the dark matter component (it is assumed collisionless and cold), its evolution is well established: DM interacts via gravity with other matter components and follows a hierarchical clustering model from the initial density perturbations set at the time of Inflation<sup>1</sup>. On the other hand, the formation and evolution of individual stars has been extensively studied and is now well understood. Nevertheless, astronomers still encounter some difficulty while describing free gas, the one which has not been incorporated in the stars. Different processes involve gas, apart from the gravitational interaction: cooling when it flows into galaxies, reheating from galactic

<sup>1</sup>Density fluctuations obey a Gaussian distribution.

feedback and conversion into stars. Baryonic gas is a multiphase medium and its properties strongly depend on the temperature and density where it is located. Due to the complexity of the problem, a complete model that provides insight on the gas in the different regimes needs to employ well-motivated physical approximations, effective schemes and/or numerical approaches.

Furthermore, the most challenging issue is to evolve all the components of the galaxy in a self-consistent way, coupling gas + dark matter + stars through gravity in the context of an expanding Universe. The galaxy formation paradigm proposed by White & Frenk (1991) set all the elements needed in a competent model: cold dark matter, gas cooling, star formation, feedback and stellar populations (Baugh, 2006). Observations of galaxies are just snapshots, hence, they only reveal the state of a structure at a particular  $z$ . Theoretical models should progress in such a way to match the inferred properties of galaxies at any stage of the Universe.

There are two alternatives to investigate galaxy formation and evolution: semi-analytical modelling (Baugh, 2006; Croton, 2006) and hydrodynamical simulations. In the former, heuristic models are proposed to account for observed properties of galaxies following an “analytical” approach, but supported by numerical routines (either N-body simulations or Monte-Carlo realizations) to build-up the dark matter halos according to the Press-Schechter model. Instead, hydrodynamical simulations solve the gas hydrodynamic equations directly, employing numerical methods, which makes them more computationally expensive. In the latter case, the treatment for DM is the same as in semi-analytical models, but preferentially N-body simulations are used because they provide a more faithful representation of the merger trees of DM halos than Monte-Carlo runs, especially if the timesteps are large.

Since this work is devoted to hydrodynamical simulations, there is brief description of both models in the following section, but the developments for gas and collisionless components are extensively discussed in the context of the hydrodynamical code GADGET.

### **Semi-analytic modelling**

In semi-analytical models the goal is to predict the properties of the galaxies and their evolution in cosmic time assuming a hierarchical clustering of structure formation (Lacey, 2001). The method assumes a cosmological model, a spectrum of primordial density fluc-

tuations and some simplifying assumptions about the physics of the gas located in galaxies and the geometry of the system. Semi-analytical models include recipes to provide an effective description of galaxy properties “ab initio” and generate galaxy luminosities with dust extinction models and self-consistent star-formation histories (Baugh, 2006).

Refined semi-analytical models should include additional physical processes to reconcile the overwhelming amount of observations of galaxies in different bands and redshifts, without introducing more degrees of freedom to the modelling. This is probably the most critical point to the approach, although most of the parameters in the models are physically motivated.

### **Hydrodynamical simulations**

Hydrodynamical simulations calculate directly the evolution of the dark matter, gas and stars based on the fundamental equations of gravity and hydrodynamics, without any assumption about the geometry<sup>2</sup>. The DM is a cold and non-relativistic component (CDM), therefore it can be described by the collisionless Boltzmann equation and its solution in the phase-space determines the evolution of the particles with an N-body method. On the other hand, baryonic gas requires the inclusion of hydrodynamical interactions in addition to the N-body scheme used to traced DM- and star-particles.

Since gas can be treated as a continuous fluid, there are two canonical methods to solve the hydrodynamical equations: the Eulerian approach discretizes the space in order to get the evolution of the hydrodynamical quantities in a fixed mesh (e.g. Cen & Ostriker, 1992, 1993; Yepes et al., 1996). On the other hand, the Lagrangian scheme discretizes the mass, following the variables of the particles in the fluid (see works from Evrard, 1988; Hernquist & Katz, 1989; Navarro & White, 1993; Couchman et al., 1995; Katz et al., 1996).

One of the main advantages with mesh codes is their accuracy when describing contact discontinuities and shocks. However, static Eulerian methods do not offer much flexibility in terms of scale, a notably withdrawal in the cosmological context, where a large dynamical range is necessary. To alleviate this issue, adaptive mesh refinement (AMR) algorithms have been introduced in the simulations (Abel et al., 2002; Kravtsov et al., 2002; Refregier & Teyssier, 2002). It is however possible that the intrinsic problem with mesh treatments is still unsolved: mesh codes present spurious generation of entropy due to mixing and a non-Galilean invariance in the Euler equations.

---

<sup>2</sup>A complete compilation of different hydrodynamical routines is found in (Kang et al., 1994).

Codes based on the Lagrangian approach, on the other hand, have proved to work well in the description of the large scale structure, because of their extensive dynamical scaling. The method treats self-gravity of the gas in a similar fashion to the dark matter implementation, making the model quite robust (Springel, 2005). Nevertheless, these models suffer of poor mixing between fluid elements, and shocks are not well-described, unless artificial viscosity is placed.

The vast majority of Lagrangian codes employs SPH (Smoothed particle hydrodynamics) method to describe the gas component. The SPH technique approximates the continuum nature of the fluid in the hydrodynamical equations to a discrete form (with particles, or interpolation points, Springel, 2010). To estimate physical quantities of the fluid, SPH relies on kernel interpolants (Gingold & Monaghan, 1982). This allows the method to have enough local resolution to follow the mass flow, a property particularly convenient to compute the density contrast in different media.

For this thesis, I used the SPH code GADGET. The main features of the model will be described in the next sections.

## 2.3 GADGET

---

GADGET (GALaxies with Dark matter and Gas intERacT) is one of the most widely used theoretical models to investigate the formation of structures in the framework of the CDM cosmology. As described in Springel (2005), the code follows the evolution of collisionless fluids (dark matter, stars, BH, etc.) and an ideal gas (mostly H and He), coupled by gravity. The collisionless component is solved with an N-body algorithm and the baryonic gas with the SPH method. The merit of SPH in GADGET is that the treatment conserves energy and entropy with a fully adaptive smoothing length. The gravitational forces are computed with a TreePM method (Xu, 1995): short-distance interactions are calculated with the *Tree* method, whereas long-range forces are determined with Fourier expansions adapted in the PM algorithm.

The non-interacting DM in the continuum limit is described by a collisionless Boltzmann equation coupled with the Poisson equation in the framework of the Friedmann-Robertson-Walker metric, that accounts for the expansion of the Universe.

The dynamics of the discretized system with N-particles is described by the Hamiltonian

$H$ :

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i a^2(t)} + \frac{1}{2} \sum_{ij} \frac{m_i m_j \varphi(\mathbf{x}_i - \mathbf{x}_j)}{a(t)}, \quad (2.1)$$

where  $\mathbf{x}_i$  are comoving coordinates and  $\mathbf{p}_i = m_i \dot{\mathbf{x}}_i a^2(t)$ , the canonical momentum of the particle  $i$ . If periodic boundary conditions are assumed inside a cube with linear size  $L$ , the gravitational potential  $\varphi(\mathbf{x})$  satisfies a Poisson's equation given by:

$$\nabla^2 \varphi(\mathbf{x}) = 4\pi G \left[ -\frac{1}{L^3} + \sum_{\mathbf{n}} \tilde{\delta}(\mathbf{x} - \mathbf{n}L) \right], \quad (2.2)$$

with  $\mathbf{n} = (n_1, n_2, n_3)$ , a three dimensional triplet. The signs in equation (2.2) warrant that this solution correspond to the 'peculiar potential' of the usual Poisson equation<sup>3</sup>. The discrete version of the peculiar potential for each particle is governed by:

$$\phi(\mathbf{x}) = \sum_i m_i \varphi(\mathbf{x} - \mathbf{x}_i). \quad (2.3)$$

The density distribution function for a single particle is the convolution of the Dirac function  $\delta$  with a normalized gravitational softening kernel of comoving scale  $\epsilon$ . GADGET applies the spline kernel (Monaghan & Lattanzio, 1985) and sets  $\tilde{\delta}(\mathbf{x}) = W(|\mathbf{x}|, 2.8\epsilon)$ , with  $W(r)$ :

$$W(r, h) = \frac{8}{\pi h^3} \begin{cases} 1 - 6 \left(\frac{r}{h}\right)^2 + 6 \left(\frac{r}{h}\right)^3 & 0 \leq \frac{r}{h} \leq \frac{1}{2}, \\ 2 \left(1 - \frac{r}{h}\right)^3 & \frac{1}{2} < \frac{r}{h} \leq 1, \\ 0 & \frac{r}{h} > 1. \end{cases} \quad (2.4)$$

With this functional form of the kernel, the Newtonian potential of a point-mass at zero lag in a non-periodic space is  $-\frac{Gm}{\epsilon}$ , the same as in a Plummer sphere of radius  $\epsilon$ .

In the code, the TreePM method is used to calculate the gravitational interactions between particles. The potential  $\phi(\mathbf{x})$  (equation (2.3)) can be expanded in the Fourier modes  $k$ , such that there is a term associated with the long- and short-range interactions, as  $\phi_{\mathbf{k}} = \phi_{\mathbf{k}}^{\text{long}} + \phi_{\mathbf{k}}^{\text{short}}$ . The long-range term is given by:

$$\phi_{\mathbf{k}}^{\text{long}} = \phi_{\mathbf{k}} \exp(-\mathbf{k}^2 r_s^2), \quad (2.5)$$

where  $r_s$  is threshold of the force split. The long-term potential is efficiently calculated with the mesh-based Fourier schemes. On the other hand, the short-distance potential is

<sup>3</sup>For further details, see (Springel, 2005).

computed under the assumption of  $r_s \ll L$ , with  $L$  the size of the box. The solution of this potential in real space is given by:

$$\phi^{\text{short}}(\mathbf{x}) = -G \sum_i \frac{m_i}{r_i} \operatorname{erfc} \left( \frac{r_i}{2r_s} \right). \quad (2.6)$$

The distance  $r_i$  is chosen such that  $r_i = \min(|\mathbf{x} - \mathbf{r}_i - \mathbf{n}L|)$  is the smallest separation between the particle  $i$  to the coordinate  $\mathbf{x}$ . The short-range potential is solved with the tree algorithm. As long as the tree walks in small regions around the particles (to satisfy the condition for the cut-off  $r_s \ll L$ ), no boundary conditions corrections are required.

The hydrodynamical treatment of the gas employs the SPH technique. Gas particles are identified by their position  $\mathbf{r}_i$ , velocity  $\mathbf{v}_i = \dot{\mathbf{x}}_i$  and mass  $m_i$ . Each particle is treated as a fluid element. The formulation of SPH in GADGET conserves energy and entropy. The density estimate is defined as:

$$\rho_i = \sum_{j=1}^N m_j W(|\mathbf{r}_{ij}|, h_i), \quad (2.7)$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  and  $W(r, h)$ , the term in equation (2.4). The adaptive smoothing length  $h_i$  is defined such the mass contained in a kernel volume  $4/3\pi h_i^3$  is constant, when weighted by the density estimate:

$$4/3\pi h_i^3 \rho_i = \bar{m} N_{\text{SPH}}, \quad (2.8)$$

with  $N_{\text{SPH}}$  the number of smoothing neighbours and  $\bar{m}$  an average mass particle. The equations of motion for the SPH particles are obtained from the Lagrangian of the fluid in its discrete version (Springel & Hernquist, 2002), as:

$$\frac{d\mathbf{v}_i}{dt} = - \sum_{j=1}^N m_j \left[ f_i \frac{P_i}{\rho_i^2} \nabla_i W(|\mathbf{r}_{ij}|, h)(h_i) + f_j \frac{P_j}{\rho_j^2} \nabla_j W(|\mathbf{r}_{ij}|, h)(h_j) \right]. \quad (2.9)$$

The particle pressure follows the equation of state  $P_i = A_i \rho_i^\gamma$ , with  $A_i$  the entropy of each particle<sup>4</sup>. The coefficients  $f_i$  are given by:

$$f_i = \left( 1 + \frac{h_i}{3\rho_i} \frac{\partial \rho_i}{\partial h_i} \right)^{-1}. \quad (2.10)$$

---

<sup>4</sup>In absence of shocks or external sources of heat, the entropy is conserved with the dynamical evolution of the fluid. Nonetheless, if a discontinuity develops, the corresponding shock is treated with an artificial viscosity. For further details, see Springel & Hernquist (2002); Springel (2005, 2010).

The choice of a suitable time integration-method is very important to avoid numerical perturbations when the equations of state are evolved in time. In order to preserve the Hamiltonian structure of the system described in the equation (2.1), at each integration step, the Hamiltonian of the system is splitted in two terms:

$$H = H_{\text{kin}} + H_{\text{pot}}. \quad (2.11)$$

With this form of  $H$ , the time-evolution for both terms can be computed exactly. Two operators are now associated with the time-evolution of the system, the ‘drift’ and ‘kick’ functions (Quinn et al., 1997):

$$D_i(\Delta t) : \begin{cases} \mathbf{p}_i \rightarrow \mathbf{p}_i \\ \mathbf{x}_i \rightarrow \mathbf{x}_i + \frac{\mathbf{p}_i}{m_i} \int_t^{t+\Delta t} \frac{dt}{a^2} \end{cases} \quad (2.12)$$

$$K_i(\Delta t) : \begin{cases} \mathbf{x}_i \rightarrow \mathbf{x}_i \\ \mathbf{p}_i \rightarrow \mathbf{p}_i + \mathbf{f}_i \int_t^{t+\Delta t} \frac{dt}{a} \end{cases} \quad (2.13)$$

Here,  $\mathbf{f}_i$  is the force associated to each particle, and it is expressed as:

$$\mathbf{f}_i = - \sum_j m_i m_j \frac{\partial \phi(\mathbf{x}_{ij})}{\partial \mathbf{x}_i}. \quad (2.14)$$

Both the operators  $D_i(\Delta t)$  and  $K_i(\Delta t)$  are symplectic (i.e. they conserve the canonical transformations of the Hamiltonian during each integration step, thus, the Hamiltonian (2.11) has an exact solution every  $\Delta t$ ). With the definition of these operators, it is possible to adopt a time integration method to approximate the time evolution operator  $U(\Delta t)$  at an arbitrary  $\Delta t$ :

$$\tilde{U}(\Delta t) = D\left(\frac{\Delta t}{2}\right)K(\Delta t)D\left(\frac{\Delta t}{2}\right) \quad \text{or} \quad K\left(\frac{\Delta t}{2}\right)D(\Delta t)K\left(\frac{\Delta t}{2}\right). \quad (2.15)$$

These operators are the drift-kick-drift (DKD) and kick-drift-kick (KDK) leapfrog integrators. The result of their operation on the system is symplectic, since  $D(\Delta t)$  and  $K(\Delta t)$  are symplectic by definition.

GADGET has proven its efficiency in different aspects: in particular, there is a self-consistent implementation of the self-gravity, that makes it an excellent code to capture the gravitational collapse. As the Lagrangian method is used to solve the hydrodynamical

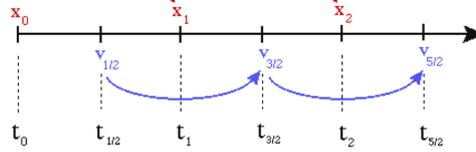


Figure 2.2 Integration time scheme used in GADGET to evolve the equations of motion in the system. When a force is computed in the code, it drifts the particle to the next time step. At that instant, the momentum of the particle is changed by the action of the ‘kick’ operator. In this way, long- and short-range interactions can be calculated. Taken from <http://cvarin.github.io/CSci-Survival-Guide/leapfrog.html>.

equations of the gas, it handles the advection processes very well. In addition, GADGET is highly adaptive, hence, it can be used to describe the large dynamical range of cosmological phenomena: from white dwarfs dynamics (Pakmor et al., 2008), formation of stars (Bürzle et al., 2011), magnetic fields in galaxies (Kotarba et al., 2009) to cluster of galaxies (Borgani et al., 2004; Dolag et al., 2009), among many others. There is low dissipation in the code and this leads to the non-existence of numerical mixing. Moreover, the code is quite efficient at computing the gravitational interaction through an N-body method, therefore large cosmological volumes can be simulated. GADGET-2 was used for the Millennium simulation (Springel et al., 2005c) and the current improved version, GADGET-3, is intensively used to study different cosmological eras.

### Additional physics included in P-GADGET3 (XXL)

The suite of numerical simulations used in this work are run with a customized version of GADGET-3: P-GADGET3(XXL). The model is an extension at high redshift of the AustraliaN GADGET-3 early Universe Simulations (ANGUS) project (Tescari et al., 2014; Maio & Tescari, 2015; Katsianis et al., 2015, 2016, 2017). P-GADGET3(XXL) takes into account the following physical processes:

- a multiphase star formation criterion (Springel & Hernquist, 2003);
- self-consistent stellar evolution and chemical enrichment modeling (Tornatore et al., 2007);
- Supernova (SN) momentum- and energy-driven galactic winds (Springel & Hernquist, 2003; Puchwein & Springel, 2013);
- AGN feedback (Springel et al., 2005a; Fabjan et al., 2010; Planelles et al., 2013);

- metal-line cooling (Wiersma et al., 2009);
- low-temperature cooling by molecules and metals (Maio et al., 2007, 2010; Maio & Tesconi, 2015).
- a low-viscosity SPH scheme to allow the development of turbulence within the intracluster medium (Dolag et al., 2005).

Moreover, the model is supported by:

- A parallel Friends-of-Friends (FoF) algorithm to identify collapsed structures.
- A parallel SUBFIND algorithm to identify substructures within FoF halos.

### 2.3.1 Multiphase Star formation

The mechanism by which star-forming gas particles transform into star-type particles is based on the algorithm described in Springel & Hernquist (2003) and first developed by Katz et al. (1996). The formation of stars is a stochastic process (Tornatore et al., 2007). At a given time, the star formation rate (SFR) of a gas particle is computed with the Schmidt law (Kennicutt, 1998):

$$\dot{m}_\star = \frac{xm}{t_\star}, \quad (2.16)$$

where  $x$  is the fraction of cool gas, hence  $xm$  is the mass in cold clouds that contributes to the star formation. The time-scale for the star formation process  $t_\star(\rho)$  is given by:

$$t_\star(\rho) = t_0^\star \left( \frac{\rho}{\rho_{\text{th}}} \right)^{-1/2}. \quad (2.17)$$

Here, the star-formation time-scale,  $t_0^\star$ , and  $\rho_{\text{th}}$  (the density threshold above which star formation can occur) are parameters of the model. The variable  $t_0^\star$  is chosen to match the observed relation between the disk-averaged SFR per unit area and the gas surface density, according to the Kennicutt law (Kennicutt, 1998). Following Springel & Hernquist (2003), the numerical value of  $t_0^\star$  is 1.5 Gyr.

From (2.16), the stellar mass expected to form in an interval  $\Delta t$  is:

$$m_\star = m \left[ 1 - \exp \left( -\frac{x\Delta t}{t_\star} \right) \right]. \quad (2.18)$$

The number of stellar generations  $N_\star$  is defined as the number of star particles produced by a single gas particle. If  $m_{g,0}$  is the initial mass of a gas particle, the mass of a star

particle is given by:

$$m_{\star,0} = \frac{m_{g,0}}{N_{\star}}. \quad (2.19)$$

A multiphase gas particle generates a star particle if a random number drawn in the range  $[0, 1]$  is lower than the probability:

$$p = \frac{m_{g,0}}{m_{\star,0}} \left[ 1 - \exp\left(-\frac{x\Delta t}{t_{\star}}\right) \right]. \quad (2.20)$$

Once many star formation events have taken place, the stochastic SFR converges to its continuous history.

### 2.3.2 Chemical enrichment

P-GADGET3(XXL) self-consistently follows the evolution of Hydrogen, Helium and 9 metal elements (C, Ca, O, N, Ne, Mg, S, Si and Fe) released from supernovae (SNIa and SNII) and low and intermediate mass stars. The chemical evolution scheme is based on the stochastic star formation model implemented in the simulations (Tornatore et al., 2007), and described in the previous subsection. It accounts for the age of stars of different mass, such that the amount of metals released over time varies with the mass of the stars. A simple stellar population (SSP) -a star particle with the same initial metallicity- has a mass in the range  $0.1 M_{\odot} \leq m \leq 100 M_{\odot}$ . Stars with mass  $m \leq 40 M_{\odot}$  explode as SNe before turning into a black hole, while stars that have masses above this threshold collapse into a black hole without passing through the SN stage. The evolution of a SSP is described once a lifetime function, stellar yields and initial mass function (IMF) are selected.

The lifetime function from Padovani & Matteucci (1993)<sup>5</sup> for stars with mass  $m$  is adopted. The stellar yields quantify the amount of different metals which is released during the evolution of the stellar population, as follows:

- SNIa: Thielemann et al. (2003). The mass range for the SNIa originating from binary systems is  $0.8 M_{\odot} < m \leq 8 M_{\odot}$ , with a binary fraction of 7 per cent.
- SNII: Woosley & Weaver (1995). The mass range for the SNII is  $m > 8 M_{\odot}$ .
- Low and Intermediate Mass Stars: van den Hoek & Groenewegen (1997).

On the other hand, the initial stellar mass function determines the ratio between SNII and SNIa, and thus, the abundance of  $\alpha$ -elements and Fe-peak elements. The shape of the

---

<sup>5</sup>Further details of the lifetime function can be found in (Matteucci, 2003).

IMF establishes how many long-lived stars form with respect to massive short-lived stars. As a consequence, the ratio of long-to-short lived stars determines the energy released by SN and the current luminosity of galaxies which is dominated by low-mass stars. More precisely, the IMF  $\xi(m)$  determines the distribution of stellar masses  $dN$  per logarithmic mass interval that form in one star formation event in a certain volume:

$$\xi(m) = \frac{dN}{d \log m}. \quad (2.21)$$

The IMF is commonly expressed as a power law, as suggested originally by Salpeter (1955):

$$\xi(m) = A m^{-x}, \quad (2.22)$$

with  $A$ , the normalization factor, given by the condition:

$$\int_{m_{\text{inf}}}^{m_{\text{sup}}} \xi(m) dm = 1, \quad (2.23)$$

In this thesis,  $m_{\text{inf}} = 0.1 M_{\odot}$  and  $m_{\text{sup}} = 100 M_{\odot}$ . Different IMFs can be considered:

- Salpeter (1955): single sloped,

$$\xi(m) = 0.172 \times m^{-1.35} \quad (2.24)$$

- Kroupa et al. (1993): multi sloped,

$$\xi(m) = \begin{cases} 0.579 \times m^{-0.3} & 0.1 M_{\odot} \leq m < 0.5 M_{\odot} \\ 0.310 \times m^{-1.2} & 0.5 M_{\odot} \leq m < 1 M_{\odot} \\ 0.310 \times m^{-1.7} & m \geq 1 M_{\odot} \end{cases} \quad (2.25)$$

- Chabrier (2003): multi sloped<sup>6</sup>,

$$\xi(m) = \begin{cases} 0.497 \times m^{-0.2} & 0.1 M_{\odot} \leq m < 0.3 M_{\odot} \\ 0.241 \times m^{-0.8} & 0.3 M_{\odot} \leq m < 1 M_{\odot} \\ 0.241 \times m^{-1.3} & m \geq 1 M_{\odot} \end{cases} \quad (2.26)$$

The scheme implemented by Tornatore et al. (2007) allows for different IMFs, as well as the study of IMFs that change with time. Tornatore et al. (2007) showed that the

---

<sup>6</sup>The original Chabrier (2003) IMF is a power-law for stellar masses  $m \geq 1 M_{\odot}$  and has a log-normal form at low masses.

choice of the IMF produces different enrichment patterns, due to the fact that each IMF predicts a different number of a given mass-type of stars.

### 2.3.3 Galactic feedback

Galactic scale winds were introduced in GADGET simulations by Springel & Hernquist (2003) to regulate the star formation, spread metals from the galaxies to the IGM and prevent the overcooling of gas. Each form of feedback contribute differently, depending on the time of the Universe when the number density of galaxies raises, and their wind efficiency. Kinetic supernova-driven winds are a form of galactic feedback, particularly important at high redshift when AGNs do not play a major role, due to the low number of massive halos at early times.

#### Energy-driven winds

The phenomenological model for this type of wind feedback is presented in Springel & Hernquist (2003). Assuming that the mass-loss rate associated with the winds  $\dot{M}_w$  is proportional to the star formation rate  $\dot{M}_\star$ :

$$\dot{M}_w = \eta \dot{M}_\star, \quad (2.27)$$

with  $\eta$ , the wind mass loading factor that accounts for the efficiency of the wind. The wind carries a fixed fraction  $\chi$  of the SN energy. The kinetic energy of the wind is defined by the energy input from the supernova,

$$\frac{1}{2} \dot{M}_w v_w^2 = \chi \epsilon_{SN} \dot{M}_\star, \quad (2.28)$$

such that the wind velocity is:

$$v_w = \sqrt{\frac{2\chi\epsilon_{SN}}{\eta}}. \quad (2.29)$$

Stochastically, star-forming gas particles are designated to be part of the wind. While a gas particle is tagged as “wind”, it is completely decoupled from the hydrodynamics for a time  $t_{\text{dec}}$ . The maximum decoupling time depends both on the wind free travel length  $l_w$  and the wind velocity, given by equation (2.29), such that  $t_{\text{dec}} = l_w/v_w$ . At a given time  $t < t_{\text{dec}}$ , the gas particle can re-couple to the hydrodynamical interactions if its density is  $\rho < \delta_w \rho_{\text{th}}$ , where  $\rho_{\text{th}}$  is the density threshold for star-forming particles (see equation (2.17)) and  $\delta_w$ , the wind-free travel density factor (Tescari et al., 2014).

If the velocity of the wind is now chosen to match observational results from Martin (2005), it depends on the mass of the host halo and  $R_{200}$ , the radius where the density is 200 times larger than the critical density at redshift  $z$ :

$$v_w = 2\sqrt{\frac{GM_h}{R_{200}}} = 2 \times v_{\text{circ}}. \quad (2.30)$$

Here,  $v_{\text{circ}}$  is the halo circular velocity. The term  $R_{200}$  is defined as:

$$R_{200} = \sqrt[3]{\frac{3}{4\pi} \frac{M_h}{200\rho_c\Omega_{0m}} (1+z)^{-1}}, \quad (2.31)$$

with  $\Omega_{0m}$  and  $\rho_c$  the matter density and the critical density today.

Due to the conservation of the wind energy, the wind mass-loading factor  $\eta$  scales as the square of the inverse of  $v_w$ ,

$$\eta = 2 \times \left(\frac{v_{\text{fid}}}{v_w}\right)^2, \quad (2.32)$$

with  $v_{\text{fid}}$  a free parameter of the model that can be calibrated to account for different efficiencies.

### Momentum-driven winds

As discussed by Puchwein & Springel (2013), the mass carried by the wind is not necessarily proportional to the SFR of the galaxy (2.27). Hence, the radiative cooling loss can depend on the mass of the galactic halo. In such case, it would be more natural to assume that there is a relation between the momentum flux (instead of the energy flux) of the wind and the SFR of the galaxy, such that  $\eta$  is proportional to the inverse of the wind velocity  $v_w$ :

$$\eta = 2 \times \frac{v_{\text{fid}}}{v_w}. \quad (2.33)$$

The velocity of the wind  $v_w$  has the same functional form as the one described in the energy-driven winds case (2.30). However, what makes completely different these two models is the efficiency of the winds.

#### 2.3.4 AGN feedback

Although AGN feedback is not included in the simulations used in this thesis, the model is briefly explained for completeness, as part of the features of P-GADGET3(XXL).

This model of AGN feedback was first described by Springel et al. (2005a), with a few modifications explored in Fabjan et al. (2010); Planelles et al. (2013). A different treatment is proposed in Booth & Schaye (2009) and a comparative study between different AGN feedback mechanisms in Wurster & Thacker (2013). In this particular model, once a star-type particle reaches the critical mass, it turns stochastically into a super-massive black hole (SMBH) with mass  $M_{\text{seed}}$  that releases energy from the gas accretion process and contributes to the overall feedback. The black hole grows while gas is accreted into it at a Bondi rate, and the Eddington limit, as:

$$\dot{M}_{\text{SMBH}} = \min(\dot{M}_B, \dot{M}_{\text{Edd}}). \quad (2.34)$$

where  $\dot{M}_B$  and  $\dot{M}_{\text{Edd}}$  are the Bondi (Bondi, 1952) and Eddington accretion rates, respectively. In the model, it is assumed that when the accretion occurs, each gas particle contributes with a quarter of its mass.

The energy radiated in terms of the energy associated with the accreted mass is given by:

$$\epsilon_r = \frac{L_r}{\dot{M}_{\text{SMBH}}c^2}, \quad (2.35)$$

with  $\epsilon_r$ , the radiative efficiency of the SMBH. Following Tescari et al. (2014), the energy radiated is thermally coupled with the gas by:

$$\dot{E}_{\text{feed}} = \epsilon_f \epsilon_r \dot{M}_{\text{SMBH}}c^2, \quad (2.36)$$

where  $\epsilon_f$  is the fraction of energy that is coupled (or feedback efficiency) and its value has been set in the model as 0.2 (radio mode) and 0.05 (quasar mode).

### 2.3.5 Molecular and low-temperature metal cooling

Chemical pollution caused by star formation contributes to the cooling of gas. Some metal line cooling efficiencies peak at  $T \sim 10^4$  K (mostly low ionization transitions, Gnat & Ferland, 2012). These transitions are privileged in metal poor high-density environments, as DLAs, where H is mostly neutral or in its molecular form.

As discussed in Maio et al. (2007, 2010), molecular and low temperature metal cooling is particularly important when collapsed structures reach temperatures  $T < 10^4$  K due to the formation of molecules. At this temperature, atomic cooling is not efficient and highly

suppressed, yet, molecular H carries on cooling the gas.

A complete non-equilibrium chemical treatment is presented in Maio & Tescari (2015) for  $e^-$ , H,  $H^+$ ,  $H^-$ , He,  $He^+$ ,  $He^{++}$ ,  $H_2$ ,  $H_2^+$ , D,  $D^+$ , HD,  $HeH^+$ , following the reaction network shown in Table 2.1. The temporal evolution for each species  $i$  in terms of the number density  $n_i$  and the temperature  $T$  is given by:

$$\frac{dn_i}{dt} = \sum_p \sum_q k_{pq,i}(T)n_q n_p - \sum_l k_{li}(T)n_l n_i, \quad (2.37)$$

with  $dt$  the time interval,  $k_{pq,i}(T)$  is the creation coefficient of species  $i$  as a result of the interchange of species  $p$  and  $q$  and  $k_{li}(T)$ , the destruction coefficient of the species  $i$  due to their interaction with the species  $l$ . These coefficients are temperature-dependent.

Due to the non-equilibrium scheme, each abundance is computed at  $\sim 1/10$  of the hydrodynamical timestep. This condition assures convergence in the calculations for molecular gas.

Findings from Maio & Tescari (2015) show that the presence of molecules not only contribute to the cooling process (in particular at high redshift), but is also a signature of dense material at low temperature. Once star formation takes place, it causes local heating and chemical enrichment in the nearby regions. The UV and metal feedback from the new born stars is the main cause for the destruction of the Hydrogen molecules and metal pollution in the ISM and CGM.

Table 2.1 Reaction network taken from Maio &amp; Tescari (2015).

Reactions	References for the rate coefficients
$\text{H} + \text{e}^- \rightarrow \text{H}^+ + 2\text{e}^-$	A97 / Y07 / M07
$\text{H}^+ + \text{e}^- \rightarrow \text{H} + \gamma$	A97 / Y07 / M07
$\text{H} + \gamma \rightarrow \text{H}^+ + \text{e}^-$	A97 / Y07 / M07
$\text{He} + \text{e}^- \rightarrow \text{He}^+ + 2\text{e}^-$	A97 / Y07 / M07
$\text{He}^+ + \text{e}^- \rightarrow \text{He} + \gamma$	A97 / Y07 / M07
$\text{He} + \gamma \rightarrow \text{He}^+ + \text{e}^-$	A97 / Y07 / M07
$\text{He}^+ + \text{e}^- \rightarrow \text{He}^{++} + 2\text{e}^-$	A97 / Y07 / M07
$\text{He}^{++} + \text{e}^- \rightarrow \text{He}^+ + \gamma$	A97 / Y07 / M07
$\text{He}^+ + \gamma \rightarrow \text{He}^{++} + \text{e}^-$	A97 / Y07 / M07
$\text{H} + \text{e}^- \rightarrow \text{H}^- + \gamma$	GP98 / Y07 / M07
$\text{H}^- + \gamma \rightarrow \text{H} + \text{e}^-$	A97 / Y07 / M07
$\text{H}^- + \text{H} \rightarrow \text{H}_2 + \text{e}^-$	GP98 / Y07 / M07
$\text{H} + \text{H}^+ \rightarrow \text{H}_2^+ + \gamma$	GP98 / Y07 / M07
$\text{H}_2^+ + \gamma \rightarrow 2\text{H}^+ + \text{e}^-$	A97 / Y07 / M07
$\text{H}_2^+ + \gamma \rightarrow \text{H} + \text{H}^+$	A97 / Y07 / M07
$\text{H}_2^+ + \text{H} \rightarrow \text{H}_2 + \text{H}^+$	A97 / Y07 / M07
$\text{H}_2 + \text{H} \rightarrow 3\text{H}$	A97 / M07
$\text{H}_2 + \text{H}^+ \rightarrow \text{H}_2^+ + \text{H}$	S04 / Y07 / M07
$\text{H}_2 + \text{e}^- \rightarrow 2\text{H} + \text{e}^-$	ST99 / GB03 / Y07 / M07
$\text{H}^- + \text{e}^- \rightarrow \text{H} + 2\text{e}^-$	A97 / Y07 / M07
$\text{H}^- + \text{H} \rightarrow 2\text{H} + \text{e}^-$	A97 / Y07 / M07
$\text{H}^- + \text{H}^+ \rightarrow 2\text{H}$	P71 / GP98 / Y07 / M07
$\text{H}^- + \text{H}^+ \rightarrow \text{H}_2^+ + \text{e}^-$	SK87 / Y07 / M07
$\text{H}_2^+ + \text{e}^- \rightarrow 2\text{H}$	GP98 / Y07 / M07
$\text{H}_2^+ + \text{H}^- \rightarrow \text{H} + \text{H}_2$	A97 / GP98 / Y07 / M07
$\text{H}_2 + \gamma \rightarrow \text{H}_2^+ + \text{e}^-$	A97 / Y07 / M07
$\text{H}_2 + \gamma \rightarrow 2\text{H}$	A97 / R01 / Y03 / M07
$\text{D} + \text{H}_2 \rightarrow \text{HD} + \text{H}$	WS02 / M07
$\text{D}^+ + \text{H}_2 \rightarrow \text{HD} + \text{H}^+$	WS02 / M07
$\text{HD} + \text{H} \rightarrow \text{D} + \text{H}_2$	SLP98 / M07
$\text{HD} + \text{H}^+ \rightarrow \text{D}^+ + \text{H}_2$	SLP98 / M07
$\text{H}^+ + \text{D} \rightarrow \text{H} + \text{D}^+$	S02 / M07
$\text{H} + \text{D}^+ \rightarrow \text{H}^+ + \text{D}$	S02 / M07
$\text{D}^+ + \text{e}^- \rightarrow \text{D} + \gamma$	GP98
$\text{D} + \gamma \rightarrow \text{D}^+ + \text{e}^-$	GP98
$\text{He} + \text{H}^+ \rightarrow \text{HeH}^+ + \gamma$	RD82 / GP98 / M07
$\text{HeH}^+ + \text{H} \rightarrow \text{He} + \text{H}_2^+$	KAH79 / GP98 / M07
$\text{HeH}^+ + \gamma \rightarrow \text{He} + \text{H}^+$	RD82 / GP98 / M07

Notes:  $\gamma$  stands for photons; P71 = Peterson et al. (1971); KAH79 = Karpas et al. (1979); RD82 = Roberge & Dalgarno (1982); SK87 = Shapiro & Kang (1987); A97 = Abel et al. (1997); GP98 = Galli & Palla (1998); SLP98 = Stancil et al. (1998); ST99 = Stibbe & Tennyson (1999); R01 = Ricotti et al. (2001); WS02 = Wang & Stancil (2002); S02 = Savin (2002); GB03 = Glover & Brand (2003); Y03 = Yoshida et al. (2003); S04 = Savin et al. (2004); Y07 = Yoshida et al. (2007); M07 = Maio et al. (2007).

## 2.4 Overview of the simulations

For the purpose of this thesis, a set of hydrodynamical simulations were run using the Raijin distributed-memory cluster from the National Computational Infrastructure (NCI) facility<sup>7</sup>, located at the ANU in Canberra. The assumed cosmology is a flat  $\Lambda$ CDM model with cosmological parameters from the latest release of the Planck Collaboration et al. (2015),  $\Omega_{0m} = 0.307$ ,  $\Omega_{0b} = 0.049$ ,  $\Omega_{\Lambda} = 0.693$ ,  $n_s = 0.967$ ,  $H_0 = 67.74 \text{ km s}^{-1}\text{Mpc}^{-1}$  (or  $h = 0.6774$ ) and  $\sigma_8 = 0.816$ . An overview of the numerical simulations is shown in Table 2.2.

Table 2.2 Summary of the simulations used in this work. Column 1: run name. Column 2: box size. Column 3: Plummer-equivalent comoving gravitational softening length. Columns 4 and 5: mass of gas and dark matter particles. All the simulations have the same initial number of gas and DM particles ( $2 \times 512^3$ ). Column 6: feedback model. Column 7: inclusion of low-temperature metal and molecular cooling (Maio et al., 2007; Maio & Tescari, 2015). The first run, Ch 18 512 MDW, is the fiducial model. The second one in the list, Ch 18 512 MDW mol, has exactly the same configuration as the reference run, but includes low-T metal and molecular cooling.

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

### Galaxy properties in our simulations

P-GADGET3 (XXL) was first tested in the context of the ANGUS project. In Tescari et al. (2014) and Katsianis et al. (2015, 2016, 2017), the authors showed that their simulations are compatible with observations of the cosmic star formation rate density and the galaxy stellar mass function at  $1 < z < 7$ . However, the set of simulations presented here differs from the ones mentioned above by the normalization of the mass loading factor for SN-driven winds, set to  $v_{\text{fid}} = 600 \text{ km/s}$  in this work (see subsection 2.3.3). Since this suite of numerical simulations does not include AGN feedback (and in particular the “early AGN feedback in low-mass galaxies” model), the galactic wind model is recalibrated to match the cosmic star formation rate density history and galaxy stellar mass function up to  $z = 8$ <sup>8</sup>.

<sup>7</sup><http://nci.org.au/systems-services/peak-system/raijin/>

<sup>8</sup>Tescari et al. (2014) considers three cases of EDW: weak, strong and very strong winds, with constant  $\eta = 2$  and velocities  $v_{\text{wind}} = 350, 450$  and  $550 \text{ km/s}$ , respectively, and  $v_{\text{fid}} = 450 \text{ km/s}$  in the MDW feedback case. They also include AGN feedback.

The cosmic star formation rate density is the mass of new stars per unit volume per year and is measured from galaxy surveys and luminosity functions. The SFR can be calculated in two different ways in our simulations. With the first method, the star formation rate of each gas particle is counted. The SFR density is the sum of the SFR contribution from all gas particles per comoving volume. The comoving volume is given by  $V = (\text{box size}(\text{Mpc}/h))^3$ . The resulting SFR density is shown in Figure 2.3.

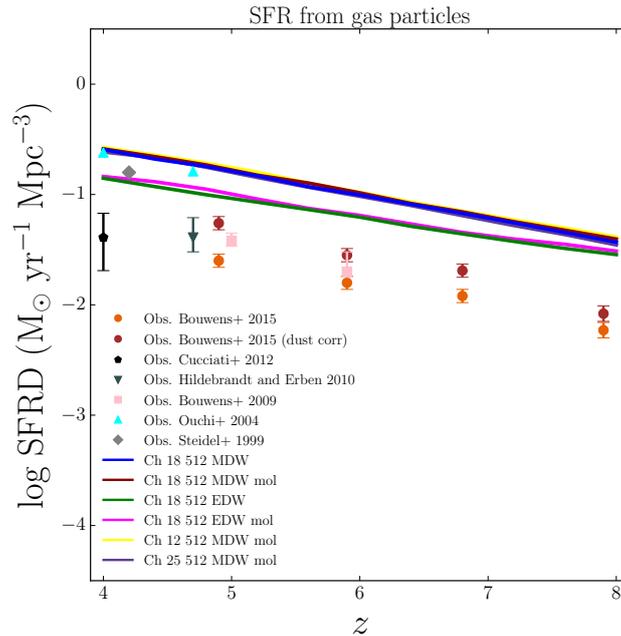


Figure 2.3 Evolution of the cosmic star formation rate density at  $4 < z < 8$  and comparison with observations by Bouwens et al. (2015) in brown and orange circles (with and without dust corrections, respectively), Cucciati et al. (2012) in black pentagon, Hildebrandt et al. (2010) in green inverted triangle, Bouwens et al. (2009) in pink square, Ouchi et al. (2004) in cyan triangles and Steidel et al. (1999) in grey diamond. We sum the SFR in each gas particle in the simulation and divide by the comoving volume. Hereafter, these colors are used to represent the simulations.

Figure 2.3 shows an excess of the calculated SFR density with respect to the observations at all redshifts. This is due to the fact that the observed SFR density is measured just for the most luminous galaxies ( $M_{\text{UV}} < -17$ ), while the SFR in the gas particles takes into account all the objects in the box. A cut in the luminosity  $M_{\text{UV}} < -17$  corresponds to a minimum SFR  $> 0.331 M_{\odot}/\text{yr}$ , calculated from FoF objects. When this criterion is imposed on the simulations, the calculated SFR density is in good agreement with available observations at these redshifts, as shown in Figure 2.4. The type of wind feedback plays an important role in the formation of stars (see subsection 2.3.3). If the feedback mechanism does not prevent the overcooling of the gas, more stars will be formed. In this

sense, energy-driven winds (EDW) quench the SFR more effectively than momentum-driven winds (MDW). For a further discussion see Tescari et al. (2014).

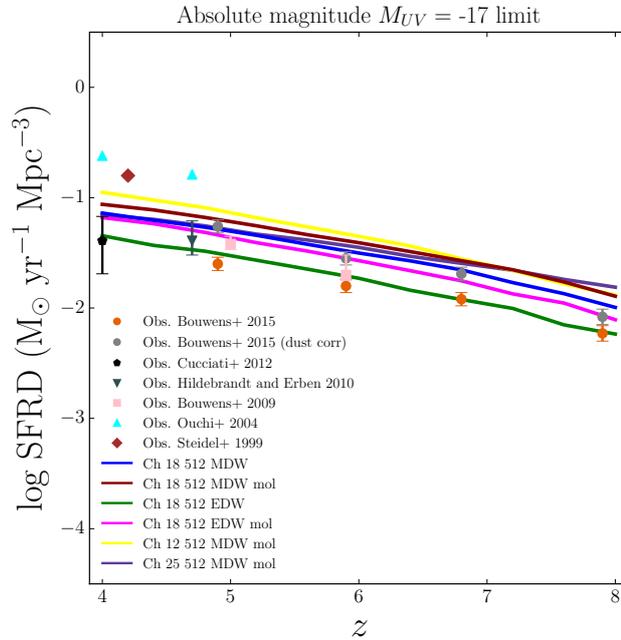


Figure 2.4 Evolution of the cosmic star formation rate density at  $4 < z < 8$  and comparison with observations by Bouwens et al. (2015) in brown and orange circles (with and without dust corrections, respectively), Cucciati et al. (2012) in black pentagon, Hildebrandt et al. (2010) in green inverted triangle, Bouwens et al. (2009) in pink square, Ouchi et al. (2004) in cyan triangles and Steidel et al. (1999) in grey diamond. To provide better match with the observations, the SFR calculation is limited to FoF objects with  $\text{SFR} > 0.331 M_{\odot}/\text{yr}$ .

In addition, the galaxy stellar mass function is calculated using galaxies above the mass resolution limit<sup>9</sup>. The stellar mass function  $\Phi(z)$  counts the number of galaxies per mass bin  $\Delta M$  per comoving volume  $V$ .

$$\Phi(z) = \frac{\#_{\text{gal}}(\Delta M)}{V \cdot \Delta M}, \quad (2.38)$$

Figures 2.5 and 2.6 show the stellar mass functions at  $z = 8$  and  $6$ , and a comparison with observational data by Song et al. (2016) and González et al. (2011).

<sup>9</sup>The mass resolution limit determines whether a galaxy has a resolved halo with mass equivalent to  $\sim 470$  dark matter particles ( $M_{\text{FoF}} = 10^{9.1} M_{\odot}$  for cosmological simulations with box-sizes of 18 Mpc/h). If the mass of the galactic halo is below that threshold, the object is considered unresolved and it is dismissed in the analysis.

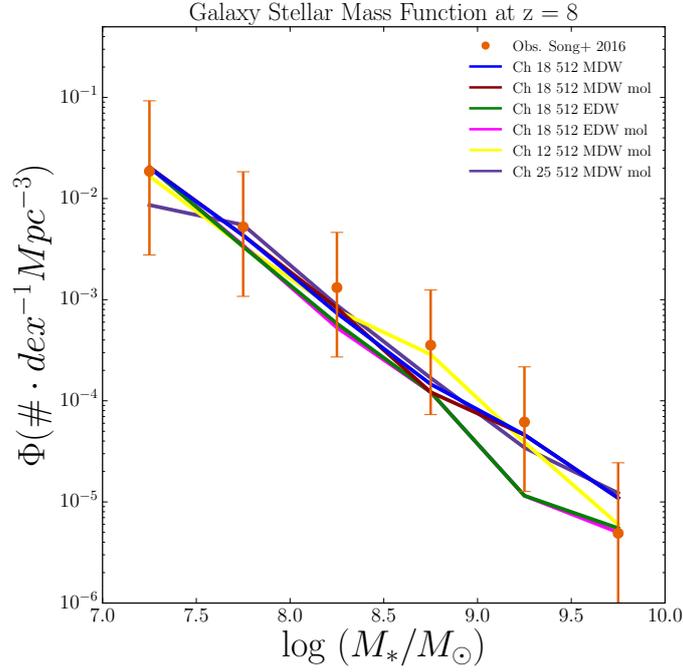


Figure 2.5 Simulated galaxy stellar mass function at  $z = 8$  and comparison with observations by Song et al. (2016) in orange circles.

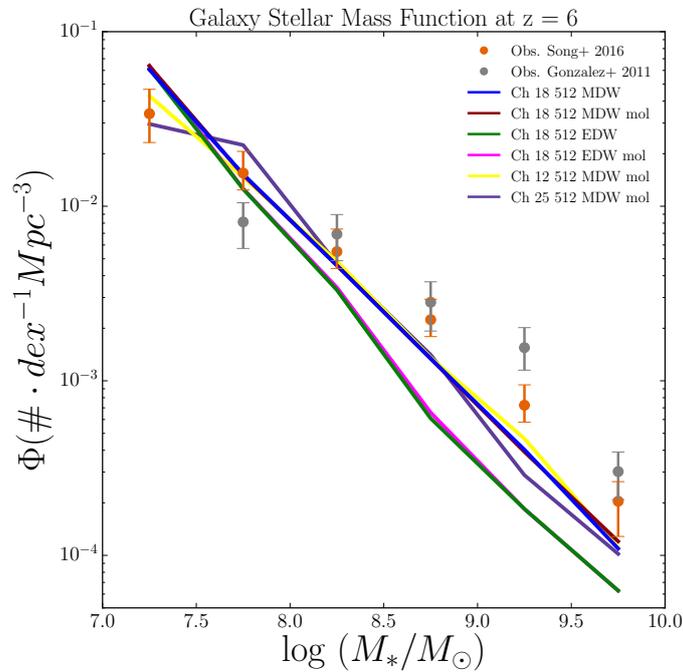


Figure 2.6 Simulated galaxy stellar mass function at  $z = 6$  and comparison with observations by Song et al. (2016) in orange circles and González et al. (2011) in grey circles.

The simulated galaxy stellar mass function at  $z = 8$  and  $6$  is compatible with the observations at high redshift. Nonetheless, the simulations differ from the observations in the high mass end at  $z = 6$ , due to the underproduction of high-mass galaxies in the simulations. This is mostly caused by the volume that is probed with the numerical runs and the fact that massive galactic halos are quite rare at high redshift.

### Chemical enrichment in the simulations

As discussed in subsection 2.3.2, one of the most important components of P-GADGET3 (XXL) is the self-consistent chemical enrichment module. The assumed stellar lifetime function, stellar yields and IMF determine the metal abundance predicted by the simulations. In this thesis, we concentrate on the evolution of three metals: Oxygen, Silicon and Carbon (the latter is extensively discussed in Chapter 3).

Figures 2.7 and 2.8 show the cosmic evolution of the Oxygen and Silicon mass densities, from  $z = 8$  to  $z = 4$ , when an important part of the chemical pollution has occurred in the Universe from stars and supernovae. The curves have been calculated at each redshift in the simulation, by summing the amount of Oxygen (or Silicon) in each gas particle in the simulation and dividing by the comoving volume.

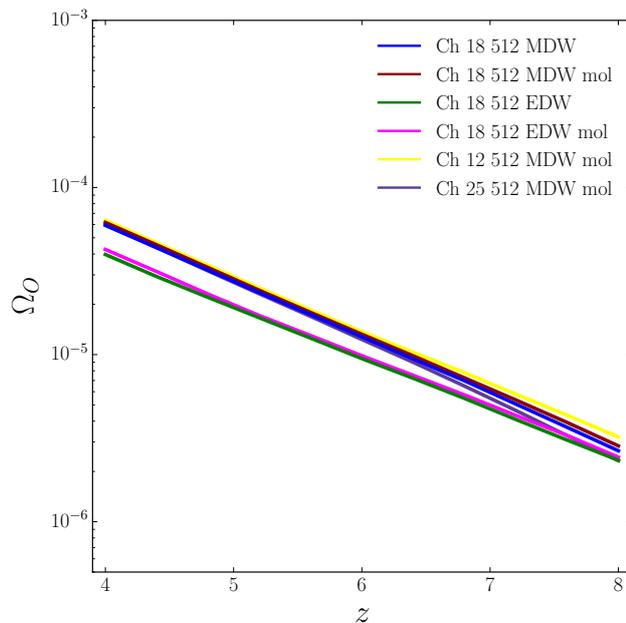


Figure 2.7 Evolution of the total Oxygen cosmological mass density  $\Omega_{\text{O}}$  for our simulations.

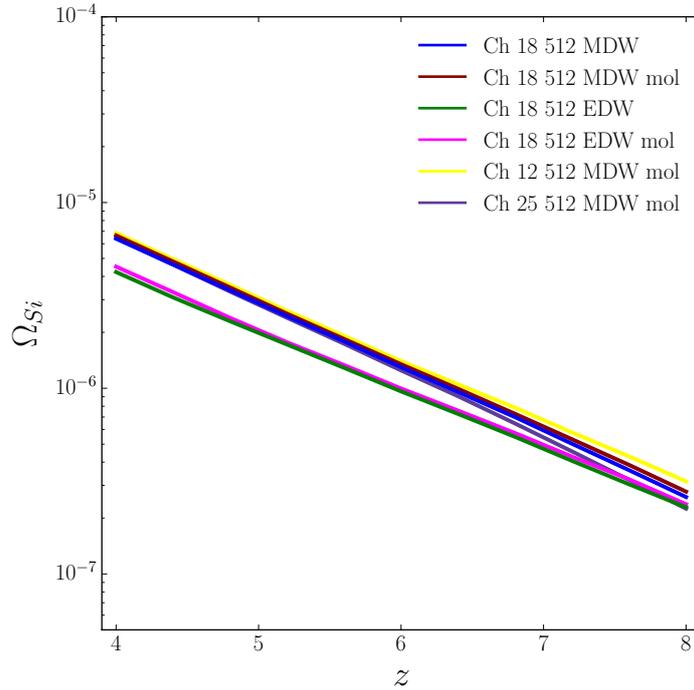


Figure 2.8 Evolution of the total Silicon cosmological mass density  $\Omega_{\text{Si}}$  for our simulations.

An interesting analysis arises while comparing Figures 2.7 and 2.8 with Figure 14 from (Madau & Dickinson, 2014), that shows the mean metallicity of the Universe. In Madau & Dickinson (2014), the mass of heavy metals per baryon density produced over the cosmic history with a given SFR model and an assumed IMF-averaged yield of  $y = 0.02$  is displayed with the solid curve. In the range of  $z = 4 - 7$ , the simulated metal abundances over this period grows 1 dex with respect to the solar metallicity, in concordance with Madau & Dickinson (2014) findings. This is a clear indication that the models used in this work describe well the cosmic chemical enrichment, driven by stars and supernovae.

To date, there is not a method to measure, directly or indirectly, the total mass density of an element  $\Omega_X$  in the Universe, especially at high redshift. Estimates of the relative amount of chemical elements with respect to Hydrogen have been made in DLAs, yet, we still lack observational constrains on the cosmic mass density of O or Si (among other metals). Therefore, the results shown in Figures 2.7 and 2.8 are purely theoretical predictions.

---

## 2.5 Post-process

---

The set of hydrodynamical simulations presented in the last section traces the properties of the gas at redshifts where the Reionization is occurring. Also, these numerical runs successfully reproduce the properties of the gas in presence of the dark matter halos, the formation of galaxies and consequently, the chemical enrichment. However, a completely self-consistent description of Reionization must take into account the progress of ionizing fronts, recombinations, the position of sources and sinks of UV photons, etc. Unfortunately, the complete picture is computationally too expensive: one can sacrifice radiative transfer modelling and resolve self-consistently the physical processes in the IGM, or, viceversa, achieve high accuracy on the regions nearby the sources with RT in small boxes, at the expenses of a poor description of the large scale structure.

There are two general ways to simulate the Reionization of the Universe. In the first one, a fiducial model is proposed to describe the emission, propagation and absorption of ionizing radiation from the sources. The second standpoint ignores absorption and propagation of the UV photons and counts the number of photons per atom, without a full description of the progression of the ionizing fronts during the Reionization, as discussed by Volonteri & Gnedin (2009). In the latter case, a UV ionizing background is introduced to post-process the simulations, and its evolution with redshift is given with a photoionization rate  $\Gamma_{\text{phot}}$  and the UV emissivity. In this work, the assumed uniform UV background is Haardt & Madau (2012) (hereafter HM12).

### 2.5.1 The uniform UV background

The ionizing UV background is the effective contribution from all the photons released from different sources at a given time in a certain volume. The HM12 model used in this thesis, is an improved version of HM2001 (Haardt & Madau, 2001). HM12 describes the evolving spectrum of the UV/X-ray diffuse background, from radiative transfer results with the code CUBA. The latest release introduces new features:

- the sawtooth modulation of the background intensity from resonant line absorption in the Lyman series of H and He;
- the X-ray emission from the obscured and unobscured quasars that gives rise to the X-ray part of the background;
- a parameterization of the distribution in redshift and column density of intergalactic

absorbers that fits recent measurements of the mean free path of 1 Ryd photons (Meiksin, 2005; Bongiorno et al., 2007; Siana et al., 2008; Cowie et al., 2009; Willott et al., 2010);

- an accurate treatment of the photoionization structure of absorbers, which enters in the calculation of the He continuum opacity and recombination emissivity;
- the UV emission from star-forming galaxies at all redshifts.

The HM12 model also presents a galaxy UV emissivity that matches the cosmic star formation rate history, introduces a luminosity-weighted escape fraction for the radiation that ionized H, which increases rapidly with  $z$  and it also confirms results from other numerical simulations with respect to the value of the clumping factor (e.g. Pawlik et al. 2009).

The photoionization and photoheating rates of HM12 used to post-process the hydrodynamical simulations are given in Table 2.3.

The UVB spectrum is incorporated in the pipeline in two steps: first, the photoionization model with CLOUDY assumes a UVB grid as a function of redshift and energy to compute the ionization fractions of the chemical elements in the gas. Second, the photoionization rates from HM12 (presented in Table 2.3) are used to feed the self-shielding prescription with best-fit parameters from Rahmati et al. (2013).

In order to run CLOUDY models, the UVB data<sup>10</sup> is split in 60 redshift directories ( $0 < z < 15.9$ ), each of them containing a file with 575 rest-frame wavelengths ( $2.5 \times 10^{-4} < \lambda(\text{\AA}) < 1 \times 10^5$ ) and their corresponding emissivity  $J_\nu$  (erg/s/cm<sup>2</sup>/Hz/sr).

At each redshift, wavelengths are converted into energies (in Ryd) through the relation  $E = \frac{hc}{\lambda} \left( \frac{1\text{Ryd}}{13.6} \right)$ , where  $h$  is the Planck's constant,  $c$ , the speed of light,  $\lambda$ , the wavelength. We include an additional line with the normalized emissivity at 1 Ryd in any redshift file. At high redshift ( $z \geq 5.8$ ), the saw-tooth attenuation produces extremely low values in emissivity, so these values are re-assigned by a threshold of  $\log J_\nu = -55$ .

---

<sup>10</sup>available in [www.ucolick.org/~pmadau/CUBA/Media/UVB.out](http://www.ucolick.org/~pmadau/CUBA/Media/UVB.out).

Table 2.3 The cosmic background photoionization  $\Gamma_{\text{HI}}$  and photoheating  $\mathcal{H}_{\text{HI}}$  rates in HM12. Taken from (Haardt & Madau, 2012).

$z$	$\Gamma_{\text{HI}}$ ( $\text{s}^{-1}$ )	$\mathcal{H}_{\text{HI}}$ ( $\text{eV s}^{-1}$ )	$z$	$\Gamma_{\text{HI}}$ ( $\text{s}^{-1}$ )	$\mathcal{H}_{\text{HI}}$ ( $\text{eV s}^{-1}$ )
0.00	0.228E-13	0.889E-13	3.21	0.765E-12	0.302E-11
0.05	0.284E-13	0.111E-12	3.42	0.705E-12	0.279E-11
0.10	0.354E-13	0.139E-12	3.64	0.647E-12	0.257E-11
0.16	0.440E-13	0.173E-12	3.87	0.594E-12	0.236E-11
0.21	0.546E-13	0.215E-12	4.11	0.546E-12	0.218E-11
0.27	0.674E-13	0.266E-12	4.36	0.504E-12	0.202E-11
0.33	0.831E-13	0.329E-12	4.62	0.469E-12	0.189E-11
0.40	0.102E-12	0.405E-12	4.89	0.441E-12	0.178E-11
0.47	0.125E-12	0.496E-12	5.18	0.412E-12	0.167E-11
0.54	0.152E-12	0.605E-12	5.49	0.360E-12	0.148E-11
0.62	0.185E-12	0.734E-12	5.81	0.293E-12	0.123E-11
0.69	0.223E-12	0.885E-12	6.14	0.230E-12	0.989E-12
0.78	0.267E-12	0.106E-11	6.49	0.175E-12	0.771E-12
0.87	0.318E-12	0.126E-11	6.86	0.129E-12	0.583E-12
0.96	0.376E-12	0.149E-11	7.25	0.928E-13	0.430E-12
1.05	0.440E-12	0.175E-11	7.65	0.655E-13	0.310E-12
1.15	0.510E-12	0.203E-11	8.07	0.456E-13	0.219E-12
1.26	0.585E-12	0.232E-11	8.52	0.312E-13	0.153E-12
1.37	0.660E-12	0.262E-11	8.99	0.212E-13	0.105E-12
1.49	0.732E-12	0.290E-11	9.48	0.143E-13	0.713E-13
1.61	0.799E-12	0.317E-11	9.99	0.959E-14	0.481E-13
1.74	0.859E-12	0.341E-11	10.50	0.640E-14	0.323E-13
1.87	0.909E-12	0.360E-11	11.10	0.427E-14	0.217E-13
2.01	0.944E-12	0.374E-11	11.70	0.292E-14	0.151E-13
2.16	0.963E-12	0.381E-11	12.30	0.173E-14	0.915E-14
2.32	0.965E-12	0.382E-11	13.00	0.102E-14	0.546E-14
2.48	0.950E-12	0.375E-11	13.70	0.592E-15	0.323E-14
2.65	0.919E-12	0.363E-11	14.40	0.341E-15	0.189E-14
2.83	0.875E-12	0.346E-11	15.10	0.194E-15	0.110E-14
3.02	0.822E-12	0.325E-11			

### 2.5.2 Photoionization modelling with CLOUDY

A photoionization model studies the interaction of ionizing continuum radiation from an energetic source into a plasma, as well as the progression of the ionizing photons through the internal structure of the diffuse gas. Due to the several physical reactions that can be triggered by the radiation in the gas and the large number of particles that constitute the gas, the photoionization treatment can be computationally extremely expensive. Most photoionization models to date describe in an effective way the macroscopic phenomena, instead of focusing on the ionization of atoms and ions by single photons in the microscale.

The model chosen in this work is CLOUDY (Ferland et al., 1998, 2013, 2017). CLOUDY is an open source code that follows the chemical, thermal and ionization state of matter exposed to external radiation or a source of heating, in order to calculate the spectroscopic lines and other features that describe different astrophysical events<sup>11</sup>.

The conditions of the gas are set up for (non-)equilibrium with the external field of radiation (UV, optical or infrared) in a broad range of densities (up to  $10^{15} \text{ cm}^{-3}$ ) and temperatures (from CMB to  $10^{10}\text{K}$ ). The large extent of the parameter space allows the code to reproduce multiple astrophysical environments: stellar coronas, IGM, accretion disk nearby a super-massive BH in a luminous quasar, as well as the properties of ionized, atomic (in different level populations) and molecular gas in the presence of AGN radiation.

An ionizing background (shape and intensity of the external field striking the gas) is required as a start-up point. Others inputs to initialize the code are: the geometry of the cloud<sup>12</sup>, the chemical composition of the gas<sup>13</sup> and the velocity structure of the motions that broad the spectral lines (either thermal, turbulence or flows).

CLOUDY computes the abundance of each ionization state using the *two-level approximation*. According to Osterbrock & Ferland (2006), the ionization balance of an ion is equivalent to the two-level system (Ferland et al., 2017):

$$\frac{n(i+1)}{n(i)} = \frac{\Gamma(i)}{\alpha(i+1)n_e}, \quad (2.39)$$

<sup>11</sup>The sources files from different versions are available in [nublado.org](http://nublado.org).

<sup>12</sup>By default, the geometry assumed by the code is 1D spherical. A plane parallel configuration can be achieved by setting the inner radius much larger than the thickness of the gas.

<sup>13</sup>The code includes the 30 lightest elements, but the user decides which metals are tracked in the internal configuration. Molecules and grains are not considered in the plasma, unless this flag is activated.

with  $n(i+1)$  and  $n(i)$ , the number densities of two contiguous ionization states,  $\alpha(i+1)$  is the recombination rate coefficient of the ion in  $\text{cm}^3 \text{s}^{-1}$  and  $\Gamma(i)$  is the photoionization rate. In photoionization equilibrium (as in the case here studied), the photoionization rate is given by:

$$\Gamma(i) = \int \phi_\nu \sigma_\nu d\nu. \quad (2.40)$$

Here,  $\phi_\nu$  is the flux of ionizing photons, that is calculated from the UV emissivity input from HM12.  $\sigma_\nu$  is the photoionization cross section and the integral is calculated in the range of energies spanned in the UVB.

If instead, there is collisional ionization equilibrium,  $\Gamma(i)$  is expressed as:

$$\Gamma(i) = q(i)n_e, \quad (2.41)$$

where  $q(i)$  is the collisional ionization rate coefficient. The two-level approximation assumes that recombinations to all excited states will eventually decay to the ground state, and that all ionizations occur out of the ground state. Only the ionization rate from the ground state and the sum of recombination coefficients to all excited states need be considered (Ferland et al., 2017).

For doublets such as CIV, CLOUDY assumes the ionization / emission approximation, but this feature is not used in the density scales relevant to the IGM at the redshifts followed in this work. However, it is quite important in very high density regions.

For the purpose of this thesis, CLOUDY is used to calculate the ionization fractions of the metals: C, Ca, O, N, Ne, Mg, S, Si, Fe. The pipeline has been fixed to recover only the following ionic transitions: HI, CI, CII, CIII, CIV, CV, OI, OVI, SiII, SiIII, SiIV. The output from the code is a grid of temperature, density and ionization fractions, as a function of redshift. The code was run separately using the supercluster Raijin from the NCI facilities.

### 2.5.3 HI self-shielding

As the outskirts of the galaxies are approached, predominately ionized Hydrogen in the IGM/CGM becomes denser and starts to self-shield. This phenomenon has been observed in high density regions as Lyman limit systems (LLS, with column densities  $N > 10^{17} \text{cm}^{-2}$ ) and DLAs, with  $N \geq 2 \times 10^{20} \text{cm}^{-2}$ . Theoretical works from Keating & Miller (2006) confirms that HI starts to self-shield in the LLS regime, independently of the pho-

toionization model assumed.

When a medium is optically thick ( $\tau_{\text{HI}} = 1$ ), Hydrogen starts to be self-shielded with respect to the UVB radiation at a column density (Schaye, 2001):

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

The sub-index SSh stands for self-shielding.  $\hat{\sigma}_{\nu\text{HI}}$  corresponds to the cross section of the ionizing UVB radiation. The densities at which the HI self-shielding begins are (e.g. Furlanetto et al., 2005):

$$n_{\text{H,SSh}} \sim 6.73 \times 10^{-3} \text{cm}^{-3} \left( \frac{\hat{\sigma}_{\nu\text{HI}}}{2.49 \times 10^{-18} \text{cm}^2} \right)^{-2/3} \left( \frac{T}{10^4 \text{K}} \right)^{0.17} \left( \frac{\Gamma}{10^{-12} \text{s}^{-1}} \right)^{2/3} \left( \frac{\Omega_b/\Omega_m}{0.17} \right)^{1/2},$$

where  $\Gamma$  is the photoionization rate.

Different studies have focused their efforts on accurately calculate the total amount of HI in different structures, especially when the environments where HI is found are optically thick due to the large number of recombinations that occur in dense regimes. In particular, Rahmati et al. (2013) use cosmological simulations combined with RT calculations using the code TRAPHIC (a photon-conservative RT method coupled directly to the irregular distribution of SPH particles, Pawlik & Schaye, 2008, 2011) to describe the distribution of neutral Hydrogen at  $z = 0-5$ . The fitting functions provided by Rahmati et al. (2013) are extensively used in the literature to compute the HI fraction as a function of density and the best fitting parameters are included in post-process in multiple works.

As described in appendix A2 from Rahmati et al. (2013), if the number of ionizations per unit time per unit volume is equal to the total number of recombinations per unit time per unit volume, a state of ionization equilibrium is established:

$$n_{\text{HI}}\Gamma_{\text{tot}} = \alpha_A n_e n_{\text{HII}}, \quad (2.42)$$

with  $n_e$ ,  $n_{\text{HI}}$  and  $n_{\text{HII}}$  are the number densities of electrons, neutral and ionized Hydrogen, respectively. On the other hand,  $\Gamma_{\text{tot}}$  is the total ionization rate per neutral Hydrogen atom and  $\alpha_A$  is the Case A recombination rate (Hui et al., 1997),

$$\alpha_A = 1.269 \times 10^{-13} \frac{\lambda^{1.503}}{(1 + (\lambda/0.522)^{0.47})^{1.923}} \text{cm}^3 \text{s}^{-1}. \quad (2.43)$$

If the contribution from Helium is ignored and the neutral Hydrogen fraction  $\eta$  is defined as the ratio of the number densities of HI and total H,  $\eta = \frac{n_{\text{HI}}}{n_{\text{H}}}$ , the ionization equilibrium condition is now:

$$\eta\Gamma_{\text{tot}} = \alpha_A(1 - \eta^2)n_{\text{H}}, \quad (2.44)$$

where  $\Gamma_{\text{tot}} = \Gamma_{\text{phot}} + \Gamma_{\text{col}}$ , *i.e.* the total ionization rate has contributions from the total photoionization rate and total collisional rate (first and second terms, respectively). The collisional rate has a functional form  $\Gamma_{\text{col}} = \Lambda_T(1 - \eta)n_{\text{H}}$ , with  $\Lambda_T$ , the collisional ionization rate of HI, taken from Theuns et al. (1998):

$$\Lambda_T = 1.17 \times 10^{-10} \frac{T^{-1/2} \exp(-157809/T)}{1 + \sqrt{T/10^5}} \text{cm}^3 \text{s}^{-1}. \quad (2.45)$$

One can re-write the ionization equilibrium condition (2.44) as a quadratic equation:

$$A\eta^2 - B\eta + C = 0, \quad (2.46)$$

with  $A = \alpha_A + \Lambda_T$ ,  $B = 2\alpha_A + \frac{\Gamma_{\text{phot}}}{n_{\text{H}}} + \Lambda_T$  and  $C = \alpha_A$ .  $\Gamma_{\text{phot}}$  can be calculated using RT results, such that:

$$\Gamma_{\text{phot}} = \left( (1 - f) \left[ 1 + \left( \frac{n_{\text{H}}}{n_0} \right)^\beta \right]^{\alpha_1} + f \left[ 1 + \left( \frac{n_{\text{H}}}{n_0} \right) \right]^{\alpha_2} \right) \Gamma_{\text{UVB}}, \quad (2.47)$$

the photoionization rate as a function of redshift  $\Gamma_{\text{UVB}}$  is assumed from the UVB field. The number density  $n_{\text{H}}$  and the temperature  $T$  are taken directly from the numerical simulation used. The values of the parameters calculated with RT at  $1 < z < 5$  are  $\alpha_1 = -2.28 \pm 0.31$ ,  $\alpha_2 = -0.84 \pm 0.11$ ,  $\beta = 1.64 \pm 0.19$ ,  $n_0 = (1.003 \pm 0.005) n_{\text{H,SSh}}$  (where  $n_{\text{H,SSh}}$  is the self-shielding density threshold) and  $f = 0.02 \pm 0.0089$ .

The positive solution of the equation (2.46) gives the evolution of the neutral Hydrogen fraction in ionization equilibrium as a function of the gas density and temperature.

#### 2.5.4 VPFIT

Once the physical conditions of the baryonic gas are achieved with the hydrodynamical simulations and the evolution of the ions is imposed with a photoionization model, the next step is to produce as many mock spectra as possible to have a large statistical sample.

The quasar spectra are a collection of absorption features along the line of sight from

a distant quasar to the observer. The best way to fit these spectroscopic features is with multiple Voigt profiles. A Voigt profile is the convolution of two broadening mechanisms: a Gaussian (caused by a thermal phenomena) and a Lorentzian profile.

$$V(x; \sigma, \gamma) = \int_{-\infty}^{\infty} G(x'; \sigma) L(x - x'; \gamma) dx' \quad (2.48)$$

where  $x$  is the shift from the line centre,  $G(x; \sigma)$  a centred Gaussian profile and  $L(x; \gamma)$ , a Lorentzian centred profile.

VPFIT (Carswell & Webb, 2014) is an automatic algorithm that uses  $\chi^2$  fits for the absorption lines (until it reaches a tolerance threshold) and provides the most accurate estimate of the parameters of each profile. The fact that it automatically fits a extremely large number of spectra, makes it quite a powerful tool. VPFIT is used to fit Voigt profiles both in observational or synthetic spectra. the difference is that the mock spectrum of each ion is calculated separately, therefore, blending with other ionic transitions do not occur, as it is the case with observational quasar spectra. Mock spectra do not suffer from sky lines or any deficiency of the observing instrument.

Along with the number of spectra that can be handled by VPFIT, the main reason why the code was introduced in this pipeline is that observers use VPFIT to analyse their quasar spectra. In order to obtain the most faithful comparison with the observational sample at high redshift, we chose to use VPFIT. The pipeline proposed here closely mimics the one used by astronomers that routinely work with absorption features imprinted in quasar spectra. Thus, the systematic uncertainties between observational and simulated spectra are minimized by using the same analysis software.

VPFIT uses FITS files with the fluxes, flux errors estimates, wavelengths and continuum values. Additional files are required, for instance, relevant atomic data (oscillator strengths and rest-frame wavelengths) and an initial guess for each transition line and its corresponding absorbed wavelength, as well as the desired tolerance, the precision and number of iterations of the fitting.

The output file contains the calculated  $\chi^2$  per iteration, the stopping criteria and best fitting parameters for each component of the profile: number of features and column density, systemic redshift, doppler parameter and corresponding errors.

The list of ions used in this work and their corresponding physical features, is given in

Table 2.4. The metal abundance  $Z$  reported in the fifth column is calculated as follows:

$$\left[ \frac{X}{H} \right] = \log \frac{(n_X/n_H)}{(n_X/n_H)_\odot} = \log(n_X/n_H) - \log(n_X/n_H)_\odot. \quad (2.49)$$

Here, X is an arbitrary species and  $n_X$  is its corresponding number fraction. The logarithm are 10-based and the suscript  $\odot$  accounts for solar values.

Table 2.4 List of the ion lines included in this work and reduced with VPFIT. The first column contains the ions, the second one the rest-frame wavelength  $\lambda$  of the transition with the highest oscillator strength. The third column, the oscillator strength  $f$  of each absorption line, the fourth one shows the ionization energy  $E$  associated to each state, and the fifth column, the metal abundance  $\log Z$  (in solar units), taken from Asplund et al. (2009).

Ion $i$	$\lambda$ ( $\text{\AA}$ )	$f$	$E_{i \rightarrow i+1}$ (eV)	$\log Z$ ( $Z_\odot$ )
HI	1215.67	0.4164	13.6	0
CII	1334.53	0.1278	24.38	-3.57
CIII	977.020	0.7570	47.89	-3.57
CIV	1548.21	0.1899	64.49	-3.57
SiIII	1526.71	0.1330	16.35	-4.49
SiIV	1393.76	0.513	45.14	-4.49
OI	1302.17	0.0480	13.62	-3.31

Note: The energy  $E$  shown in the fourth column is the energy required to reach the next ionization state  $i + 1$  from the state  $i$ .



# 3

## Metal and HI absorption lines at high redshift

### 3.1 Abstract

---

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

### 3.2 Introduction

---

Understanding the epoch of Reionization (EoR) is one of the current challenges of extragalactic astronomy. It will complete the big picture of the thermal history of the Universe (e.g. McQuinn, 2016). The EoR commenced when the first stars switched on, producing UV radiation that transformed the neutral Hydrogen (HI) in the surrounding circumgalactic and intergalactic media (CGM and IGM) into ionized hydrogen (HII). Observations of

Lyman- $\alpha$  photons ( $\text{Ly}\alpha$  1216 Å) in absorption towards  $z_{\text{em}} \gtrsim 6$  quasars (QSOs) show fluctuations in flux consistent with Reionization concluding at  $z \lesssim 6$  (e.g. Fan, 2006; Becker et al., 2015b). Measuring Lyman series absorption or UV emissivity at wavelengths blueward of  $\text{Ly}\alpha$  becomes almost impossible at redshifts greater than 5.5 due to the increasing density of matter and neutral Hydrogen fraction (Becker et al., 2015a).

Metal absorption lines are an alternative proxy to the Lyman series for probing the evolution of the IGM during the EoR and offer many advantages at high redshift. For example, they can be detected even when Hydrogen is completely saturated in the spectra of the background quasars, since some of these transitions occur redward of  $\text{Ly}\alpha$  and are thus unaffected by Lyman series absorption and Gunn–Peterson troughs (Gunn & Peterson, 1965). Also, ionic transitions in the QSO spectra at high redshift provide important constraints on the model of the ionizing background that drove Reionization (e.g. Furlanetto & Mesinger, 2009; Becker et al., 2011, 2015b; Finlator et al., 2016). Low ionization transitions (OI, CII, SiII, MgII, FeII) trace the location of neutral Hydrogen at high redshift. On the other hand, high ionization states (CIV, SiIV, OVI) have a comparably larger ionization potential than H, therefore the energy required to produce these transitions is reached in regions where Hydrogen is highly ionized at early times.

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

An extensive observational effort has been pursued with CIV, the triply ionized state of Carbon. CIV offers many advantages that make its detection easier than other ionic species: large oscillator strength, wavelength redward of  $\text{Ly}\alpha$  emission and a doublet tran-

sition. Observations of CIV in the foreground of quasar spectra at high redshift by Songaila (2001, 2005), Pettini et al. (2003), Ryan-Weber et al. (2006, 2009), Simcoe (2006), Simcoe et al. (2011), D’Odorico et al. (2010, 2013, 2016) and Boksenberg & Sargent (2015) have built the largest sample of metal absorbers as a function of redshift to date. The cosmological mass density  $\Omega_{\text{CIV}} = \frac{\rho_{\text{CIV}}}{\rho_{\text{crit}}}$  (the density of CIV with respect to the critical density today) drops at high redshifts during the progression of the EoR, due to the changing ionization state of the IGM and a decrease of its metallicity.

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

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

Likewise, CII has been observed by Becker et al. (2006) and modelled by D’Odorico et al. (2013) at  $z = 5.7$  to be best fitted by low density gas,  $\delta = \frac{\rho_{\text{gas}}}{\langle \rho \rangle} - 1 = 10$ , where  $\langle \rho \rangle$  is the mean density at the considered redshift. Low ionization metals as CII and MgII are routinely detected at low redshifts in regions close to the centre of galaxies or in damped Ly $\alpha$  systems (DLAs, *i.e.* absorption systems with  $N_{\text{HI}} > 10^{20.3} \text{cm}^{-2}$ , Wolfe et al., 1986, 2005). These latter are optically thick structures that constitute the main reservoir of neutral Hydrogen after the EoR. DLAs have been studied both observationally (e.g. Péroux et al., 2003; Prochaska et al., 2005; O’Meara et al., 2007; Prochaska & Wolfe, 2009; Crighton et al., 2015) and theoretically/numerically (e.g. Nagamine et al.,

2004; Pontzen et al., 2008; Barnes & Haehnelt, 2009; Tescari et al., 2009; Bird et al., 2014; Rahmati et al., 2015; Maio & Tescari, 2015) in order to understand their physical properties and statistical distribution. However, many open questions still remain on the nature of these systems and their connection with the chemical enrichment of the Universe.

The aim of this chapter is to i) compare high- $z$  observations of metal and HI absorbers with synthetic spectra and ii) use simulations to gain insight into the underlying physical explanation for the observations.

### 3.3 Cosmological simulations

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

Each simulation generates a cosmological box (including periodic boundary conditions) with the same initial number of gas and dark matter (DM) particles for a total of  $2 \times 512^3$ . Initial masses of the gas and DM particles are given in Table 3.1. Whenever the gas density is above a threshold  $\rho_{\text{th}}$ , there is a probability that a gas particle will turn into a star<sup>2</sup>. Stochastically, a new star-type particle is introduced in the simulation. Each star particle represents a simple stellar population with mass  $0.1M_{\odot} \leq m \leq 100 M_{\odot}$ . The stars with mass  $m \leq 40 M_{\odot}$  explode as SNe before turning into a black hole, while stars that have masses above this threshold collapse into a black hole without passing through

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

<sup>2</sup>This probability is calibrated to reproduce the Kennicutt-Schmidt law (Kennicutt, 1998).

the SN stage.

A flat  $\Lambda$ CDM model with cosmological parameters from the latest release of the Planck Collaboration et al. (2015) is assumed:  $\Omega_{0m} = 0.307$ ,  $\Omega_{0b} = 0.049$ ,  $\Omega_{\Lambda} = 0.693$ ,  $n_s = 0.967$ ,  $H_0 = 67.74 \text{ km s}^{-1}\text{Mpc}^{-1}$  (or  $h = 0.6774$ ) and  $\sigma_8 = 0.816$ . The simulations were calibrated according to the parameters used in Tesconi et al. (2014) and Katsianis et al. (2015), and are compatible with observations of the cosmic star formation rate (SFR) density history and the galaxy stellar mass function at  $z = 6$  to 8.

The subgrid scheme takes into account the lifetimes of stars of different mass and follows the evolution of Hydrogen, Helium, molecules and metals (C, Ca, O, N, Ne, Mg, S, Si and Fe) released from SNIa, SNII and low and intermediate mass stars. It is possible to vary the initial mass function (IMF), the lifetime function and stellar yields. The algorithm is ideally suited to modeling IGM enrichment produced by galactic winds blown by “starburst” galaxies at high redshift. Radiative cooling and heating processes are included according to Wiersma et al. (2009) and Maio et al. (2007). The IMF considered is the Chabrier multi-sloped (Chabrier, 2003), as described in equation (3.1), that produces a large number of intermediate- and high-mass stars, expected to play an important role during Reionization.

$$\zeta(m) = \begin{cases} 0.497 \times m^{-0.2} & 0.1M_{\odot} \leq m < 0.3M_{\odot}, \\ 0.241 \times m^{-0.8} & 0.3M_{\odot} \leq m < 1M_{\odot}, \\ 0.241 \times m^{-1.3} & m \geq 1M_{\odot}. \end{cases} \quad (3.1)$$

We adopted the following stellar yields:

- SNIa: Thielemann et al. (2003). The mass range for the SNIa originating from binary systems is  $0.8 M_{\odot} < m \leq 8 M_{\odot}$ , with a binary fraction of 7%.
- SNII (massive stars): Woosley & Weaver (1995). The mass range for SNII is  $m > 8 M_{\odot}$ .
- Asymptotic giant branch (low and intermediate mass) stars: van den Hoek & Groenewegen (1997).

### 3.3.1 Feedback mechanisms

In order to regulate the star formation and chemically enrich the IGM, the simulations have been set up to account for kinetic supernova-driven winds. According to the nature

Table 3.1 Summary of the simulations used in this work. Column 1: run name. Column 2: box size. Column 3: Plummer-equivalent comoving gravitational softening length. Columns 4 and 5: mass of gas and dark matter particles. All the simulations have the same initial number of gas and DM particles ( $2 \times 512^3$ ). Column 6: feedback model. Column 7: inclusion of low-temperature metal and molecular cooling (Maio et al., 2007; Maio & Tescari, 2015). The first run, Ch 18 512 MDW, is the fiducial model. The second one in the list, Ch 18 512 MDW mol, has exactly the same configuration as the reference run, but includes low-T metal and molecular cooling.

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

of the feedback and its effectiveness, galactic winds are classified as momentum- or energy-driven winds.

### Momentum-driven galactic winds (MDW)

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

$$\dot{M}_w = \eta \dot{M}_\star. \quad (3.2)$$

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

$$v_w = 2 \sqrt{\frac{GM_h}{R_{200}}} = 2 \times v_{\text{circ}}, \quad (3.3)$$

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

$$R_{200} = \sqrt[3]{\frac{3}{4\pi} \frac{M_h}{200\rho_c\Omega_{0m}}} (1+z)^{-1}, \quad (3.4)$$

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

$$\eta = 2 \times \frac{600 \text{ km s}^{-1}}{v_w}. \quad (3.5)$$

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

### Energy-driven galactic winds (EDW)

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

$$\eta = 2 \times \left( \frac{600 \text{ km s}^{-1}}{v_w} \right)^2, \quad (3.6)$$

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

## 3.4 Methodology

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

Assuming this UVB, we compute the ionization states of each element with CLOUDY

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

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

photoionization code v8.1 (Ferland et al., 2013) for optically thin gas in ionization equilibrium, focusing in particular on the following ions: HI, CII, CIV, SiII, SiIV and OI. We consider only the transition with the highest oscillator strength and rest-frame wavelength  $\lambda_{\text{rest}} \geq 1216 \text{ \AA}$ , which can be observed at high redshift. The assumed metallicity in the CLOUDY tables is solar.

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

$$\frac{\Gamma_{\text{phot}}}{\Gamma_{\text{UVB}}} = (1 - f) \left[ 1 + \left( \frac{n_{\text{H}}}{n_0} \right)^\beta \right]^{\alpha_1} + f \left[ 1 + \frac{n_{\text{H}}}{n_0} \right]^{\alpha_2}, \quad (3.7)$$

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

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

$$\rho_X(j) = a^3 \sum_i X(i) W_{ij}, \quad (3.8)$$

$$(\rho T)_X(j) = a^3 \sum_i X(i) T(i) W_{ij}, \quad (3.9)$$

$$(\rho v)_X(j) = a^3 \sum_i X(i) (a \dot{x}(i) + \dot{a} [x(i) - x(j)]) W_{ij}, \quad (3.10)$$

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

mass and:

$$q_{ij} = \frac{a|x(i) - x(j)|}{h_{ij}}, \quad (3.11)$$

$$h_{ij} = \frac{1}{2} [h(i) + h(j)], \quad (3.12)$$

with  $h$  the physical softening scale.

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

$$\tau(u) = \frac{\sigma_{0,I}c}{H(z)} \int_{-\infty}^{\infty} n_I(x) V [u - x - v_{\text{pec,IGM}}^{\text{IGM}}(x), b(x)] dx, \quad (3.13)$$

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

In numerical works, it is a common practice to normalize HI fluxes averaged over a large number of random lines of sight to the observed mean normalized flux of the Ly $\alpha$  forest at a given redshift,  $\langle F(z) \rangle = \exp(-\tau_{\text{eff}})$ , through a constant rescaling factor  $A_{\text{HI}}$ . For consistency, also the fluxes of metal ions are rescaled by the same constant factor. However, at  $z > 5$  reliable measurements of  $\tau_{\text{eff}}$ , and therefore  $A_{\text{HI}}$ , are not possible, due to the thickening of the Ly $\alpha$  forest. To overcome this problem, we adopted a reversed approach. First, we calculated the CIV column density distribution function at  $z = 4.8$  and  $5.6$  (see section 3.6.1) using three different simulations (Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW) and a range of scaling factors for the CIV optical depths,  $\tau_{\text{CIV}}$ . Then, we selected the scaling factor,  $A_{\text{CIV}}$ , which provided the best chi-by-eye agreement between the simulated and observed distribution functions. Finally, we rescaled all the ionic fluxes (including HI) in all the simulations and at all the redshifts considered (i.e. also at  $z \leq 5$ , for consistency) by  $A_{\text{CIV}}$ . The best fit value is  $A_{\text{CIV}} = 0.85$ , very close to unity.

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

VPFIT v.10.2 (Carswell & Webb, 2014).

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

As a final remark, we note that radiative transfer effects are not included in our simulations. The models use the evolving HM12 uniform UVB to quantify the ionization state of the CGM/IGM at a given redshift. The aim of this work is not to follow the progression of Reionization, nor the evolution of the HII bubbles or their topology. The implicit assumption is that our boxes (that are small compared to the size of the HII bubbles at the redshifts of interest) represent a region of the Universe already reionized at a level given by the HM12 UVB. At  $6 < z < 8$ , chemical enrichment occurs mostly inside and in close proximity of galaxies (interstellar medium, CGM and high density IGM) where, assuming an inside-out progression of Reionization, the gas in which metals lie should be ionized. Although proper RT calculations would be more accurate, they are extremely expensive from the computational point of view. Using a uniform (evolving) UVB is a common approach in the literature (see e.g. Oppenheimer et al., 2009; Keating et al., 2016). Moreover, Finlator et al. (2015) studied the evolution of Carbon absorption from  $z = 10$  to  $z = 5$  with cosmological hydrodynamic simulations that include a self-consistent multifrequency, inhomogeneous UVB. They found that the difference between their more realistic UVB and the uniform HM12 is within  $\sim 2 - 4$  times, which according to them is fair agreement given the uncertainties.

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

### 3.5 Evolution of the IGM during Reionization

Figure 3.1 shows the distribution of the gas particles in the fiducial simulation Ch 18 512 MDW in diagrams of overdensity–temperature at redshifts  $z = 8, 6$  and  $4$ , comparing total metallicity, number of particles and HI fraction (left, middle and right panels, respectively). These redshifts are chosen to study the IGM at i)  $z = 8$  when Reionization is progressing;

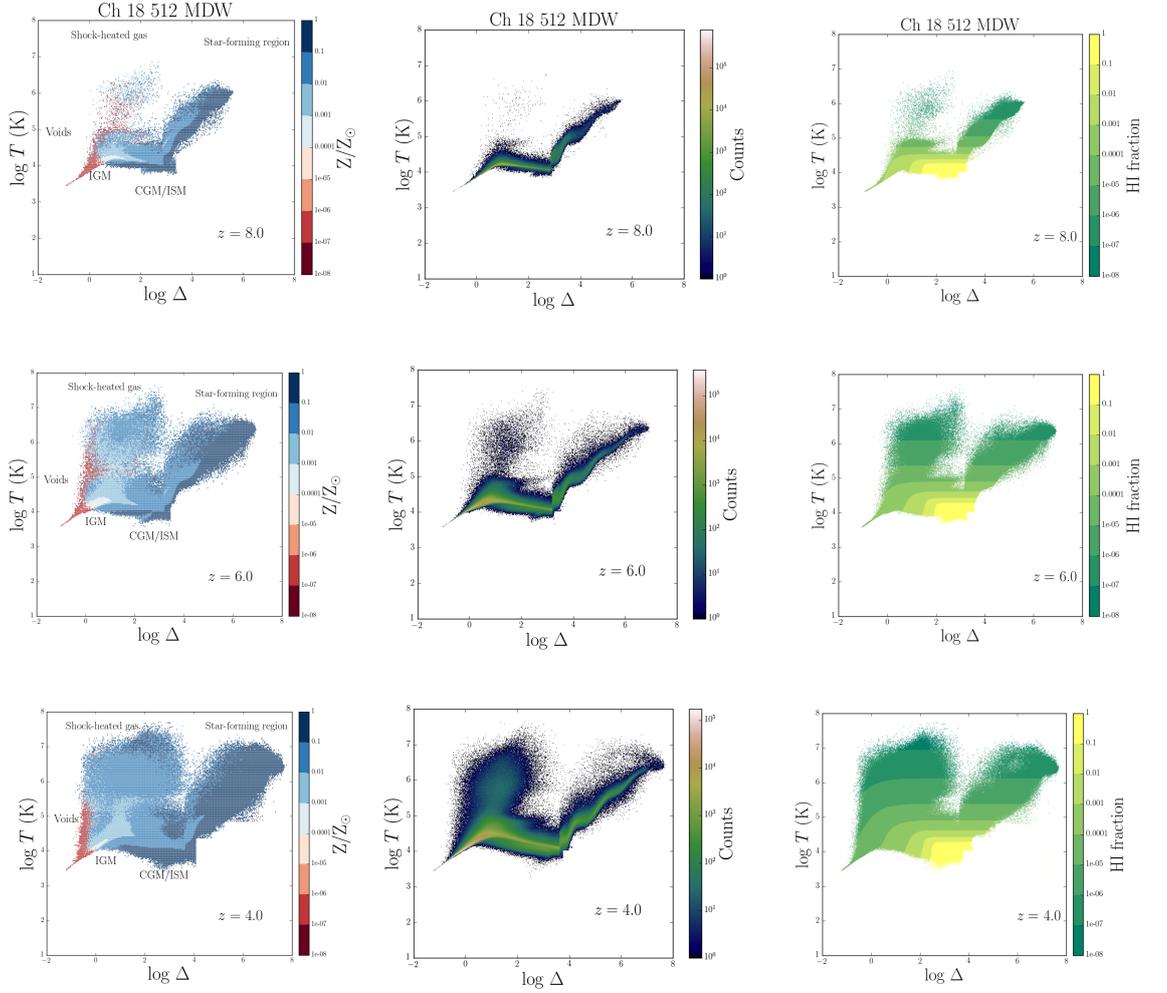


Figure 3.1 Density–temperature diagram for gas particles at redshifts  $z = 8, 6$  and  $4$  for the fiducial model Ch 18 512 MDW. The left panels corresponds to the distribution of the total metallicity (in solar units). The dark red region represents zero metallicity. As a reference, the different regions are tagged (voids, IGM, CGM/ISM, shock–heated gas and star–forming region). In the middle panels, the colour maps represent the number of particles with overdensity and temperature in a given  $(\log \Delta_i, \log T_i)$  bin. On the right hand side, we show the distribution of HI fraction. The top right panel was included for illustrative purposes and does not represent a realistic description of the neutral Hydrogen content of the Universe at  $z = 8$  (see the discussion at the end of section 3.4).

ii)  $z = 6$ , when the overlap of the HII bubbles is nearly complete at the end of the EoR; and  
 iii)  $z = 4$  when the Reionization of Hydrogen has fully concluded and the only remaining reservoirs of neutral Hydrogen are self-shielded regions of HI and DLA systems.

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

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

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

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

The right panels show the distribution of Hydrogen in the reference simulation. Yellow indicates Hydrogen completely neutral ( $X_{\text{HI}} \geq 10^{-1}$ ), dark green particles containing HII

( $X_{\text{HI}} < 10^{-7}$ ) and the intermediate colors are used to illustrate the transition between the two regimes. It is clear that most of the reservoirs of HI are located at low temperature and intermediate- to high-density regions, that are presumably where DLAs or isolated self-shielded regions are. On the other hand, the shock-heated gas with temperatures  $T > 10^5$  K contains a large amount of ionized gas. The fraction of this gas that has been chemically enriched is expected to be found in high ionization states, such as CIV or SiIV, consistent with Cen & Chisari (2011).

Please note that, for the reasons discussed at the end of section 3.4, the top right panel should not be regarded as an accurate representation of the neutral Hydrogen content in the Universe at  $z = 8$ , since a uniform UVB does not properly describe the progression of Reionization. We included this panel just for illustrative purposes.

## 3.6 Evolution of Carbon

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

### 3.6.1 CIV column density distribution function

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

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

$$f(N, X) = \frac{n_{\text{sys}}(N, N + \Delta N)}{n_{\text{lov}} \Delta X}, \quad (3.14)$$

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

$$\Delta X = \frac{H_0}{H(z)} (1 + z)^2 \Delta z. \quad (3.15)$$

The equivalent redshift path of our cosmological boxes is defined as:

$$\Delta z = (1 + z) \frac{\Delta v}{c}, \quad (3.16)$$

where  $\Delta v$  is the box size in  $\text{km s}^{-1}$  at a given redshift  $z$ .

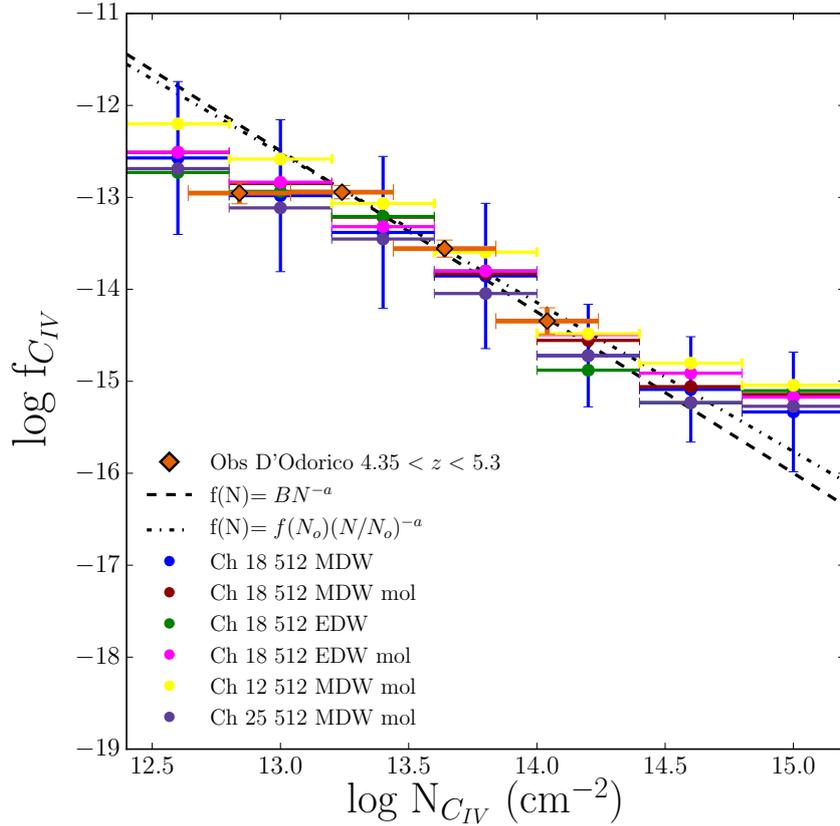


Figure 3.2 CIV column density distribution function at  $z = 4.8$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 10.29 \pm 1.72$  and  $\alpha = 1.75 \pm 0.13$  and the dotted–dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 13.56$  and  $\alpha = 1.62 \pm 0.2$ , from the same observational work. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. Hereafter, these colors are used to represent the simulations.

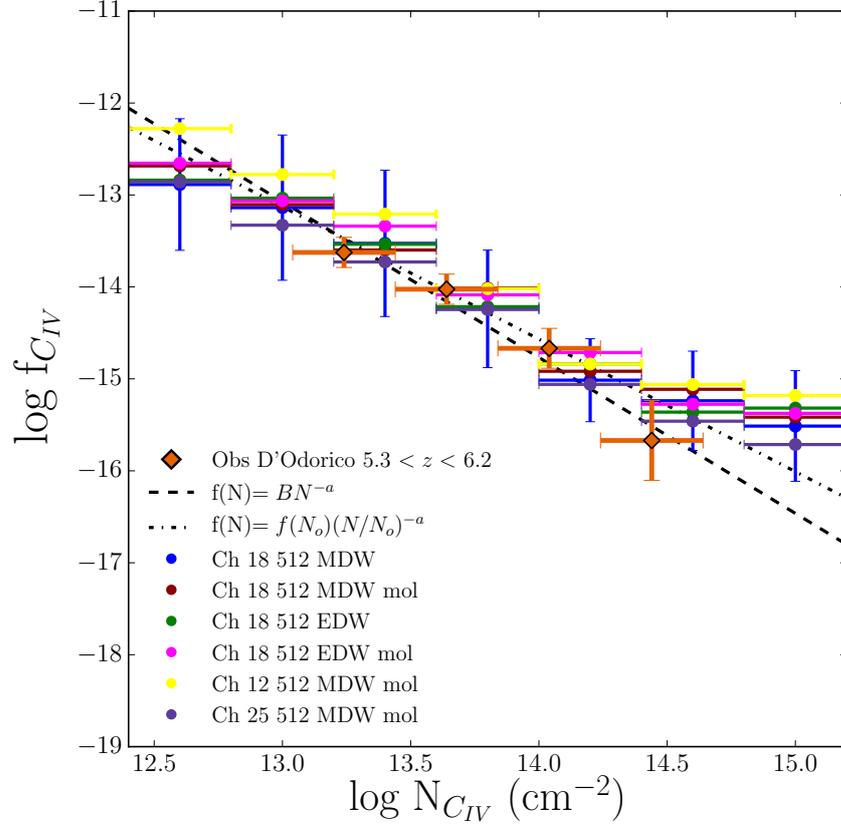


Figure 3.3 CIV column density distribution function at  $z = 5.6$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 8.96 \pm 3.31$  and  $\alpha = 1.69 \pm 0.24$  and the dotted–dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 14.02$  and  $\alpha = 1.44 \pm 0.3$ , from the same observational work. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. At large column densities, simulations predict values of the CIV–CDDF slightly higher than observations. However, the error bars on all of the theoretical models overlap with the observational data.

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

Figure 3.2 shows the predicted column density distribution function of CIV at  $z = 4.8$  and compares with observations by D’Odorico et al. (2013) and two fitting functions proposed by the authors:  $f(N) = BN^{-\alpha}$  with  $B = 10.29 \pm 1.72$  and  $\alpha = 1.75 \pm 0.13$  and

$f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 13.56$  and  $\alpha = 1.62 \pm 0.2$ . Although we used this statistics as a rough guide to calibrate the CIV optical depths in three of our simulations (see the final part of section 3.4), there is good agreement among *all* the simulations and the observational data at this redshift, and small deviations from the observations are within the error bars. The error bars in CIV–CDDF shown correspond to the reference run and are representative for all the simulations.

The same observable is computed at  $z = 5.6$  in Figure 3.3. The values from the numerical simulations are shown in points and they are compared to systems detected by D’Odorico et al. (2013). The observational fitting functions parameters for  $f(N) = BN^{-\alpha}$  are  $B = 8.96 \pm 3.31$ ,  $\alpha = 1.69 \pm 0.24$  and for  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  are  $f(N_0) = 14.02$  and  $\alpha = 1.44 \pm 0.3$ . There is good agreement between the observational and theoretical points except at large column densities, where simulations deviate to values higher than observations. We point out that the function fits at  $z = 4.8$  and  $5.6$  extend to high column densities, even though observations are only available up to  $N_{\text{CIV}} = 10^{14.4} \text{ cm}^{-2}$ . Therefore, the outcome of the simulations gives additional information in a range of  $N_{\text{CIV}}$  where the CIV–CDDF has not been measured.

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

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

In similar works by Finlator et al. (2015, 2016) and Keating et al. (2016), strong absorbers are highly disfavored for different simulated UV backgrounds and feedback prescriptions. In Finlator et al. (2015, 2016), the feedback mechanisms are quite efficient at enriching the IGM with metals. However, their simulated boxes are relatively small. Therefore, galaxies with large halo masses are suppressed. If these structures host the high column density systems, such simulations will also lack strong absorptions. On the other hand, Keating et al. (2016) explore various feedback prescriptions. Their Sherwood run has a similar hydrodynamics and configuration as our models, whereas HVEL implements a more aggressive version of the energy–driven winds used in this work (including also a minimum wind speed of  $600 \text{ km s}^{-1}$ ). Nevertheless, all their models struggle to reproduce the high CIV column density absorbers. This issue is not alleviated by varying the UVB.

The presence of rare strong CIV absorbers in our theoretical models results from an appropriate level of resolution on the scale of the absorbers and a post-processing pipeline that closely mimics the method used by observers and accounts for individual features to calculate the column densities from Voigt profile fits to the absorption lines. The latter increases the accuracy in the estimation of the column densities.

In addition to the results published in García et al. (2017a), here we present the CIV column density distribution function at  $z \sim 6.4$  to compare with the latest data from Bosman et al. (2017) in the redshift range  $6.2 < z < 7.0$  (see Figure 3.4). In Bosman et al. (2017), there is just one detection of CIV at low column density, therefore, the authors provide upper limits for the observed CDDF. The simulations are consistent with these upper limits, in particular, the run Ch 12 512 MDW mol, for the same reasons explained above.

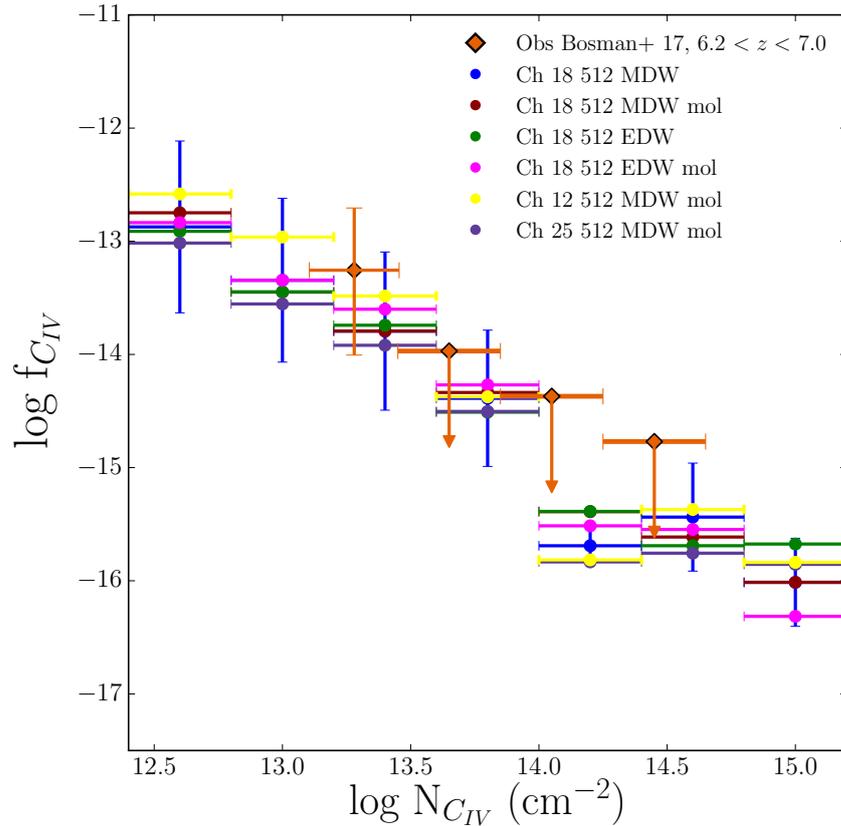


Figure 3.4 CIV column density distribution function at  $z = 6.4$  and comparison with observational data by Bosman et al. (2017) in orange diamonds. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models.

### 3.6.2 CIV cosmological mass density

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

$$\Omega_{\text{ion}}(z) = \frac{H_0 m_{\text{ion}}}{c \rho_{\text{crit}}} \frac{\sum N(\text{ion}, z)}{n_{\text{lov}} \Delta X}, \quad (3.17)$$

where  $m_{\text{ion}}$  is the mass of the ionic species,  $n_{\text{lov}}$  is the number of lines of view (lov),  $\rho_{\text{crit}}$  is the critical density today and  $\Delta X$  is defined in equation (3.15).

In the case of CIV, we restricted the column densities in the sum to the range  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , to compare directly with the results of D’Odorico et al. (2013). The results for all of the simulations are shown in Figure 3.5. The theoretical results are compared to observations by Pettini et al. (2003) and Ryan-Weber et al. (2009), Songaila (2001, 2005), Simcoe et al. (2011), D’Odorico et al. (2013), Boksenberg & Sargent (2015), Bosman et al. (2017) and Díaz et al. (in prep). It is worth mentioning that the observations by Pettini, Ryan-Weber and Díaz have been recalibrated to the cosmology adopted in this paper. In the other cases, missing details of the precise pathlength probed do not allow us to convert the observations to the Planck cosmology.

The rise of CIV over the redshift period from 8 to 4 reflects both the increase in the chemical enrichment of the intergalactic medium through galactic feedback mechanisms, and the evolution of the ionization state of the IGM. At  $z = 8$ , CIV is largely suppressed and most of the Carbon in the box is in its neutral state. As time passes, more SN events take place and pollute the IGM with metals (as we will show in section 3.6.4). Simultaneously, the specific intensity of the HM12 UVB around the wavelength where the CIV transition occurs ( $\lambda_{\text{CIV}} = 192 \text{ \AA}$ ) increases by  $\sim 4$  orders of magnitude between  $z = 8$  and 4. At all redshifts, each simulation is broadly consistent with current observational data.

Using results from a hybrid wind model, Finlator et al. (2015, 2016) integrated a power law fit to the column density distribution function to calculate the evolution of the comoving mass density of CIV. Instead, we calculate  $\Omega_{\text{CIV}}$  by summing the column density of the absorbers using equation (3.17). This method is more accurate when there are uncertainties in the slope of the CIV–CDDF, especially at high column densities.

### 3.6.3 CII cosmological mass density

We extend the analysis to CII and compare with the observational lower limits of Becker et al. (2006) in Figure 3.6, taking into account systems whose column densities are in the

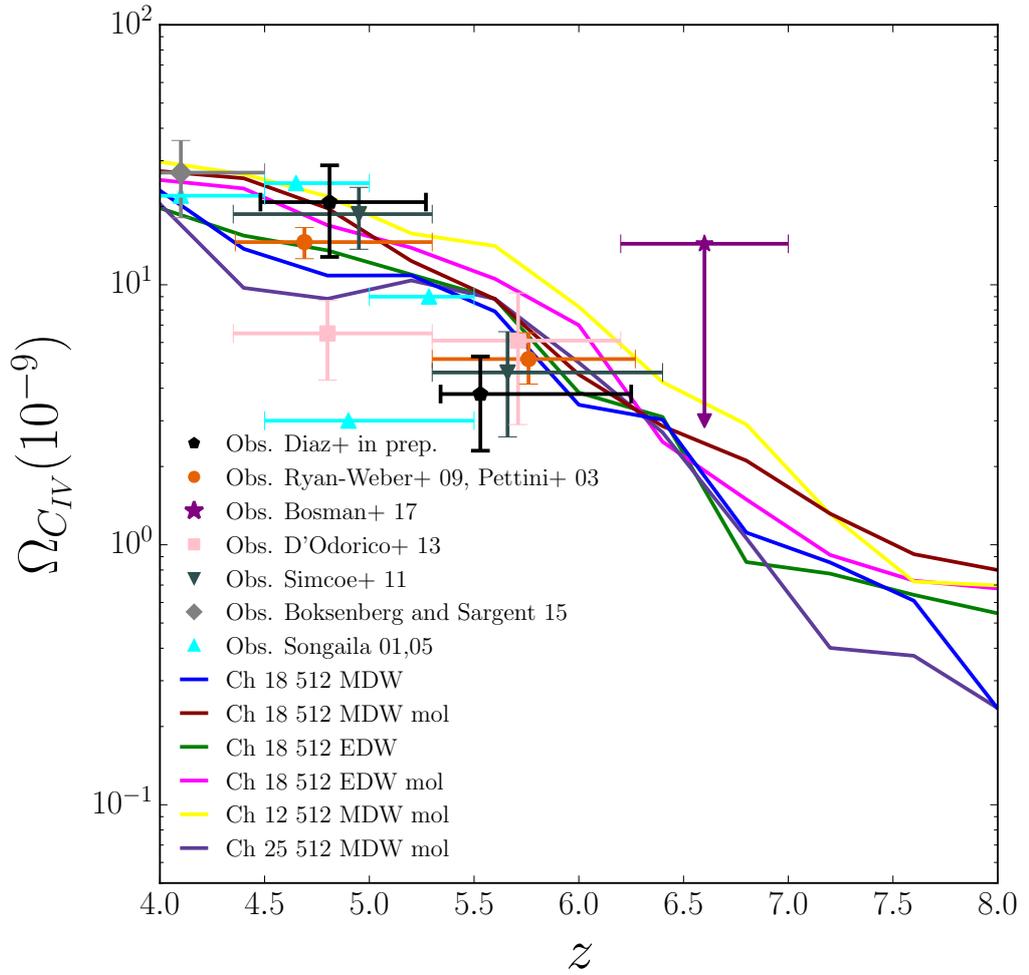


Figure 3.5 CIV cosmological mass density at  $4 < z < 8$ . Comparison between the simulated data and observations by Pettini et al. (2003) and Ryan-Weber et al. (2009) in orange circles, Songaila (2001, 2005) in cyan triangles, Simcoe et al. (2011) in dark green inverted triangles, D’Odorico et al. (2013) in pink squares, Boksenberg & Sargent (2015) in grey diamond, upper limits from Bosman et al. (2017) in purple star and Díaz et al. (in prep) in black pentagons. Pettini, Ryan-Weber and Díaz measurements are converted to the Planck cosmology, while for the others this recalibration was not possible due to missing details of the precise pathlength probed.

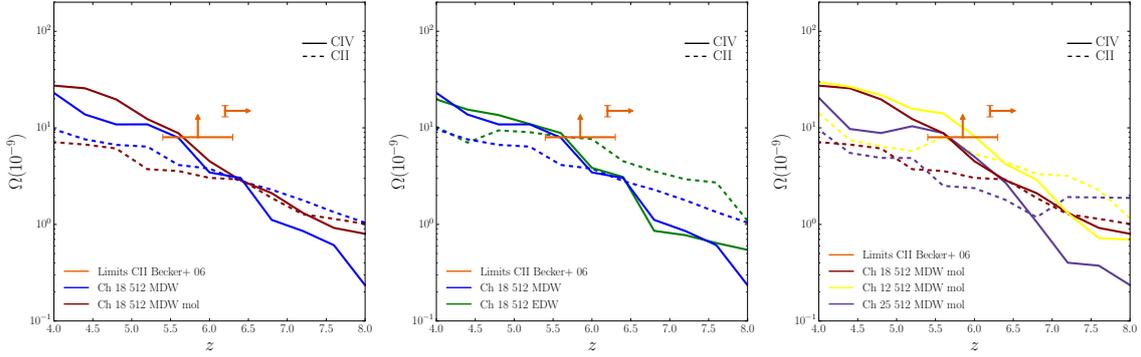


Figure 3.6 Evolution of the CII and CIV cosmological mass density. The left and central panel compare simulations with 18 cMpc/ $h$  box size. On the left is displayed the case of simulations with MDW feedback with and without low temperature metal and molecular cooling included. In the middle, we show the configuration without low-T metal and molecular cooling for the runs with MDW and EDW feedback. On the right hand side, a resolution and box size test with the following models is shown: Ch 12 512 MDW mol, Ch 18 512 MDW mol and Ch 25 512 MDW mol. In all the panels, the solid lines show the evolution of  $\Omega_{\text{CIV}}$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , and the dashed lines  $\Omega_{\text{CII}}$  in the range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$ . The orange points with errors represent the observational lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006).

range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$ . In the context of Reionization, it is interesting to consider the evolution of CII (a low ionization state of Carbon with ionization energy of 11.26 eV) with respect to CIV, whose ionization energy is more than 4 times larger (47.89 eV). Since the UVB is softer at earlier times, CII is expected to be more dominant than higher ionization states of Carbon at high redshifts.

On the left and central panels of Figure 3.6 we draw a comparison between different physical scenarios among the simulations with boxsize 18 cMpc/ $h$ : the further left relates simulations with the same feedback prescription (MDW) with and without low temperature metal and molecular cooling. The run with low-T cooling included produces slightly more CIV (and less CII) than the fiducial run, but the difference is only significant at  $z > 6.5$ . In the middle panel, we contrast simulations without low temperature metal and molecular cooling with different feedback prescriptions: MDW and EDW. The different physical scenarios do not have a strong impact on the evolution of CIV, as seen also in Figure 3.5. The trends for CII are more dependent on the model, because this low-ionization state is very sensitive to the densities of the absorbers (which are correlated to the strength of the feedback prescription), and the Ch 18 512 EDW run is in good agreement with the observations of Becker et al. (2006).

Although the run with lowest resolution, Ch 25 512 MDW mol, produces slightly degraded results, the evolution of  $\Omega_{\text{CIV}}$  does not change significantly with the spatial resolution or size of the simulations (see the rightmost panel of Figure 3.6).

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

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

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

### 3.6.4 Total Carbon cosmological mass density

To conclude section 3.6, in Figure 3.7 we analyse the evolution of the cosmological mass density of (total) Carbon in the simulated boxes,  $\Omega_{\text{C}}$ . This is calculated differently than  $\Omega_{\text{CIV}}$  and  $\Omega_{\text{CII}}$ : here we sum the amount of Carbon in each gas particle in the simulation and divide by the comoving volume. Thus,  $\Omega_{\text{C}}$  gives an estimation of the total amount of Carbon at high redshift. The trend is consistent with the enrichment history of the Universe, increasing by 1.5 orders of magnitude from  $z = 8$  to 4. At  $z = 4$ , our results are compatible with those of Tescari et al. (2011). Given the different methods used to calculate  $\Omega_{\text{C}}$  and  $\Omega_{\text{CIV}}$  &  $\Omega_{\text{CII}}$ , the first quantity should not be directly compared with the other two. Moreover, the evolution of the total Carbon cosmological mass density is a purely theoretical prediction of our simulations, since  $\Omega_{\text{C}}$  is not (directly or indirectly) observable.

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

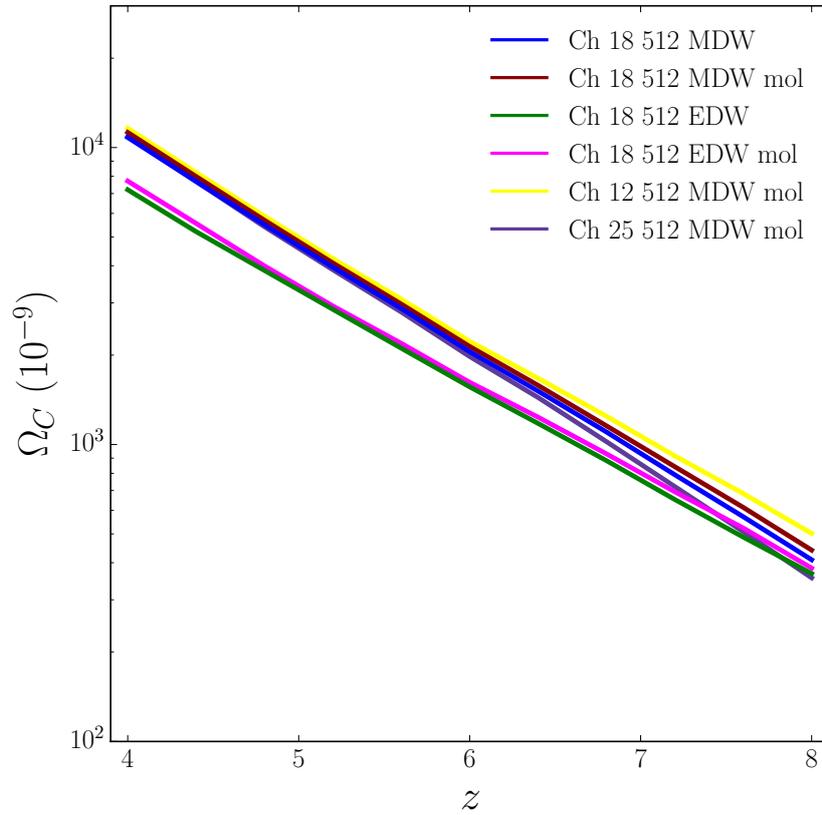


Figure 3.7 Evolution of the total Carbon cosmological mass density for our simulations. This is a purely theoretical prediction, since  $\Omega_C$  is not (directly or indirectly) observable.

The EDW models quench the formation of stars in galaxies more effectively than MDW, therefore less Carbon is produced and  $\Omega_C$  is lower.

Figure 3.7 also compares simulations with different spatial resolutions (the comoving softening of the different configurations scales with the size of the box). The Ch 12 512 MDW mol, Ch 18 512 MDW mol and Ch 25 512 MDW mol runs have comoving softening of 1.0, 1.5 and 2.0 cMpc/h (yellow, red and purple lines), respectively. These simulations predict a similar trend for the mass density of Carbon and at low redshift converge to the same  $\Omega_C$ . However, at high redshift the resolution affects the amount of metals. The run with the highest resolution (Ch 12 512 MDW mol, yellow line) can resolve higher densities at earlier times and therefore better describes star formation in galaxies and produces more Carbon than Ch 18 512 MDW mol (red) and, particularly, Ch 25 512 MDW mol (purple).

### 3.7 Column density relationships of metal absorption lines

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

As mentioned in section 3.6.1, this procedure produces a very large number of systems with low column densities, that display non physical linear correlations among the metal absorption features. To avoid this issue, a cut-off in column density,  $\log N_{\text{th}} (\text{cm}^{-2}) \geq 12.5$ , has been imposed to the synthetic data, and systems below this noise threshold have not been taken into account in the statistics.

We explore relationships in column densities among different ionic species in a uniform UV background HM12 at  $z = 6$  using random lines of sight to emulate the observational method and compare with the observations available to date in Figure 3.8. In panel A,  $N_{\text{CII}}$  vs.  $N_{\text{CIV}}$  are compared, panel B displays  $N_{\text{SiIV}}$  vs.  $N_{\text{CIV}}$  and panel C the ratio of  $N_{\text{CII}}/N_{\text{CIV}}$  vs.  $N_{\text{SiIV}}/N_{\text{CIV}}$ . All these quantities are compared to systems detected by D’Odorico et al. (2013), while panels D, E and F draw a comparison with low ionization states observed by Becker et al. (2006). The comparisons  $N_{\text{CII}}$  vs.  $N_{\text{SiII}}$ ,  $N_{\text{CII}}$  vs.  $N_{\text{OI}}$  and  $N_{\text{SiII}}$  vs.  $N_{\text{OI}}$  are shown in panels D, E and F, respectively. In all cases, we display the simulations with 18 cMpc/h box size, to analyze possible variations among the predicted values with different prescriptions of feedback and cooling. However, the dispersion of the data points do not allow clear distinctions between the models.

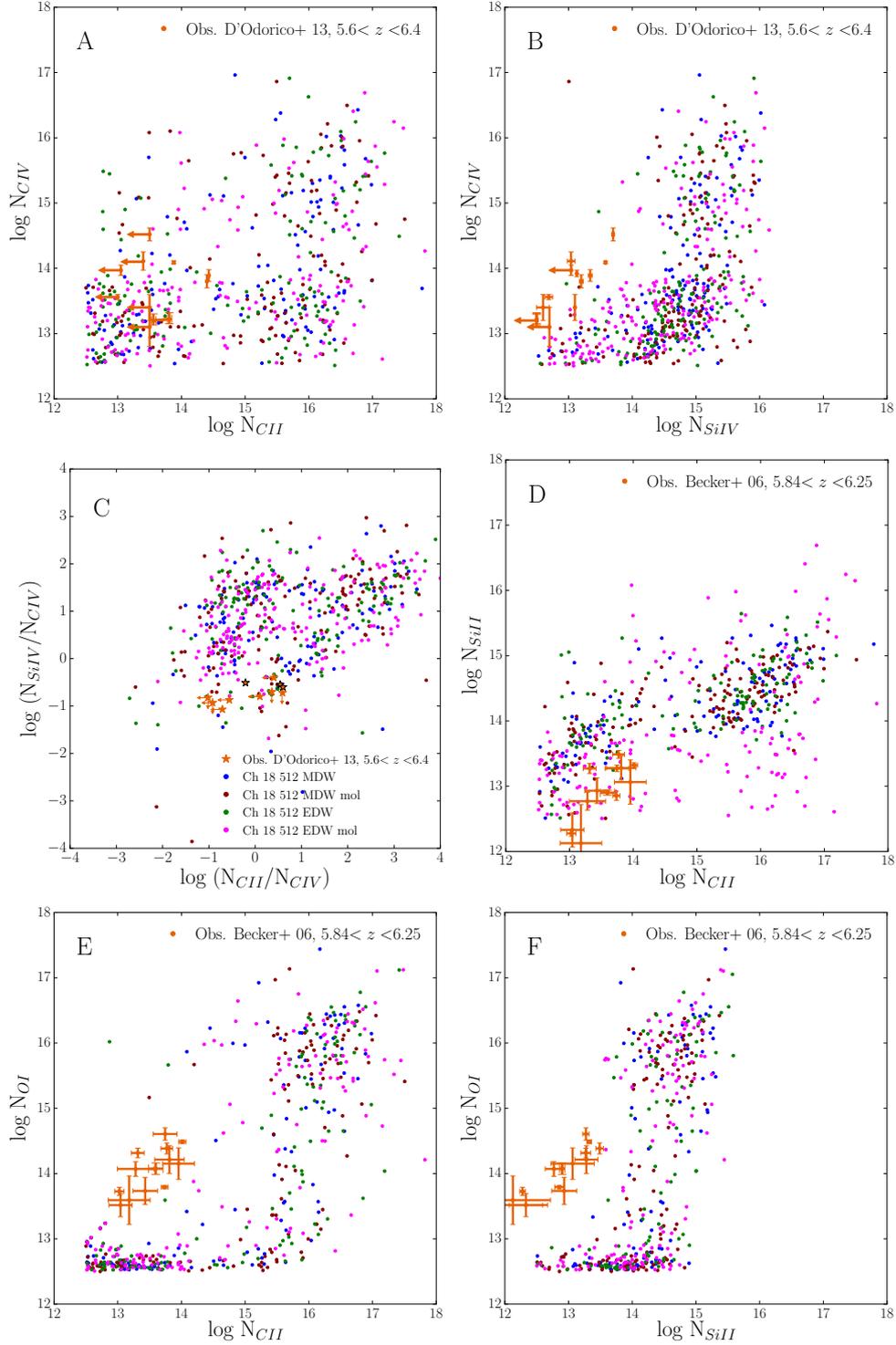


Figure 3.8 Column density relationships among metal absorption lines at  $z = 6$ . Panels A, B and C compare with systems observed by D’Odorico et al. (2013) in the redshift range of (5.6–6.4) and D, E and F with absorbers detected by Becker et al. (2006) in the redshift range of (5.84–6.25), both in orange. From the synthetic spectra of each ion, the column densities are calculated and displayed for the simulations with box size of 18 Mpc/ $h$  (the label keys are shown in the panel C), but there is not a remarkable difference between the different prescriptions (feedback and cooling) imposed in each case. The simulated data displayed are above the threshold noise at  $\log N_{\text{th}} = 12.5$  ( $\text{cm}^{-2}$ ). This selection criterion results in an effective number of synthetic systems of  $\sim 18\%$  from the initial 1000 random lines of sight traced. In panel C, the observations by D’Odorico et al. (2013) produced 3 systems with column densities reported (represented by the black stars) and 8 with upper limits in  $N_{\text{CII}}$  and/or  $N_{\text{SiIV}}$ , represented by arrows.

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

CII is a very interesting case in this study. The numerical estimation of CII column densities do not seem to be well reproduced by the simulations, although in most of the cases the observational data are just a lower limit, and a bimodality in the distribution of column densities appears, which will be analyzed in section 3.7.1.

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

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

### 3.7.1 Bimodality in CII column densities

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

Observations at  $z \sim 6$  exhibit  $N_{\text{CII}} \leq 10^{14} \text{ cm}^{-2}$ , while at higher redshift only lower limits in CII are available. In all the simulations, we found a remarkable bimodal distribution

---

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

in the  $N_{\text{CII}}$  calculated using random sightlines through each box (see e.g. panel A in Figure 3.8). In order to test the nature of this bimodality, in Figure 3.9 we plot the column densities of CII and CIV, using lines of sight at different impact parameters from the centre of the galaxies in the fiducial run Ch 18 512 MDW ( $d = 20, 100$  and  $500$  ckpc/ $h$  in yellow, red and green points, respectively). At each  $d$ , we extracted a thousand low around halos with masses in the range  $10^9\text{--}10 M_{\odot}/h$ . The blue points represent the column densities along 1000 random lines of sight through the box shown in panel A of Figure 3.8. The black stars are observational data from D’Odorico et al. (2013). The top panel in Figure 3.9 displays the distribution of the CII absorption features and the right-most window the corresponding distribution of CIV column densities above the noise threshold.

When the impact parameter  $d$  approaches the galaxies, the peak at high column densities becomes sharper, indicating that there are more CII absorbers in the CGM where the temperature and density are larger than in the IGM. For  $d = 20$  ckpc/ $h$ ,  $\sim 3$  pkpc at this redshift, the absorbers are located inside or in the outskirts of the galaxies (spectroscopic confirmation of galaxies at  $z = 6\text{--}8$  with the Hubble Space Telescope shows that the radii of these galaxies are in the range of  $0.6\text{--}1.1$  pkpc, Jiang et al. 2013). Absorbers are gravitationally bounded to the halos in the simulated sample, which have virial radii of  $3\text{--}30$  pkpc.

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

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

Finally, numerical results in the CGM (yellow and red cases in Figure 3.9) lead to a bimodality in CIV and a clear excess of high column density systems at low impact pa-

rameters. Oppenheimer et al. (2009) showed that the properties of the CIV absorbers depend mostly on the evolution of the environment of their hosting galaxy (distance, mass and metallicity). By tracing the surroundings of the galaxies in halos with masses  $10^9\text{--}10M_\odot/h$ , we confirm that there is a correlation between CIV column densities and the proximity of the galaxies that host the absorbers.

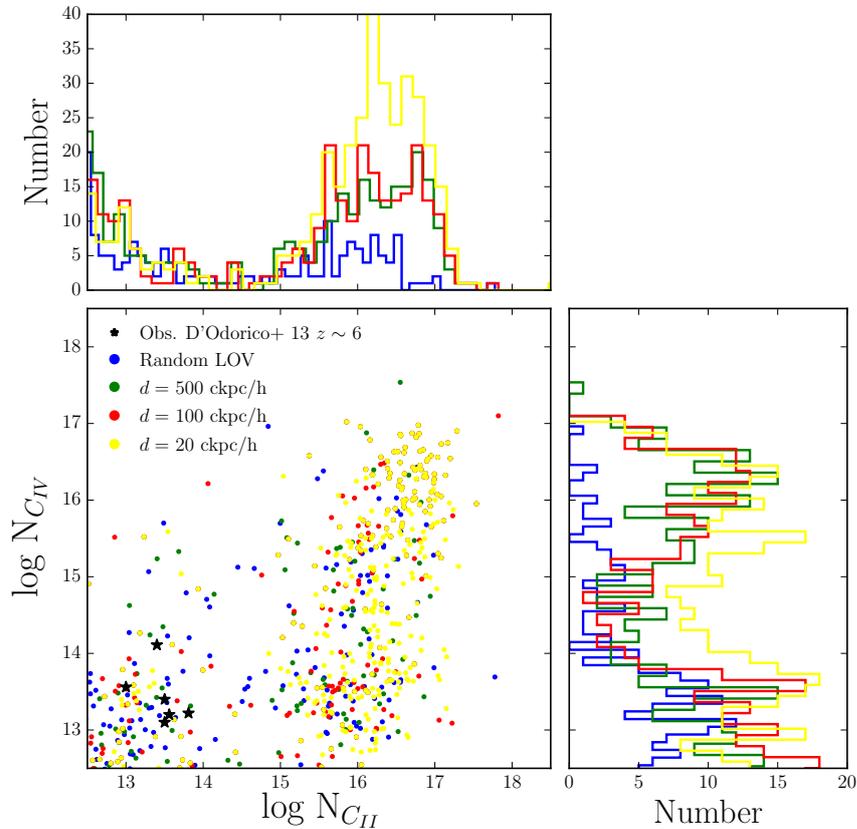


Figure 3.9 CII vs CIV column densities in different environments at  $z = 6$  for the reference simulation, Ch 18 512 MDW. Using different impact parameters from the centre of the galaxies it is possible to understand what drives a bimodality in  $N_{\text{CII}}$ . This low ionization state of C is mostly found in the CGM, but there are some traces of it in the IGM and, so far, it has been observed in the latter, with low column densities, as it is shown in the plot (the black stars are observational data at  $z = 6$  from D’Odorico et al. 2013). The blue points are theoretical predictions using random lines of sight (cf. panel A in Figure 3.8). When the sightlines are chosen to be close to the centre of the galaxies, the distribution in CII is shifted to larger column densities, indicating that high- $z$  galaxies are huge reservoirs of CII, not detected yet due to the low likelihood of reaching these overdense regions with QSOs lines of sight.

### 3.8 Cosmological mass density of HI at redshifts lower than 6

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

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

$$f_{\text{HI}}(N, X) = \frac{n_{\text{sys}}(N, N + \Delta N)}{\Delta X \cdot n_{\text{cube}} \cdot n_{\text{grid}}^2}, \quad (3.18)$$

where  $n_{\text{sys}}$  is the number of systems in the interval  $(N, N + \Delta N)$ ,  $\Delta X$  the absorption path defined in equation (3.15),  $n_{\text{cube}}$  the number of cubes generated and  $n_{\text{grid}}$  the number of bins in the grid. The parameters selected to achieve a good resolution of the systems that contain neutral Hydrogen are:  $n_{\text{cube}} = 1000$  and  $n_{\text{grid}} = 64$  (i.e. the grid size is 3.1 ckpc/h).

Figure 3.10 shows the HI-CDDF, with column densities in the range  $12 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$ , for all the simulations at  $z = 4$  and compares with the observations of Prochaska et al. (2005), Crighton et al. (2015), Bird et al. (2017) (all in the column densities corresponding to DLA systems at  $z \geq 4$ ) and O’Meara et al. (2007) in the super Lyman limit systems (LLS) range at  $z \sim 3.4^6$ . Note that there is a very large number of systems with low column densities in all the models and DLAs are extremely rare. The agreement with the observational data is good, indicating that the method implemented to find neutral systems works well at these redshifts.

Simulations with smaller comoving softening (or smaller box size), better resolve the Ly $\alpha$  forest. In Figure 3.10 the yellow line (simulation Ch 12 512 MDW mol) is above the red

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

line, corresponding to the run Ch 18 512 MDW mol, and the latter is above the purple line (Ch 25 512 MDW mol, the run with the lowest resolution in the suite of simulations) at  $\log N_{\text{HI}}(\text{cm}^{-2}) < 19$ .

At DLA column densities, the distributions follow a double–power law, as discussed by Zwaan & Prochaska (2006) and Prochaska & Wolfe (2009), with a knee at  $\log N_{\text{HI}}(\text{cm}^{-2}) \sim 21$  in all of the cases considered. Also, at  $N_{\text{HI}} > 10^{21} \text{ cm}^{-2}$ , where molecular cooling plays an important role in the star forming regions, the HI–CDDF in models with molecular cooling implemented drops off more rapidly than in the other models.

In Figure 3.11 a zoom of the DLA region is shown ( $\log N_{\text{HI}}(\text{cm}^{-2})$  in the range [20.3 , 22.0]). Overplotted are the most recent observational data of HI absorbers at high redshift by Prochaska et al. (2005), Crighton et al. (2015), Bird et al. (2017) and the fitting function proposed by Prochaska & Wolfe (2009) for DLA systems at redshift 4.0–5.5. The simulations are compatible with the observational data, especially runs with no molecular cooling.

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

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

$$\Omega_{\text{HI}}(z) = \frac{H_0 m_{\text{HI}}}{c \rho_{\text{crit}}} \int_{N_{\text{min}}}^{N_{\text{max}}} f_{\text{HI}}(N, z) N dN, \quad (3.19)$$

with  $m_{\text{HI}}$  the mass of the Hydrogen atom and  $f_{\text{HI}}(N, z)$  defined in equation (3.18). In theory, the integral should be computed with an upper limit  $N_{\text{max}} = \infty$ , but in reality this value is set by the maximum column density detected. Thus, the limits considered for  $\Omega_{\text{HI}}$  are  $12 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$ .

We compare the predictions from the theoretical models to observations at high redshift by Prochaska et al. (2005), Prochaska & Wolfe (2009), Zafar et al. (2013) and Crighton et al. (2015) in the left panel of Figure 3.12.

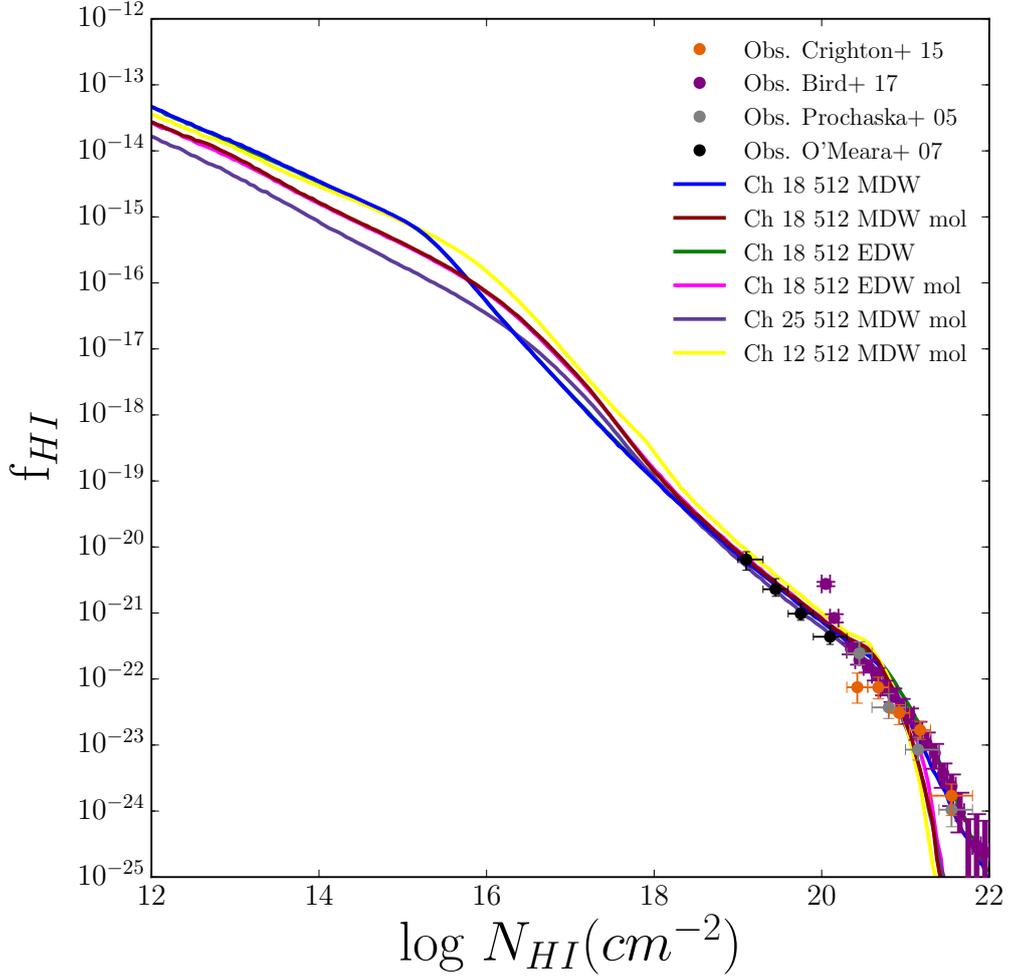


Figure 3.10 HI column density distribution function at  $z = 4$  in the range  $12 < \log N_{HI}(\text{cm}^{-2}) < 22$  for all the simulations described in Table 3.1 and comparison to observations by Prochaska et al. (2005) in grey, O’Meara et al. (2007) in black, Crighton et al. (2015) in orange and Bird et al. (2017) in purple.

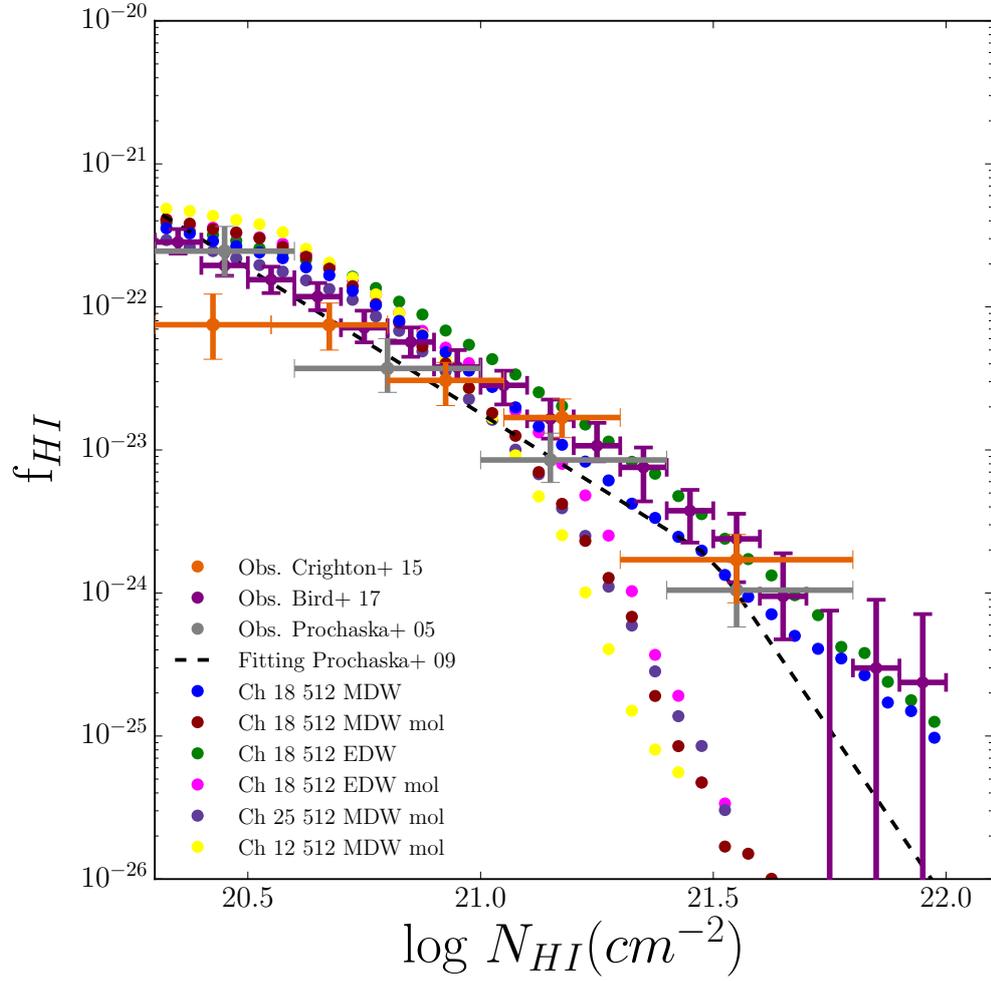


Figure 3.11 Same as Figure 3.10 but zoomed in the DLA range:  $20.3 < \log N_{HI}(\text{cm}^{-2}) < 22$ . The black dashed line represents the fitting function by Prochaska & Wolfe (2009) for DLA systems at redshift 4.0–5.5.

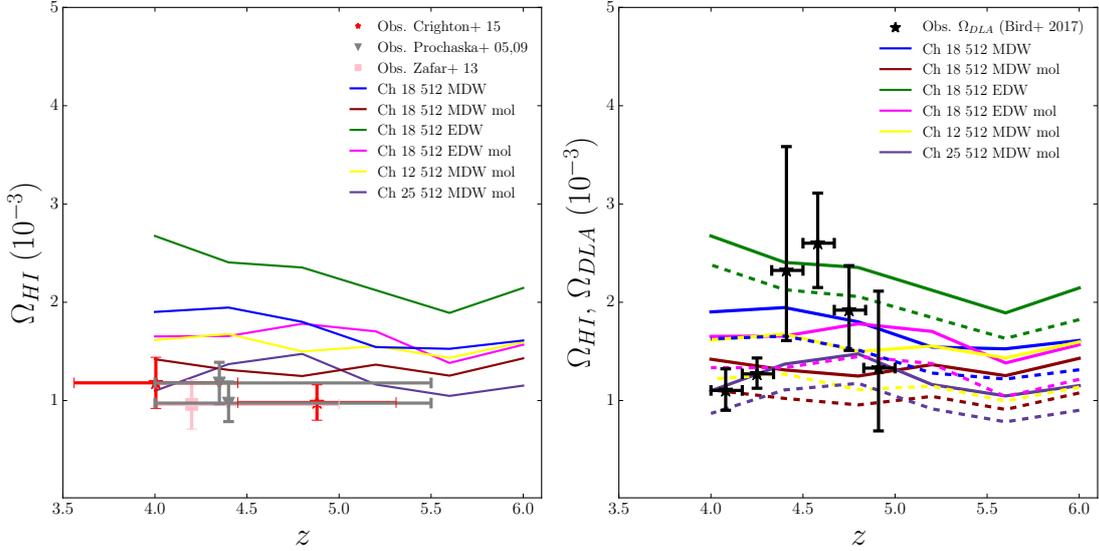


Figure 3.12 Cosmological mass density of HI. In the left panel, we display the theoretical prediction of  $\Omega_{\text{HI}}$  and compare with the data available at high redshift: Prochaska et al. (2005); Prochaska & Wolfe (2009) in grey inverted triangles and Zafar et al. (2013) in pink square and Crighton et al. (2015) in red stars. The right panel shows  $\Omega_{\text{HI}}$  and  $\Omega_{\text{DLA}}$  (solid and dashed lines, respectively). We compare the resulting  $\Omega_{\text{DLA}}$  with recent data from SDSS (Bird et al., 2017) in black stars. In the observational work of Crighton et al. (2015) it is proposed that more than 80% of the contribution to the HI cosmological mass density at  $z < 5$  comes from DLA systems. We confirm this assumption theoretically (Table 3.2).

The amount of HI at  $z = 4$  in the Ch 18 512 MDW mol (red line) and Ch 25 512 MDW mol (purple line) runs is in excellent agreement with the observational data, probably because the inclusion of molecular cooling prevents the production of too much neutral Hydrogen (by converting it into molecular Hydrogen). On the other hand, simulations without molecular cooling overproduce  $\Omega_{\text{HI}}$ , especially in the EDW case (green line). All the models predict an increasing amount of HI at  $z > 5.5$ . However, we stress again that our assumption of a uniform UVB is no longer valid to describe the diffuse HI in the IGM when approaching  $z = 6$ . At these redshifts, our simulated  $\Omega_{\text{HI}}$  only describes the amount of neutral Hydrogen in collapsed/self-shielded systems. Please note that for this calculation we also extended the HI self-shielding prescription of Rahmati et al. (2013), which was originally proposed only up to redshift  $z = 5$ .

In order to estimate the contribution of DLAs to the overall HI budget, the right panel of Figure 3.12 compares the prediction of  $\Omega_{\text{HI}}$  for all the simulations (solid lines) with the DLA comoving mass density  $\Omega_{\text{DLA}}$  (dashed lines), obtained by integrating the HI-CDDF in the range  $20.3 < \log N_{\text{HI}}(\text{cm}^{-2}) < 22$ . The theoretical estimations are compared with

Table 3.2 Contribution from DLA systems to  $\Omega_{\text{HI}}$ .

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

recent data from SDSS (Bird et al., 2017) in black stars. In this case, simulations without molecular cooling included are in better agreement with observations (as for the HI-CDDF shown in Figure 3.11).

We investigate the assumption  $\Omega_{\text{HI}} = \alpha \times \Omega_{\text{DLA}}$ , where  $\alpha$  is a factor that accounts for the numbers of systems that contain neutral Hydrogen with  $N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$ . The values for  $\alpha$  in each simulation (averaged in the redshift range 4–6) are in Table 3.2. In the observational work of Crighton et al. (2015), the authors estimate a 20 per cent contribution from systems with column density below the DLA threshold (or 80% contribution from DLAs) at  $z < 5$ . This assumption seems to be validated with the numerical results here reported.

### 3.9 Discussion and summary

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

### 3.9.1 Effect of wind feedback

The simulations we used in this work are based on the analysis of Tescari et al. (2014), where there is an extensive discussion of the feedback mechanisms implemented and how they successfully reproduce the observations of the SFR at high redshift (e.g. Madau & Dickinson, 2014). When these models are used to study the evolution of different ionic species, we find that it is hard to disentangle the effect of different SN-driven galactic wind prescriptions for some of the statistics considered, in particular the CIV and HI column density distribution functions. The reasons are mostly numerical: probably higher resolution is needed to better resolve patches of enriched gas at relatively low-density in the CGM/IGM (as suggested by our resolution tests). Moreover, the SPH algorithm is not very effective in the mixing of metals. Therefore, only extreme feedback models would produce significant differences in the observables at high redshift. As a result, when calculating the column densities of metal absorption lines with Voigt profile fitting at  $z > 5$ , it is not possible to distinguish between momentum-driven and energy-driven winds (MDW and EDW). The dispersion in the relationships among the metal absorption lines (see Figure 3.8) is too high to provide definitive hints on the evolution of the ionization states due to one particular wind model.

Nonetheless, a clear difference between the numerical estimation of the total Carbon content,  $\Omega_C$ , in EDW and MDW models arises at  $z \leq 6$  (Figure 3.7). This is due to the fact that EDW quenches the overall star formation in galaxies more effectively than MDW, and consequently the production of Carbon in the simulation is 0.2–0.3 dex lower at  $z = 4$ .

Future observational campaigns will allow us to discriminate the feedback models. By increasing the number of metal absorption lines detected, it is possible to compare the predictions from different feedback prescriptions and therefore, derive constraints on the chemical enrichment at high redshift. In addition, a larger number of observational absorption features will span in a wider column density range and provide a more significant statistical distribution of the absorptions, that ultimately gives a tighter constraint to compare the predictions from different models of the column density distribution function and cosmological mass density of the ions.

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

As expected, the implementation of low-temperature metal and molecular cooling affects mostly neutral Hydrogen statistics. In particular, molecular cooling suppresses HI in regions where the majority of it is self-shielded, mainly because of the conversion of HI to

$\text{H}_2$  at high densities. This effect is important when computing the HI column density distribution function and cosmological mass density. Interestingly, the introduction of molecular cooling improves the estimation of  $\Omega_{\text{HI}}$  at  $z \sim 4\text{--}5$  (left panel of Figure 3.12), whereas, when the HI-CDDF is computed in the DLA regime,  $20.3 < \log N_{\text{HI}} (\text{cm}^{-2}) < 22$ , runs without molecular cooling provide a better match to new SDSS observations (Bird et al., 2017) of  $\Omega_{\text{DLA}}$  (right panel of Figure 3.12) and the distribution function at  $z = 4$  (Figure 3.11). This apparent discrepancy is understood when considering the differential HI cosmic mass density (Figure 16 in Crighton et al. (2015)). At  $\log N_{\text{HI}} = 21$  there is the largest contribution of HI from DLA systems. Therefore, DLA systems with this column density contribute more to the calculation of the cosmic mass density of neutral Hydrogen. Simulations with molecular cooling are more efficient converting atomic to molecular Hydrogen and are not dominant at this column density, leading on a lower estimation of the HI mass density at  $z = 4$ .

Once the feedback prescription is fixed, simulations with low-T metal and molecular cooling included produce slightly more CIV than the other runs (Figure 3.5 and left panel of Figure 3.6). The effect is less important for CII and not visible for the total Carbon content and the dispersion in the relationships among the metal absorption lines (Figures 3.7 and 3.8).

### 3.9.3 Limitations

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

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

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

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

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

### 3.10 Conclusions

---

We have introduced high-resolution hydrodynamical simulations to study the physical and chemical state of the IGM at the end of the Epoch of Reionization. We have tested the effect of a low-temperature metal and molecular cooling model and two prescriptions for galactic wind feedback on the evolution of metal absorption lines in the redshift range  $4 < z < 8$ . We have also studied the HI cosmological mass density at  $z < 6$ . Our theoretical predictions are consistent with the available observations at high redshift.

The evolution of the high and low ionization states of Carbon is the main focus of our work. The drop in the cosmological mass density of CIV,  $\Omega_{\text{CIV}}$ , from  $z = 4$  to  $8$  is due to the combined effect of a change in the ionization state of the gas and the decreased metallicity of the IGM (Figure 3.5). In fact, the total Carbon comoving mass density at  $z = 8$  is more than a factor of 10 lower than at  $z = 4$  (Figure 3.7), and the simulations show a consistent transition to more neutral states of metals at high redshift (as shown in Figure 3.6). Most notably, the crossover between  $\Omega_{\text{CIV}}$  and  $\Omega_{\text{CII}}$  happens at  $z \sim 6-6.5$ , in agreement with available observations and the current paradigm of the tail of Reioniza-

tion.

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

The simulated CII exhibits a bimodal distribution with large absorptions in and around galaxies, and some traces in the lower density IGM. These latter correspond to systems detected with current observations (D’Odorico et al., 2013; Becker et al., 2006). We predict that, at high redshift, most of the high column density CII, with  $\log N_{\text{CII}} (\text{cm}^{-2}) > 15$ , should lie in proximity of galaxies at impact parameters of order  $20 \text{ ckpc}/h$  (Figure 3.9) and has not been detected yet due to the low likelihood of reaching these overdense regions with lines of sight towards distant QSOs.

We have studied the column density relationships among different ionic species (Figure 3.8). High ionization states (like  $N_{\text{CIV}}$ ) are reasonably well described by our simulations with the adopted HM12 UV background. On the other hand, a comparison with observations in the literature shows that the low ionization states are not well represented, regardless of the feedback model implemented.

Finally, our simulations are in good agreement with observations of the HI column density distribution function at  $z = 4$  (Figures 3.10 and 3.11) and the HI cosmological mass density,  $\Omega_{\text{HI}}$ , at  $4 < z < 6$  (Figure 3.12). We validate the estimate made in the observational work of Crighton et al. (2015) that DLA systems contribute to  $\sim 80\%$  of  $\Omega_{\text{HI}}$  (see Table 3.2).

In Chapter 4 (García et al., 2017b), we have explored the likelihood of reproducing the observed Ly $\alpha$  emitter galaxy – CIV absorption pair detected by Díaz et al. (2015) and studied the physical processes that produced the metal enrichment in the IGM at  $z \geq 5.6$ . Future work will be focused on analyzing more metals and ions to have a broader perspective on the evolution of the IGM at the end of the EoR. Some of the column density relationships are extremely sensitive to the ionizing background, therefore varying the normalization and hardness of the UVB is the next step to obtain a more realistic description of this cosmic era.

## Acknowledgements

---

Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. The numerical simulations were run using the Raijin distributed-memory cluster from the National Computational Infrastructure (NCI) facility, post-processed with the Edward HPC machine from the University of Melbourne and analyzed with the gSTAR supercomputer at Swinburne University of Technology. This work was supported by the Flagship Allocation Scheme of the NCI National Facility at the ANU. The authors acknowledge CAASTRO for funding and allocating time for the project *Diagnosing Hydrogen Reionization with metal absorption line ratios* (fy6) during 2015 and 2016. L.A. García thanks Valentina D’Odorico and Neil Crighton for providing their observational data in private communications and Jeff Cooke for the insightful discussions. E. Ryan–Weber acknowledges ARC DP 1095600. We also thank the anonymous referee for providing valuable comments that considerably improved our manuscript and Klaus Dolag for his help with the technical details of the SPH code.

# 4

## Theoretical study of an LAE-CIV absorption pair at $z = 5.7$

### 4.1 Abstract

---

We present a theoretical model to predict the properties of an observed  $z = 5.72$  Lyman  $\alpha$  emitter galaxy – CIV absorption pair separated by 1384 comoving kpc/ $h$ . We use the separation of the pair and an outflow velocity/time travelling argument to demonstrate that the observed galaxy cannot be the source of metals for the CIV absorber. We find a plausible explanation for the metal enrichment in the context of our simulations: a dwarf galaxy with  $M_{\star} = 1.87 \times 10^9 M_{\odot}$  located 119 comoving kpc/ $h$  away with a wind velocity of  $\sim 100$  km/s launched at  $z \sim 7$ . Such a dwarf ( $M_{UV} = -20.5$ ) is fainter than the detection limit of the observed example. In a general analysis of galaxy – CIV absorbers, we find galaxies with  $-20.5 < M_{UV} < -18.8$  are responsible for the observed metal signatures. In addition, we find no correlation between the mass of the closest galaxy to the absorber and the distance between them, but a weak anti-correlation between the strength of the absorption and the separation of galaxy – absorber pairs.

### 4.2 Introduction

---

There is an intrinsic connection between the formation and evolution of galaxies in the Universe and the properties of the gas in the intergalactic medium (IGM), such as its metallicity and ionization state. These properties are successfully reproduced by simulations that employ feedback models, in forms of radiative and kinetic winds, able to chemically enrich the outer regions of galaxies (Shen et al., 2013). The detection of galaxies that shape observations of the IGM provides information on the interplay of the outflowing

winds from the star-forming regions and the gas in the outskirts that is photoionized and polluted by these feedback mechanisms.

One of the best ways to explore the properties of the gas in the IGM is by the detection of CIV absorption systems. Among other high-ionization states, CIV traces low-density regions in the IGM at temperatures of  $T \geq 10^4$  K (Oppenheimer et al., 2009; Tescari et al., 2011; Cen & Chisari, 2011; Finlator et al., 2015; Rahmati et al., 2016). In addition, CIV is a good indicator of ionized regions due to its large ionization potential energy and its detection in the spectra of background high redshift quasars (e.g. Danforth & Shull, 2008; Ryan-Weber et al., 2009; Cooksey et al., 2010; Simcoe et al., 2011; D’Odorico et al., 2013; Boksenberg & Sargent, 2015; D’Odorico et al., 2016) from  $z \sim 0$  to 6.

At low and intermediate redshift, extensive observational campaigns are carried out with the goal to find, identify and characterize galaxy – CIV absorbers pairs: at  $z \sim 0.05$  by Burchett et al. (2016) and at  $z \sim 2-3$  (Steidel et al., 2010; Turner et al., 2014). Observing these galaxy – CIV absorber systems when the Universe is approaching the tail of Reionization ( $z \sim 6$ ), might give us a hint of the connection between the galaxies that drove the Reionization and the metal enrichment in the very early Universe. For example, Oppenheimer et al. (2009) explored the physical environment of the absorbers assuming different scenarios for the ionizing background and found that CIV absorbing gas is primarily intergalactic at  $z \sim 6$ . The absorbers are distributed at distances up to 200 physical kpc from the parent galaxy. In their HM2001 model, the CIV absorbers are only associated with galaxies of  $M_\star \sim 10^9 M_\odot$ . In addition, work from Oppenheimer et al. (2009); Finlator et al. (2016) showed that local sources can affect the strength of CIV absorption.

With the detection of a high redshift ( $z = 5.72$ ) Lyman  $\alpha$  emitting galaxy by Díaz et al. (2015, hereafter Díaz et al. (2015)) located 212.8 physical kpc/ $h$  from a high column density ( $\log N_{\text{CIV}}(\text{cm}^{-2}) = 14.52$ ) CIV absorber (D’Odorico et al., 2013), there is now observational evidence of a galaxy – absorber pair at this redshift. This detection opens the question: is the CIV absorber due to an outflowing wind from the nearby star forming galaxy or does it instead arise due to an undetected dwarf galaxy closer to the line of sight?

A theoretical comparison with different feedback prescriptions at  $z = 6$  by Keating et al. (2016) finds that their models are unable to reproduce the configuration observed by Díaz et al. (2015), and that their strong absorbers are linked to nearby galaxies ( $< 100$  pkpc) with very high halo masses ( $\log M_h/M_\odot \geq 10$ ). The low incidence rate of strong CIV absorption systems makes it doubly difficult to simulate the observed scenario.

This Chapter is devoted to exploring the likelihood of reproducing a system as detected

by Díaz et al. (2015) and to giving a plausible explanation for the enrichment. In addition, we make some theoretical predictions in order to guide the analysis of future observations. The work follows the methodology and models described in Chapter 3 (García et al., 2017a). The numerical simulations were run using a customized version of the smoothed particle hydrodynamics (SPH) code GADGET-3 (Springel, 2005) with cosmological parameters from Planck Collaboration et al. (2015):  $\Omega_{0m} = 0.307$ ,  $\Omega_{0b} = 0.049$ ,  $\Omega_{\Lambda} = 0.693$  and  $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (or  $h = 0.6774$ ). The simulated run Ch 18 512 MDW has comoving box size and softening of  $18 \text{ Mpc}/h$  and  $1.5 \text{ kpc}/h$ , respectively, and includes  $2 \times 512^3$  dark matter and gas particles. We assume a momentum-driven wind feedback model with a wind mass-loading factor  $\eta = 2 \times \frac{600 \text{ km s}^{-1}}{v_w}$  and a velocity of  $v_w = 2 \sqrt{\frac{GM_h}{R_{200}}} = 2 \times v_{\text{circ}}$ . The simulation has been post-processed with a field radiation due to the cosmic microwave background (CMB) and the Haardt & Madau (2012) ultraviolet/X-ray background from quasars and galaxies with saw-tooth attenuation (HM12), an effective prescription for HI self-shielding according to Rahmati et al. (2013), a photoionization modeling for CIV using CLOUDY v8.1 (Ferland et al., 2013) for optically thin gas and Voigt profile fitting with the code VPFIT v.10.2 (Carswell & Webb, 2014). As described in Chapter 3 (García et al., 2017a), our numerical simulations correctly reproduce observed global statistics of CIV absorbers, namely, the column density distribution function and the cosmological mass density  $\Omega_{\text{CIV}}$ . Throughout the chapter, we use the prefix c for comoving and p for physical distances. Please note that in our simulation the closest snapshot to the  $z = 5.72$  Díaz et al. (2015) system is at  $z = 5.6$ .

### 4.3 The LAE – CIV absorption pair of Díaz et al. 2015

In this section, we investigate the likelihood of reproducing the galaxy – CIV absorber pair detected by Díaz et al. (2015) with theoretical models. Specifically, a CIV absorption system in the line of sight to the background quasar J1030+0524, with a column density of  $\log N_{\text{CIV}} (\text{cm}^{-2}) = 14.52 \pm 0.08$  (D’Odorico et al., 2013) at a distance of  $212.8^{+14}_{-0.4} h^{-1} \text{ pkpc}$  ( $1384 \text{ ckpc}/h$  in the adopted cosmology) from a Lyman  $\alpha$  emitter galaxy (LAE) of  $M_{\text{UV}} = -20.7$  at  $z = 5.72$ , with inferred parameters  $\log (M_{\star}/M_{\odot}) = 9.4$  (using the stellar mass function for  $z = 6$  galaxies from Song et al., 2016) and  $\log (M_h/M_{\odot}) = 10 - 11$  (derived from  $z \sim 6.6$  LAE clustering from Ouchi et al., 2010). The possible scenarios that the authors propose to explain the metal enrichment of the region (in particular with Carbon) are either a very powerful galactic outflow that left the LAE at earlier times or an undetected dwarf galaxy closer to the absorber.

In the simulations, the chemical enrichment is driven by supernovae that produce and

Table 4.1 Wind velocity ( $v_w$ ) and travel time ( $t$ ) to enrich a region at 1300 ckpc/ $h$  from our most massive galaxy ( $M_h$ ).

Box size (cMpc/ $h$ )	$M_h$ ( $\times 10^{11} M_\odot$ )	$v_w$ (km/s)	$t$ (Gyr)
18	4.9	446	0.73

expel metals to the outer regions of galaxies. Wind velocities in the adopted feedback model depend on the halo mass of the galaxies. Therefore, only the most massive objects produce winds powerful enough to pollute regions at more than 1 cMpc/ $h$ , but these objects are quite rare at high redshift, especially in small simulations. We calculate the wind velocity and travel time of the most massive galaxy in the box in Table 4.1 to set an upper limit on the time required for an outflow to reach a region at a distance of 1300 ckpc/ $h$  (comparable to the Díaz et al. (2015) example). The results from Table 4.1 suggest that not even the most massive galaxy is able to produce winds with the velocity required to travel 1300 ckpc/ $h$  in a reasonable time. The time displayed in Table 4.1 is comparable with the age of the Universe at  $z = 5.6$ . Thus, for this option to be viable, galaxies would need to be formed at an age ( $z \geq 30$ ) incompatible with the current paradigm of galaxy formation. Therefore, we rule out the possibility that an outflow produced by the LAE enriched a region at 1384 ckpc/ $h$  in the context of the feedback model implemented in our simulations. Nevertheless, alternative configurations could lead to a different result: more aggressive prescriptions for supernova-driven outflows or larger halo masses that produce higher wind velocities. The latter condition could be fulfilled in our simulations with a galaxy halo mass of  $\sim 1.5 \times 10^{12} M_\odot$ . However, galaxies have only been observed with stellar masses up to  $10^{10} M_\odot$  at  $z \sim 6$  (Song et al., 2016). Assuming a mass-to-light ratio of 10 (inferred from the simulations at this redshift), this scenario is still out of reach by at least an order of magnitude.

This test also reveals interesting details of the model: the chemical enrichment is driven by galaxies nearby the absorbers, since galactic outflows cannot travel a distance larger than about 1 cMpc in less than 0.5 Gyr and chemically enrich the IGM. At this redshift, active galactic nuclei are the only sources energetic enough to produce winds that travel that far.

Next, we explore the dwarf galaxy scenario, in which the presence of another galaxy closer to the absorber, and fainter than the detection limits for optical observations, is responsible for the enrichment. The upper limit for an undetected galaxy is:  $M_{UV} =$

-20.5, corresponding to  $M_\star = 10^{9.3} M_\odot$  (Song et al., 2016). Independent narrow band observations of Ly $\alpha$  emission set a limit on the star formation rate,  $\text{SFR} = 5\text{-}10 M_\odot/\text{yr}$ . The goal of this analysis is to find mock scenarios that resemble the conditions of the LAE detected in Díaz et al. (2015) and, if possible, confirm or rule out the hypothesis of an undetected dwarf galaxy.

At  $z = 5.6$ , the simulated galaxies have halo masses in the range of  $7.83 < \log M_h(M_\odot) < 10.83$ . We classify them in the following categories: LAE candidates ( $M_h \geq 1.48 \times 10^{10} M_\odot$ ) and dwarf galaxies ( $1.48 \times 10^9 M_\odot < M_h < 1.48 \times 10^{10} M_\odot$ ) to provide a fair comparison with the observational parameters. Galaxies with halo mass below  $1.48 \times 10^9 M_\odot$  are excluded from the statistics because of their low resolution. We stress that we do not follow the Ly $\alpha$  emission of any galaxy. The criterion to describe the galaxies is based only on their stellar and halo masses.

Stochastically, 1000 galaxies are selected from the friends-of-friends catalogue and from each one, a line of sight is projected with a random impact parameter  $d$  up to  $1500 \text{ ckpc}/h$ . The column density  $N_{\text{CIV}}$  and the position of the absorption in the box is recovered in each case. We focus on CIV absorption systems with  $\log N_{\text{CIV}}(\text{cm}^{-2})$  in the range of 14-15. We then look for all the galaxies around the absorption (our absorption spectra have an uncertainty of  $7 \text{ km/s}$ , or  $\pm 49 \text{ ckpc}/h$ , along the line of sight). This set of galaxies is identified and sorted by mass and distance (the 3D distance can be decomposed in components parallel and perpendicular to the line of sight). The array in mass allows us to tag the galaxy as LAE or dwarf (or just exclude it according to the resolution criterion mentioned above), whereas the distance array defines the closest galaxies to the absorption. The cross-matching of the arrays leaves 4 systems with a configuration LAE – CIV absorption (– dwarf galaxy), close to the observational arrangement.

In particular, one system drew our attention for its geometry being closest to the Díaz et al. (2015) observations. We show the configuration LAE – CIV absorption – dwarf galaxy in Figure 4.1. The CIV absorption feature has  $\log N_{\text{CIV}} (\text{cm}^{-2}) = 14.3 \pm 1.1$ , the LAE has  $M_h = 1.54 \times 10^{10} M_\odot$  at a 3D distance of 1296 ckpc/ $h$  from the CIV system. The closest galaxy in the field, a dwarf galaxy of  $M_\star = 1.87 \times 10^9 M_\odot$ ,  $M_h = 9.67 \times 10^9 M_\odot$ ,  $\text{SFR} = 0.07 M_\odot/\text{yr}$  is located at a 3D distance from the absorber of 119 ckpc/ $h$ . A dwarf galaxy with this stellar mass lies just below the 50% complete value for  $M_{\text{UV}}$  images of Díaz et al. (2015) at  $M_\star = 2 \times 10^9 M_\odot$ . In this case, the velocity of an outflow produced by the dwarf galaxy is  $v_w \sim 100$  km/s, leading to a travel time of  $\sim 260$  Myr (*i.e.* the outflow was launched at  $z \sim 7$ ).

The spatial configuration displayed in Figure 4.1 favors an undetected dwarf galaxy in the observations of Díaz et al. (2015) and confirms that, in the context of our feedback model, the chemical enrichment is caused by nearby galaxies. In addition, these results emphasize that the metal absorption line systems offer the best technique for detecting galaxies beyond the limits of imaging at high redshifts.

As a final point, we mention that in the Díaz et al. (2015) observations, there is a small probability that an associated galaxy of any mass is obscured by the quasar. The highest resolution image of this quasar is an HST/ACS  $z$ -band image taken by Stiavelli et al. (2005). Conservatively, we estimate that the quasar light would prevent the detection of galaxies within a radius of 1 arcsec. Thus, 11% of the circular area within 18 pkpc/ $h$  (119 ckpc) from the quasar line of sight is obscured. At this redshift, a separation of 18 pkpc/ $h$  corresponds to 3.07 arcsec.

In order to understand the incidence rate of galaxies around the CIV absorption in the particular scenario discussed above, we plot the distribution of 3D distances from the absorption to each galaxy in a range of 1500 ckpc/ $h$  in Figure 4.2. There are many more dwarf galaxies closer to the CIV absorber than the LAE.

Furthermore, the probability of finding a randomly placed dwarf galaxy at 119 ckpc/ $h$  from a metal absorber is 5.1%, knowing that the number density of dwarf galaxies in the sample is 0.46 objects per  $\text{cMpc}^3$  and the volume where these pairs are likely to occur is about  $0.02 \text{ cMpc}^3$ .

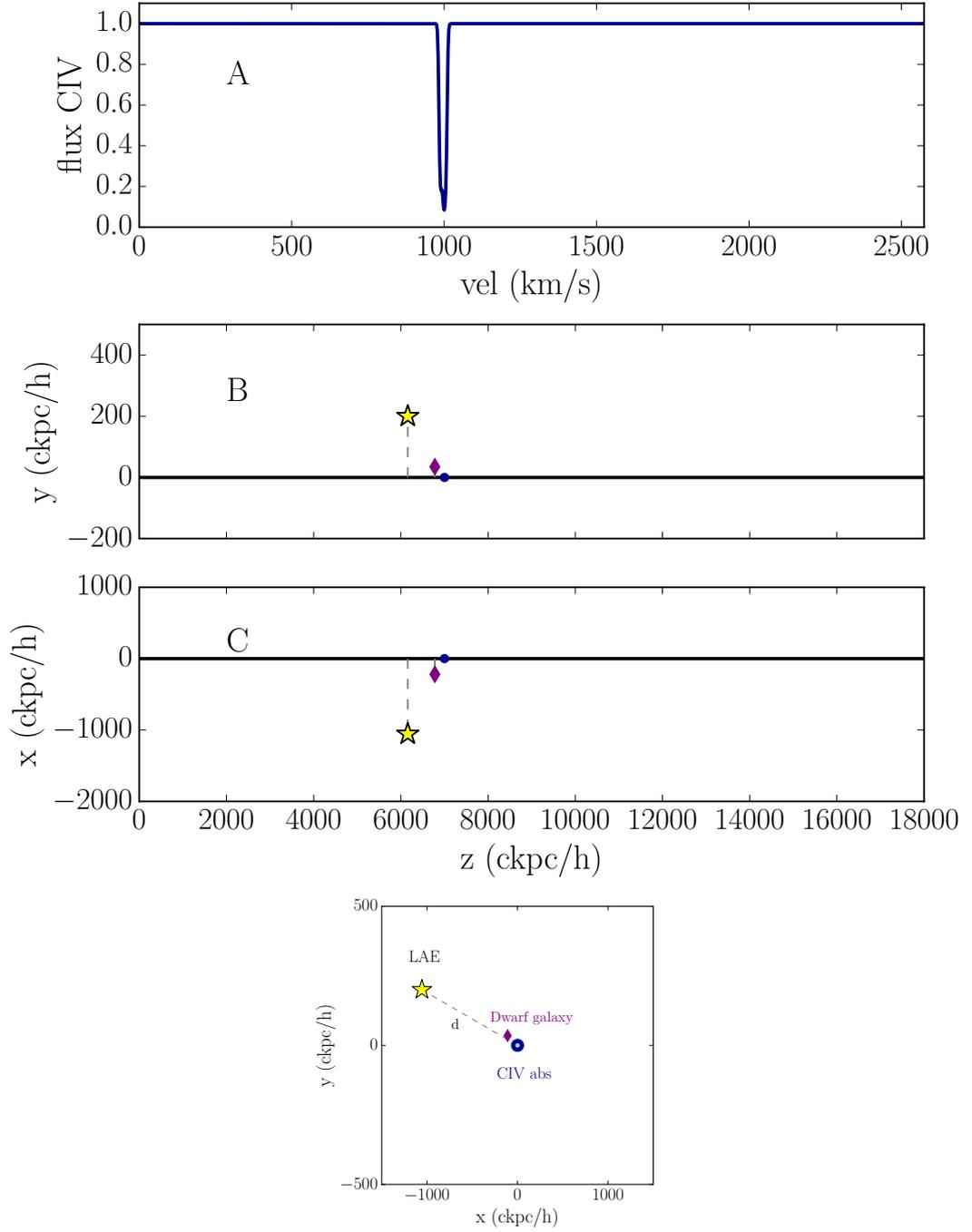


Figure 4.1 Theoretical example of the observed LAE – CIV absorption pair. In this window, we test one of the hypotheses proposed in Díaz et al. (2015): the existence of an undetected dwarf galaxy in the field of the LAE. We identify all the CIV absorptions at  $z = 5.6$  with  $\log N_{\text{CIV}} (\text{cm}^{-2})$  in the range of 14–15 and the closest galaxies in the surroundings of each absorption. The upper panel (A) shows a feature with  $\log N_{\text{CIV}} (\text{cm}^{-2}) = 14.3 \pm 1.1$  on a line of sight traced along the  $z$  direction. Panel B shows the position of the strong CIV absorption (blue point), an LAE (yellow star) with  $M_h = 1.54 \times 10^{10} M_\odot$  and a 3D distance of 1296 ckpc/h from the absorption, and the closest galaxy in the field (purple diamond), a dwarf galaxy of  $M_\star = 1.87 \times 10^9 M_\odot$  at a 3D distance from the absorber of 119 ckpc/h, that could not have been detected in the observations. Panel C shows the same configuration in a different edge-on projection ( $xz$ ). The bottom panel displays the physical disposition of the system face-on: the line of sight from the quasar (in the background) contains the absorption and it is orthogonal to the plane of the image.

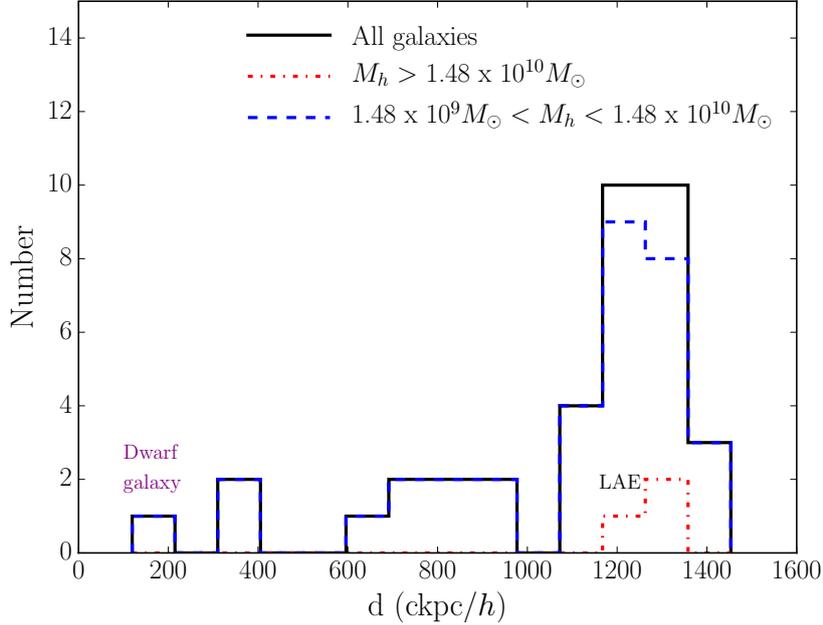


Figure 4.2 Distribution of galaxies around the CIV strong absorption shown in Figure 4.1 up to a 3D distance of 1500 ckpc/h. The black histogram accounts for all galaxies, whereas the red dashed–dotted line shows the systems with masses  $M_h \geq 1.48 \times 10^{10} M_\odot$  (LAE) and the blue dashed line the contribution from dwarf galaxies  $1.48 \times 10^9 < M_h/M_\odot < 1.48 \times 10^{10}$  in the box. The galaxies below the mass resolution limit of  $1.48 \times 10^9 M_\odot$  are not shown.

#### 4.4 Galaxy – absorber connection

We have shown a theoretical realization of the observed LAE – CIV absorption pair at  $z = 5.6$ . Now, we want to explore the physical connection of these systems (if any), using the properties of the galaxies and the column densities of the absorbers in the simulated box. For the purpose of the analysis of this section, which is made with galaxies in the simulations above the mass resolution limit, we define high ( $M_h \geq 1.48 \times 10^{10} M_\odot$ ), intermediate ( $6.64 \times 10^9 < M_h/M_\odot < 1.48 \times 10^{10}$ ) and low mass ( $1.48 \times 10^9 < M_h/M_\odot < 6.64 \times 10^9$ ) galaxies.

The connection of the galaxy – CIV systems can be studied using galaxy – absorber pairs with  $\log N_{\text{CIV}} (\text{cm}^{-2}) > 13$  and taking into account the nearest galaxy to the absorption. We implicitly assumed that this nearest galaxy is the main source of metal enrichment for the absorber.

We construct the overall distribution of the distance from each absorption to its closest galaxy in the simulated box. From the initial sample of 1000 lines of sight, 182 CIV

absorbers have a galaxy companion above the mass resolution. In Figure 4.3, we display this distribution with respect to the mass of the galaxy.

There are not high mass galaxies in the random selection from the friends–of–friends catalogue because just 2% of the objects are in the range of mass  $M_h \geq 1.48 \times 10^{10} M_\odot$ , which makes it extremely hard to encounter a random line of sight through these galaxies. On the other hand, almost 90% of the galaxies in the box are intermediate and low mass ( $1.48 \times 10^9 < M_h/M_\odot < 1.48 \times 10^{10}$ ). Therefore, it is not surprising that the closest galaxy to each  $N > 10^{13} \text{ cm}^{-2}$  CIV absorber has a mass in this range.

Additionally, intermediate and low mass galaxies are typically found at a mean distance of 520 ckpc/h (78.8 pkpc/h) and 700 ckpc/h (106.1 pkpc/h), respectively. This result is only partially in agreement with findings of Oppenheimer et al. (2009), for high stellar–mass galaxies in their HM2001 model: strong CIV absorptions at  $z = 6.0$  are associated with galaxies with  $M_\star \sim 10^9 M_\odot$  and the typical galaxy – absorption separation is less than 100 pkpc. It is important to clarify that, in addition to the intrinsic difference between HM2001 and HM12 backgrounds, their highest mass galaxies are close to our mass resolution (lower) limit. Nonetheless, we confirm that CIV is mostly found in the IGM (García et al., 2017a; Oppenheimer et al., 2009).

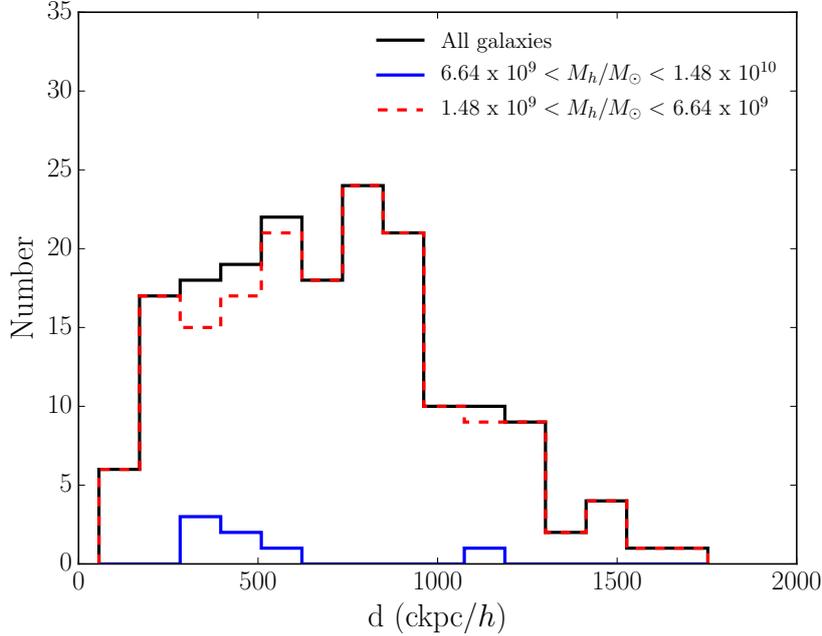


Figure 4.3 Mass distribution of the 3D distance from the CIV absorptions to the closest galaxy. The black histogram accounts for all the galaxies. There are not high mass galaxies ( $M_h \geq 1.48 \times 10^{10} M_\odot$ ), meaning that, in all cases, the closest galaxies from the absorptions lay in the category of dwarfs. Intermediate mass galaxies are presented in blue solid line (7) and low mass galaxies with dashed red (175). The galaxies below the mass resolution limit  $1.48 \times 10^9 M_\odot$  are not presented. In cases where  $d \geq 1500$  ckpc/h, the enrichment is most likely due to an unresolved galaxy.

Finally, Figure 4.3 reveals that the CIV strong absorptions at high redshift are driven by dwarf galaxies with  $M_\star$  spanning the range  $10^{8.46-9.42} M_\odot$  and  $\text{SFR} = 0.01-2.5 M_\odot/\text{yr}$ . These parameters allow us to infer an absolute magnitude  $M_{\text{UV}}$  in the range of  $(-20.5, -18.8)$  (using the  $M_\star - M_{\text{UV}}$  relation at  $z \sim 6$  from Song et al., 2016) that should be achieved with future surveys. The detection and study of these faint galaxies is fundamental, since they are believed to provide the largest contribution of photons to complete the Reionization of Hydrogen (Robertson et al., 2015; Liu et al., 2016).

Figure 4.4 shows the distribution of 3D distance from the CIV absorptions to the closest galaxy as a function of column density of the absorption. The 182 absorbers span the range of  $13 \leq \log N (\text{cm}^{-2}) \leq 16$ . From this sample, there are 73 absorbers with  $13 \leq \log N (\text{cm}^{-2}) < 14$ , 31 in the range of  $14 \leq \log N (\text{cm}^{-2}) < 15$  and 78 with  $\log N (\text{cm}^{-2}) \geq 15$ .

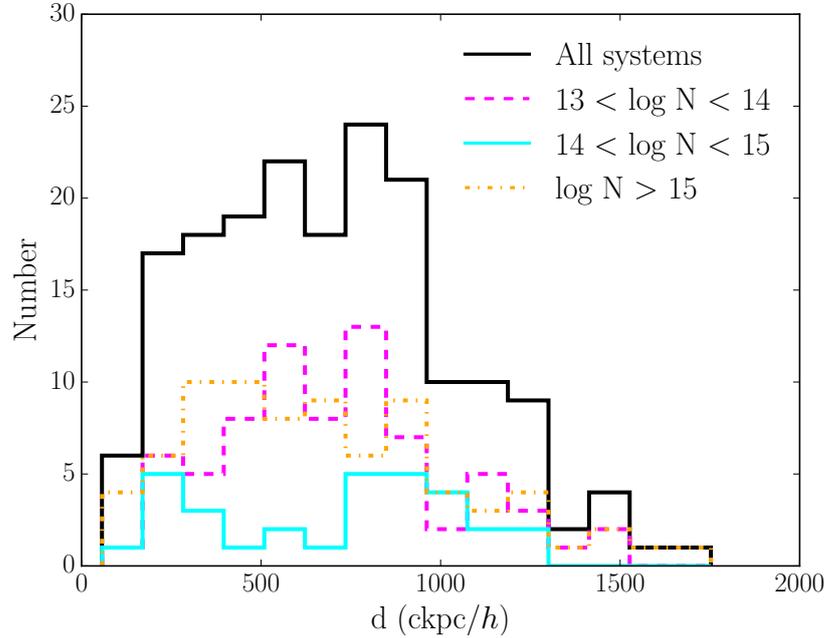


Figure 4.4 Distribution of the 3D distance from the CIV absorptions to the closest galaxy with respect to the column density of the absorption. The black histogram accounts for all the column densities in the range of  $13 \leq \log N (\text{cm}^{-2}) \leq 16$ . The magenta dashed line shows systems with  $13 \leq \log N (\text{cm}^{-2}) < 14$ , the cyan solid line the systems with  $14 \leq \log N (\text{cm}^{-2}) < 15$  and in orange dashed-dotted line systems with column densities  $\log N (\text{cm}^{-2}) \geq 15$ .

The mean distance in the three cases is respectively: 710.6, 705.5 and 680.6 ckpc/h. This result indicates that there is a weak negative correlation between the column density of the CIV absorptions and the separation of absorber – galaxy pairs. Finally, it is worth noting that strong CIV absorptions in our simulations are produced not only by very close structures with large masses ( $M_h \sim 10^{10-11} M_\odot$ , which we rarely find within this box size), but also for small galaxies.

## 4.5 Conclusions

We have explored the likelihood of reproducing the observed LAE – CIV absorption pair detected by Díaz et al. (2015) and studied the physical processes that produced the metal enrichment in the IGM at  $z \sim 5.6$ . In the context of our feedback model, we rule out the scenario of an outflow produced by the observed LAE at a distance of 1384 ckpc/h at early times.

Instead, our simulations support a scenario in which a dwarf galaxy, undetected in the field, very close to the absorber (119 ckpc/h), with  $M_\star = 1.87 \times 10^9 M_\odot$ ,  $M_h = 9.67 \times$

$10^9 M_\odot$  and  $\text{SFR} = 0.07 M_\odot/\text{yr}$ , is responsible for the CIV absorption.

Additionally, we find that the main drivers of the chemical enrichment of the IGM and CIV strong absorptions at  $z = 5.6$  are dwarf galaxies, which are mostly observed at a mean distance of  $700 \text{ ckpc}/h$  ( $106.1 \text{ pkpc}/h$ ). These galaxies have stellar masses  $M_\star$  in the range  $10^{8.46-9.42} M_\odot$  and  $\text{SFR}=0.01-2.5 M_\odot/\text{yr}$ . From these parameters, we infer an absolute magnitude  $M_{\text{UV}}$  in the range of  $(-20.5,-18.8)$ , derived by combining results of Song et al. (2016) with the Kennicutt relation. Future observations and investigations of these faint galaxies are fundamental, since they are believed to be the largest contributors of photons to the Reionization of Hydrogen.

When exploring the galaxy – absorber connection at high redshift, we find no correlation between the mass of the closest galaxy to the absorber and the distance between them. On the other hand, we encounter a weak negative correlation between the column density and the separation of the galaxy – absorber pairs. The largest column densities are preferentially seen when separations are smallest.

## Acknowledgements

---

Parts of this research were conducted by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020. This work was supported by the Flagship Allocation Scheme of the NCI National Facility at the ANU. The authors acknowledge CAASTRO for funding and allocating time for the project **Diagnosing Hydrogen Reionization with metal absorption line ratios** (fy6) during 2015 and 2016. We also thank the anonymous referee for their insightful comments. ERW acknowledges ARC DP 1095600.

# 5

## Future directions

### 5.1 Abstract

---

Here we present future directions for the use of simulated metal absorption lines at high redshift inspired by questions and limitations raised in Chapters 3 and 4. Specifically, we ask: i) what is the contribution of CIII to the total Carbon budget? ii) How do our previous results change if the UV background is modified? iii) How does the assumption of solar metallicity change the CLOUDY photoionization results?

### 5.2 CIII evolution compared with CII and CIV

---

In Chapter 3, the evolution of CII and CIV was extensively discussed and compared with observations available at high redshift. However, these ionization states are not the only ones where Carbon can be found. In particular, at high redshift, CI should be predominant, but it has not been detected yet, due to its small oscillator strength (CI 1139 has a oscillator strength of  $5.42 \times 10^{-4}$  whereas CIV 1548 is 0.1899). Also, CIII and CV might have a non-negligible contribution to the total Carbon budget but, again, this hypothesis can only be tested with photoionization models. CIII has been detected at low redshift in the Ly $\alpha$  forest, but as discussed in previous Chapters, that region of the spectra is completely saturated by HI absorption at  $z \sim 5$  and above. Therefore, CIII is often neglected in the analysis at high redshift because there are not CIII absorptions detected to calculate an observational mass density.

The CLOUDY photoionization tables run for Chapters 3 and 4 allow us to study other ionization states of Carbon. The doubly ionized state of Carbon, CIII, is quite important in our analysis because it creates a natural bridge between the states previously presented, CII and CIV. Despite the lack of an observational archive of CIII, Carbon atoms could

be significantly present in this ionization state, especially because ionizing photons in the IGM require less energy to generate a transition from CII to CIII than from CIII to CIV (from CII to CIII an energy of 24.38 eV is required, in comparison to the 47.89 eV necessary for CIII to make way to CIV).

Figure 5.1 displays the ionization fractions of CII, CIII and CIV with respect to the density and temperature of the gas at  $z = 6$  in the reference simulation Ch 18 512 MDW. The black region in the low density regime corresponds to zero metallicity, indicating that in this epoch of the Universe, low density gas had not been enriched. Interestingly, CII is

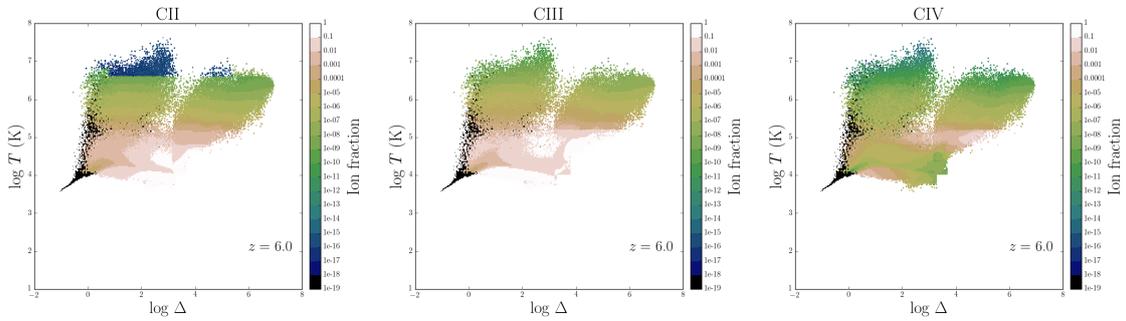


Figure 5.1 Photoionization modelling for states of Carbon at  $z = 6$  in the reference simulation. The panels display the ionization fraction of CII, CIII and CIV (as indicated in the colour bars), as a function of the overdensity and temperature of the gas. The different regions shown here, have been discussed in Chapter 3 Figure 3.1. Null metallicity in the gas is represented in black. Regions at very low densities and voids had not been chemically enriched at this redshift, therefore, no ionization states of Carbon are present. The largest fraction of CII is found in low temperature-high density gas, corresponding to DLAs and regions where HI is highly self-shielded, and, as expected, the amount of CII in shock-heated regimes ( $\log T \text{ (K)} \geq 7$ ) is highly suppressed. There is a visible transition between CII and CIV in regions in the CGM and IGM. In the middle panel, there is a large amount of CIII around  $T \sim 10^5 \text{ K}$ , where this transition has a large contribution in the cooling efficiencies of the gas. Depending on the conditions of the gas, CIII can decay into CII in the CGM, otherwise, at lower densities can favour the production of CIV in regions of the IGM where the ionizing background contribution is high.

found mostly in low temperature-high density gas (DLAs and regions where HI is highly self-shielded) and it is highly suppressed in shock-heated gas ( $\log T \text{ (K)} \geq 7$ ). Instead, CIII is found in gas at  $T \sim 10^5 \text{ K}$ , where its cooling efficiency peaks. CIII is notably important because, depending on the conditions of the gas (temperature and density), it can transition to CII in high density regions, or, it can favour the creation of CIV in regions of the IGM with low density or where the gas is shock-heated.

While comparing the three panels of Figure 5.1, there is a clear evidence that the area covered by CIII in the overdensity–temperature phase diagram is larger than those covered

by CII and CIV. The lack of observations in absorption of CIII, makes a complete analysis of all ionization states of Carbon extremely difficult.

Work done by Simcoe (2011) for Carbon at  $z = 4.3$  with CLOUDY, compares HM01 and other UVB model (Faucher-Giguère et al., 2009). The work shows a modest dependency of the ionization fractions of Carbon on the choice of UV background spectrum used. Moreover, the CIV fraction peaks in the IGM, in agreement with results discussed in Chapter 3. Although our findings at  $z = 6$  are not directly comparable with Simcoe’s work, due to difference in the UVB models and the redshift of the CLOUDY tables used, it is reassuring that his study predicts the prevalence of CIII in conjunction with CIV in the IGM (Figure 11 Simcoe, 2011), compatible with the phase diagram of CIII presented in Figure 5.1. It is important to note that analysis at  $z = 4.3$  can take into account associated HI column density measurements, that make the photoionization modelling more accurate. Instead, we cannot employ HI at  $z \geq 6$ .

There is an additional study that can be done using photoionization modelling in the framework of HM12. In Chapter 3, the cosmological mass density of CII and CIV is calculated in the range of  $13 < \log N_{\text{CII}} (\text{cm}^{-2}) < 15$  and  $13.8 < \log N_{\text{CIV}} (\text{cm}^{-2}) < 15$ , respectively, and compared with lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006). However, one can ask how CII and CIV evolve with respect to other ionization states of Carbon. Using the CLOUDY tables and the fiducial run Ch 18 512 MDW, we calculated the comoving mass density of CIII in the range  $13 < \log N_{\text{CIII}} (\text{cm}^{-2}) < 15$  (the choice of this range is consistent with the CII and CIV observational column density ranges, since there are not observational constraints available on CIII) and compared with  $\Omega_{\text{CII}}/\Omega_{\text{CIV}}$ , as shown in Figure 5.2. The Figure reveals that CIII is about an order of magnitude more abundant than the other two ionization states of Carbon.

Therefore, introducing CIII helps us to integrate our understanding of the evolution of Carbon. The photoionization models offer plenty of information, sometimes unexplored due to the lack of observations to compare with.

### 5.3 Variations of the UVB in post-process

---

Investigations carried out with a uniform UV background (Haardt & Madau, 2012) in Chapter 3 showed that the calculated column densities of the low ionization states (CII, SiII, OI) and the corresponding observables that depend on these quantities (comoving mass density, column density distribution function, etc.) have some room for improvement when compared with observations. There is general agreement that current simulations

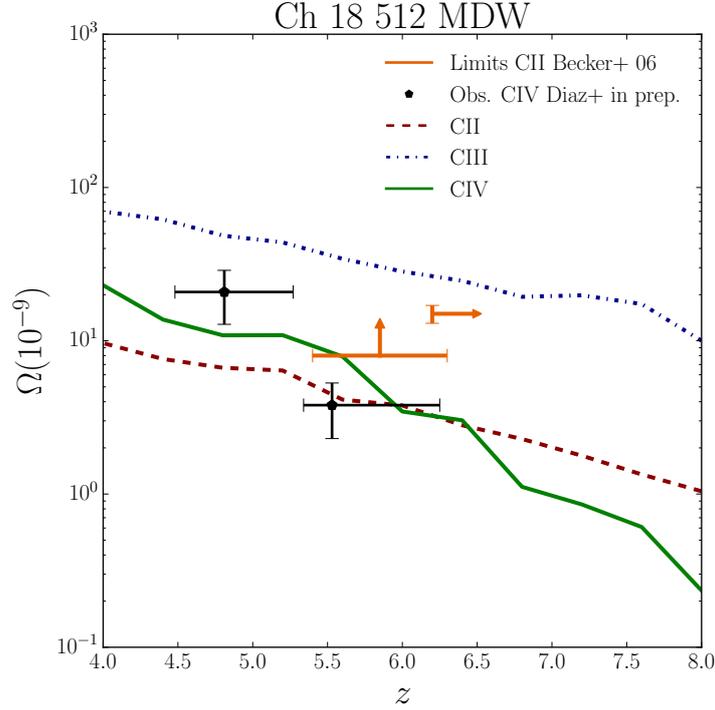


Figure 5.2 Evolution of the CII, CIII and CIV cosmological mass densities in the reference simulation. The solid green line shows the evolution of  $\Omega_{\text{CIV}}$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , the dashed dark red line  $\Omega_{\text{CII}}$  and the dashed-dotted blue line  $\Omega_{\text{CIII}}$ , both in the range  $13.0 < \log N_{\text{CII/CIII}}(\text{cm}^{-2}) < 15.0$ . The orange points with errors represent the observational lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006) and the black pentagons, an observation of  $\Omega_{\text{CIV}}$  by Díaz et al. (in prep), as a reference. The cosmological mass density of these three ions is particularly important because it indicates that other ionization states of Carbon have a larger mass density than the observed ones (CII and CIV) in the range of column densities of the detections.  $\Omega_{\text{CIII}}$  is a dex above than  $\Omega_{\text{CII}}$  in the same range of column densities, and at least 0.5 dex larger than  $\Omega_{\text{CIV}}$ . Notably, the cosmic mass density of these ions just account for absorptions in the range defined by the observations, but large column densities, not detected in the IGM, could increase by orders of magnitude the relative density of each ion transition.

do not have enough resolution on the scale of the absorbers with low ionization states. However, uncertainties on the high  $z$  UVB suggest that varying its normalization is a first step towards a better agreement with the observations at high redshift.

Here we explore the sensitivity of the results presented in Chapter 3 on different ultraviolet/X-ray ionizing backgrounds by modifying some features of the uniform HM12 UVB in post-process. Two tests were performed: 1) a variation of the normalization factor at 1 Ryd and 2) a change of the hardness of the high energy end of the HM12 spectrum. Variations of the UVB spectrum (quasars+galaxies model) from the original HM12 (see Figures 5.3 and 5.4) require modified input files to run new CLOUDY tables. The proce-

ture followed here is explained in Chapter 2 and tested in Chapters 3 and 4, with some small adjustments:

- **Varying the normalization:** the normalization parameter at 1 Ryd is reduced by an order of magnitude below the fiducial value in HM12 at all redshifts. This leads to a softer UVB. Hereafter, the test is referred as  $\log J_\nu - 1$  (see section 5.3.1).
- **Modifying the high energy end of the UVB:** by introducing a very aggressive variation in the slope of the hard part of the quasar model, it is possible to study the effect of very energetic photons (and as a result, a larger contribution in the luminosity) produced by quasars at  $z \sim 6$  on the different ionic transitions (see section 5.3.2).

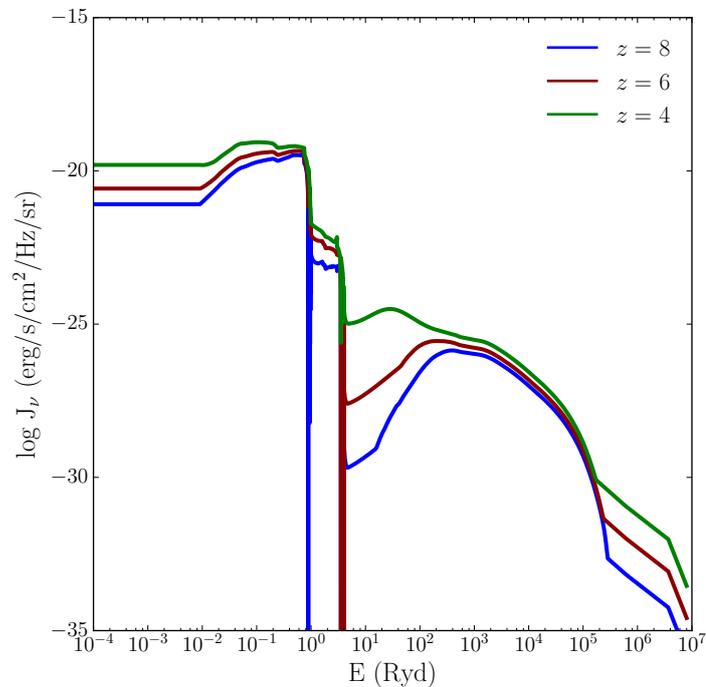


Figure 5.3 UV emissivity vs. energy for the uniform HM12 background at three different redshifts:  $z = 8, 6$  and  $4$  (blue, dark red and green, respectively).

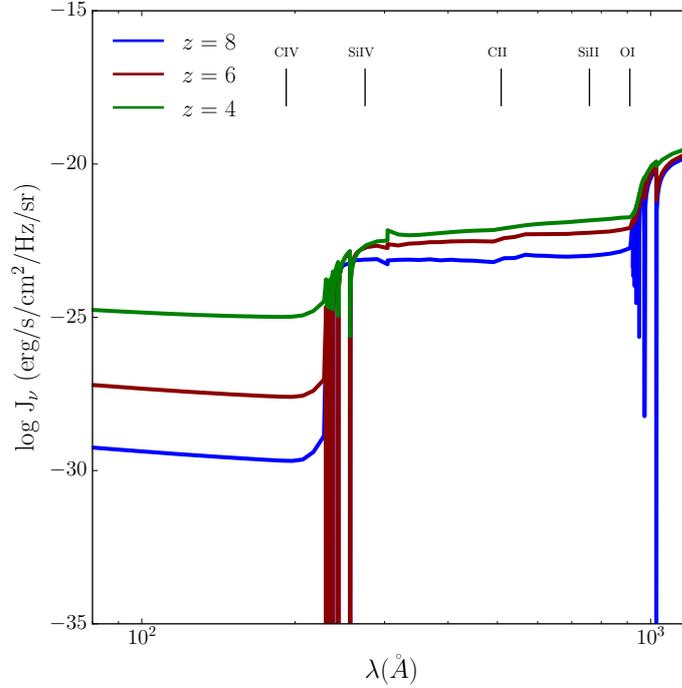


Figure 5.4 UV emissivity vs. wavelength for the uniform HM12 background at three different redshifts:  $z = 8, 6$  and  $4$  (blue, dark red and green, respectively) in the wavelength range where the ion states occur.

### 5.3.1 Change in the normalization of HM12

It is a reasonable expectation that the presence of a softer UVB input than the uniform HM12 in the photoionization model would favour low ionization states and more neutral states would show large incidence rates. In order to test this hypothesis, the UV emissivity  $J_\nu$  at 1 Ryd is reduced by one dex compared with the value defined by the uniform HM12 at all redshifts. This leads to a softer UVB. A schematic figure of the effect in the overall spectrum at  $z = 6$  is shown in Figure 5.5.

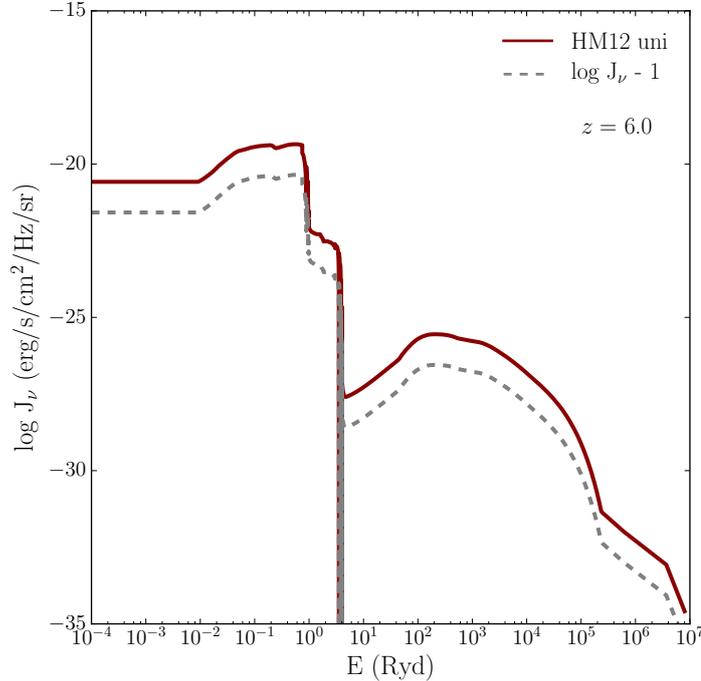


Figure 5.5 UV emissivity vs. energy at  $z = 6$ . The dark red curve shows the uniform HM12 and the grey dashed line a variation of - 1 dex in the overall normalization of the spectrum ( $\log \mathbf{J}_\nu - 1$ ).

In order to avoid introducing more variables to this test, the box size has been fixed to runs with 18 Mpc/h. The simulations used to recover the observables are Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW.

The first comparison with the observations explored here is the CIV column density distribution function. At  $z = 4.8$  and  $5.6$ , the CCDFs are compared with observations from D’Odorico et al. (2013) in Figures 5.6 and 5.7, respectively. At  $z = 6.4$ , the simulated values of the CDDF are compared with upper limits from Bosman et al. (2017) in Figure 5.8. In addition, Figure 5.9 shows a comparison of the calculated CIV cosmological mass density at  $4 < z < 8$  for absorbers in the range  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$  and observations by Pettini et al. (2003) and Ryan-Weber et al. (2009) in orange circles, Songaila (2001, 2005) in cyan triangles, Simcoe et al. (2011) in dark green inverted triangles, D’Odorico et al. (2013) in pink squares, Boksenberg & Sargent (2015) in grey diamond, upper limits from Bosman et al. (2017) in purple star and Díaz et al. (in prep) in black pentagons.

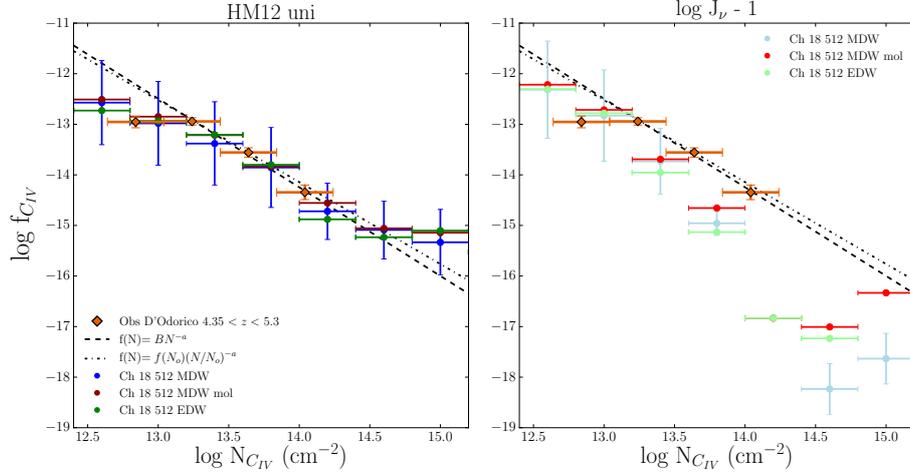


Figure 5.6 CIV column density distribution function at  $z = 4.8$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 10.29 \pm 1.72$  and  $\alpha = 1.75 \pm 0.13$  and the dotted–dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 13.56$  and  $\alpha = 1.62 \pm 0.2$ , from the same observational work. The error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. The left panel shows results with the uniform HM12 and the right panel the results with the test  $\log \mathbf{J}_\nu - 1$ , for the simulations Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW. In all cases, but in particular in simulations without molecular cooling implemented, the number of CIV absorbers in the  $\log \mathbf{J}_\nu - 1$  case is under-represented in the range of column densities considered.

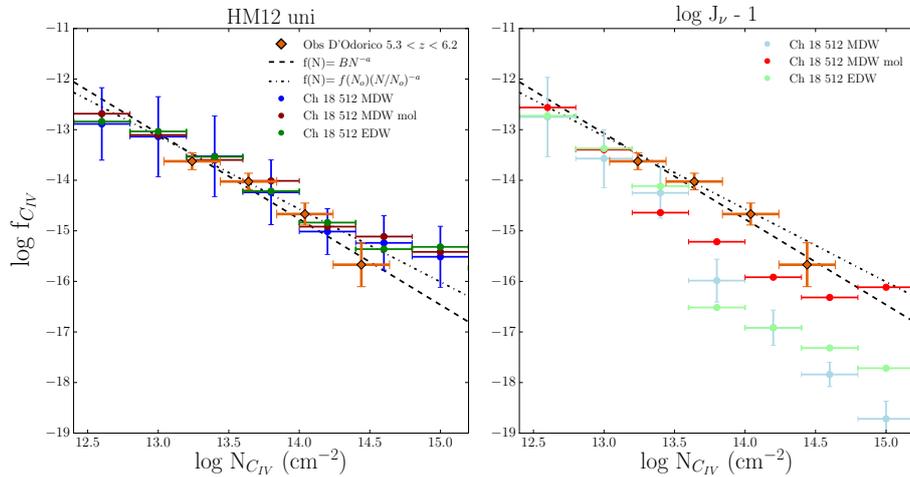


Figure 5.7 CIV column density distribution function at  $z = 5.6$  and comparison with observational data by D’Odorico et al. (2013) in orange diamonds. The black dashed line represents the fitting function  $f(N) = BN^{-\alpha}$  with  $B = 8.96 \pm 3.31$  and  $\alpha = 1.69 \pm 0.24$  and the dotted–dashed line  $f(N) = f(N_0)(N/N_0)^{-\alpha}$  with  $f(N_0) = 14.02$  and  $\alpha = 1.44 \pm 0.3$ , from the same observational work. The blue error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. The left panel shows results with the uniform HM12 and the right panel the results with the test  $\log \mathbf{J}_\nu - 1$ , for the simulations Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW.

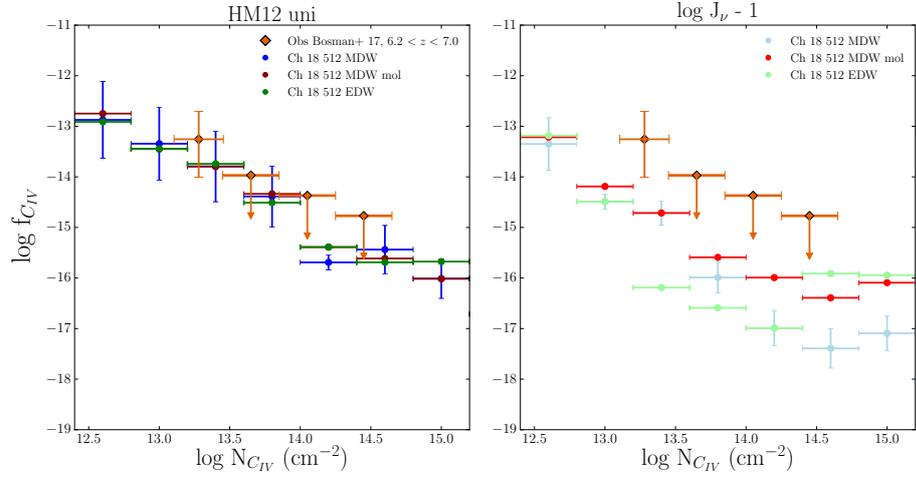


Figure 5.8 CIV column density distribution function at  $z = 6.4$  and comparison with observational data by Bosman et al. (2017) in orange diamonds. The error bars are the Poissonian errors for the reference run and are a good representation of the errors in the other models. The left panel shows results with the uniform HM12 and the right panel the results with the test  $\log \mathbf{J}_\nu - 1$ , both for the simulations Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW.

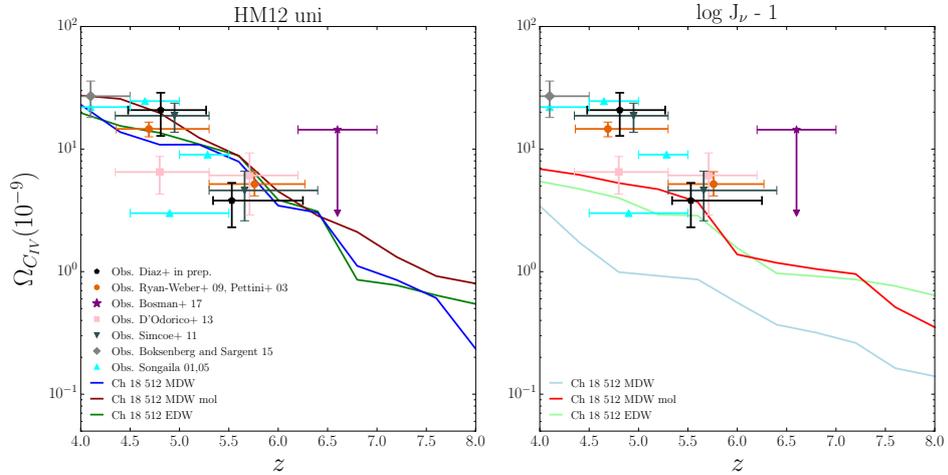


Figure 5.9 CIV cosmological mass density at  $4 < z < 8$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ . Comparison between the simulated data and observations by Pettini et al. (2003) and Ryan-Weber et al. (2009) in orange circles, Songaila (2001, 2005) in cyan triangles, Simcoe et al. (2011) in dark green inverted triangles, D'Odorico et al. (2013) in pink squares, Boksenberg & Sargent (2015) in grey diamond, upper limits from Bosman et al. (2017) in purple star and Díaz et al. (in prep) in black pentagons. Pettini, Ryan-Weber and Díaz measurements are converted to the Planck cosmology, while for the others this recalibration was not possible due to missing details of the precise pathlength probed. On the left panel the results with the uniform HM12 are presented. The right panel shows  $\Omega_{\text{CIV}}$  in the framework of the test  $\log \mathbf{J}_\nu - 1$ . In both cases, the simulations used are Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW. As a consequence of the low number of absorbers in this column density range, when the UVB normalization is varied,  $\Omega_{\text{CIV}}$  is at least an order of magnitude lower than the case with the original UVB.

The left panels in the previous plots have been discussed in Chapter 3, and they are shown here just as a reference to compare the impact of the variation in the normalization of the UVB. The key feature of Figures 5.6, 5.7, 5.8 and 5.9 is that all of them show a notable underproduction of the CIV absorbers at all redshifts when the emissivity is lowered. The calculated CDDFs are always below the observed values, and for high column densities, there is a clear departure from the fitting functions provided by D’Odorico et al. (2013). It is worth noting that Bosman et al. (2017) data are just upper limits for the CIV–CDDF at  $6.2 < z < 7.0$ , nevertheless the computed values in this test are significantly underrepresented.

As a consequence, the cosmological mass density in the right hand panel in Figure 5.9 is at least an order of magnitude below the reference case on the left panel, because the number of absorbers in the range  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$  are underproduced by the simulations in the framework of the modified UVB. The most remarkable difference is visible in the reference run Ch 18 512 MDW, with an order of magnitude shift between the blue curve (on the left) and the light blue one (on the right).

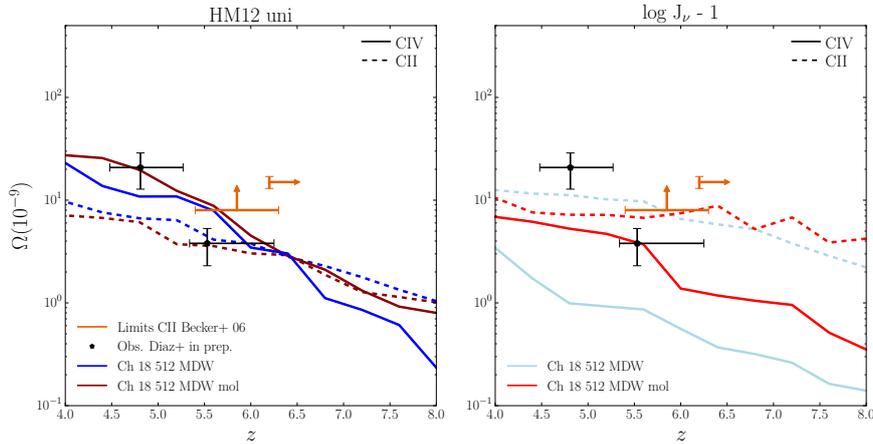


Figure 5.10 Evolution of the CII and CIV cosmological mass density when the normalization of the UVB is varied at 1 Ryd (comparison of molecular cooling content). On the left panel the results with the uniform HM12 are presented. The right panel shows results of the test  $\log J_\nu - 1$ . In both cases, the runs used are Ch 18 512 MDW and Ch 18 512 MDW mol. The solid lines mark the evolution of  $\Omega_{\text{CIV}}$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , and the dashed lines  $\Omega_{\text{CII}}$  in the range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$ . The orange points with errors represent the observational lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006). Although in the case with softer UVB there is not a crossover of CII and CIV (due to the low number of CIV absorbers), the mass density of CII matches the limits from Becker et al. (2006) in both simulations and CIV is well represented in the molecular cooling run.

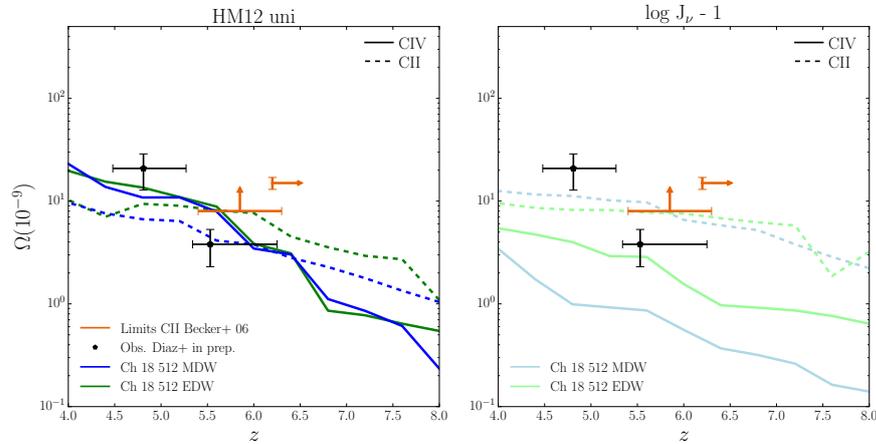


Figure 5.11 Evolution of the CII and CIV cosmological mass density when the normalization of the UVB is varied at 1 Ryd (comparison of feedback prescriptions). On the left panel the results with the uniform HM12 are presented. The right window panel shows results of the test  $\log \mathbf{J}_\nu - 1$ . In both cases, the runs used are Ch 18 512 MDW and Ch 18 512 EDW. The solid lines mark the evolution of  $\Omega_{\text{CIV}}$  for  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$ , and the dashed lines  $\Omega_{\text{CII}}$  in the range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$ . The orange points with errors represent the observational lower limits for  $\Omega_{\text{CII}}$  from Becker et al. (2006). There is not crossover of CII and CIV at any redshift, because of the different orders of magnitude of the mass densities of these ions. In addition, different feedback prescriptions do not seem to give rise to a remarkable distinction in the evolution of CII. Yet, the plot reveals that a softer UVB effectively favours low ionization states as CII, and brings down the gap between the observations from Becker et al. (2006) and the simulated column densities.

Therefore, strong variations of the UVB around 1 Ryd (specifically, normalization changes in the wavelength range where the transition occurs) seem to have a large impact on the number of CIV absorbers at all redshifts. Figures 5.10 and 5.11 draw a comparison of the evolution of CII and CIV in the redshift range  $4 < z < 8$ , with the original HM12 (on the left panel) and the test  $\log \mathbf{J}_\nu - 1$  (on the right). As discussed above, the cosmological mass density of CIV is underrepresented. However, the resulting mass density of CII significantly improves with a softer UVB, indicating that the hypothesis made to perform this test is well-motivated. Effectively, the number of large column density CII absorbers increases and the right panels of Figures 5.10 and 5.11 are now compatible with the limits measured by Becker et al. (2006).

The feedback prescription does not play a major role in the evolution of CII in this redshift range, while the molecular cooling run Ch 18 512 MDW mol shows relatively good agreement with the observational data.

Due to the different orders of magnitude between the calculated mass densities of CII and CIV, the right panels of Figures 5.10 and 5.11 show no crossover of the low and high

ionization states of Carbon. A natural conclusion from this could be that decreasing the intensity of the UV background lead to an improvement in low ionization states at the expense of a poor calculation of CIV absorbers, that are traditionally well constrained by observations.

The column density relationships compare the column density of two/three ions simultaneously and allow us to see a trend (if any) between different ion transitions, in the column density ranges where the observations are made. However, results from Chapter 3 show this is not the best way to probe the statistical distribution of the data and therefore, constrain different models (feedback prescription and/or molecular cooling content). In addition, ionic transitions with very low column densities detected (below the noise threshold limit) are not well represented by the synthetic absorbers in the column density relationships.

The column density relationships were recalculated for the purpose of this test ( $\log J_\nu - 1$ ) and they are shown in Appendix A, comparing in each plot a simulation with the uniform UVB and the test with a lower normalization in the emissivity at 1 Ryd. Again, no remarkable difference is found between the simulation runs Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW.

As pointed out before, CIV is underrepresented, but there is a relative improvement in the column densities of low ionization states. However, it is extremely difficult to draw definitive conclusions from the column density relationships because the number of absorbers depends strongly on the ion, with considerably less CIV synthetic absorbers in the case of  $\log J_\nu - 1$ .

### **An alternative way to count the number of absorbers**

Because of the difficulty in discriminating models on top of a varying UVB, a different approach is proposed in this analysis. Instead of comparing the ion species in column density relationships within each other, each transition is considered separately. The idea is to quantify how well the synthetic absorbers compare with the observed detections at  $z \sim 6$ .

The observational sample consists of 13 CIV absorbers identified by D’Odorico et al. (2013) and 14 OI absorbers reported by Becker et al. (2006). We use this sample to define

the observational column density range in  $\log N$  ( $\text{cm}^{-2}$ ): 12.8-14.3 (CII), 13.1-14.6 (CIV), 12.5-13.7 (SiII), 12.5-13.7 (SiIV) and 13.4-14.9 (OI). The difference between the number of detected absorbers in column **B** and total for CII, SiII and SiIV is due to detections reported as upper limits rather than absolute detections. The sample of synthetic absorbers, on the other hand, spans a range of column density of  $10^{12.5-18.0} \text{ cm}^{-2}$ , hence, each ion can be classified in three categories: **A**, absorbers below the range of the observations, **B**, absorbers in the range defined by the observations and **C**, absorbers above the maximum column density of the detections range. In some cases, e.g. SiII or SiIV, there are observations below the noise threshold of the synthetic absorbers, but they are ignored because the minimum column density assumed in the simulated data is  $10^{12.5} \text{ cm}^{-2}$ . By default, in these cases there are not absorbers in category **A**, as indicated with the dash (-) in Tables 5.1, 5.3 and 5.5.

The information gathered by classifying the absorbers in this way is summarized in Table 5.1. In the table, the first row in each ion shows the number of systems detected in each survey and the maximum number of detections in the quasar spectra. Then, as a reference, the simulation Ch 18 512 MDW with a uniform HM12 UVB and finally, the number of absorbers in each column density range with the test  $\log J_\nu - 1$  for the simulation runs: Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW.

The distribution of the number of absorbers per ion and the comparison with the corresponding observations is shown in Figure 5.12. The histograms have been normalized by the number of absorbers in each range in column density **A**, **B** and **C**, not by the total amount of absorbers, in order to have a fair comparison with the observations in the range **B**. The arrangement in Figure 5.12 shows the ions in each row in the following order: CII (in green), CIV (in blue), SiII (in purple), SiIV (in grey) and OI (in yellow). The observations are represented by the orange hatched histograms. The columns show, from left to right, the different cases of the test performed : Ch 18 512 MDW (HM12 original), as a reference, Ch 18 512 MDW ( $\log J_\nu - 1$ ), Ch 18 512 MDW mol ( $\log J_\nu - 1$ ) and Ch 18 512 EDW ( $\log J_\nu - 1$ ). The overlap between the observations and the synthetic absorbers probes how well each case perform.

Both in Table 5.1 and Figure 5.12, one can see that the total number of CIV absorbers significantly decreases in the test  $\log J_\nu - 1$  with respect to reference case (by a factor of three) and the absorbers in the range of the detections by D’Odorico et al. (2013) are definitely not well represented, explaining the orders of magnitude difference from the

Table 5.1 Number of absorbers at  $z = 6$  when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd compared with the observations, in the range of column density defined by the detections of CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). The first column refers to the ion. In the second column, the observational reference or simulation runs used. Column **A** corresponds to the number of absorbers that are below the column density range of the observations. Column **B**, the number of absorbers in the range defined by the observations and **C**, the absorbers that are above the maximum column density of the detections range. The final column displays the total number of absorbers in the range  $12.5 < \log N(\text{cm}^{-2}) < 18$ .

ion		<b>A</b>	<b>B</b>	<b>C</b>	Total
<b>CII</b>	Obs. Becker et al. (2006)		13		14
	Ch 18 512 MDW (HM12 original)	-	70	102	172
	Ch 18 512 MDW ( $\log J_\nu -1$ )	-	77	114	191
	Ch 18 512 MDW mol ( $\log J_\nu -1$ )	-	75	129	204
	Ch 18 512 EDW ( $\log J_\nu -1$ )	-	65	129	194
<b>CIV</b>	Obs. D’Odorico et al. (2013)		13		13
	Ch 18 512 MDW (HM12 original)	46	68	39	153
	Ch 18 512 MDW ( $\log J_\nu -1$ )	17	3	30	50
	Ch 18 512 MDW mol ( $\log J_\nu -1$ )	20	4	25	49
	Ch 18 512 EDW ( $\log J_\nu -1$ )	21	7	23	51
<b>SiII</b>	Obs. Becker et al. (2006)		13		14
	Ch 18 512 MDW (HM12 original)	-	31	122	153
	Ch 18 512 MDW ( $\log J_\nu -1$ )	-	38	140	178
	Ch 18 512 MDW mol ( $\log J_\nu -1$ )	-	51	168	219
	Ch 18 512 EDW ( $\log J_\nu -1$ )	-	42	147	189
<b>SiIV</b>	Obs. D’Odorico et al. (2013)		10		13
	Ch 18 512 MDW (HM12 original)	-	38	148	186
	Ch 18 512 MDW ( $\log J_\nu -1$ )	-	45	85	130
	Ch 18 512 MDW mol ( $\log J_\nu -1$ )	-	63	84	147
	Ch 18 512 EDW ( $\log J_\nu -1$ )	-	42	80	122
<b>OI</b>	Obs. Becker et al. (2006)		14		14
	Ch 18 512 MDW (HM12 original)	616	21	53	690
	Ch 18 512 MDW ( $\log J_\nu -1$ )	543	29	87	659
	Ch 18 512 MDW mol ( $\log J_\nu -1$ )	552	27	101	680
	Ch 18 512 EDW ( $\log J_\nu -1$ )	599	23	89	711

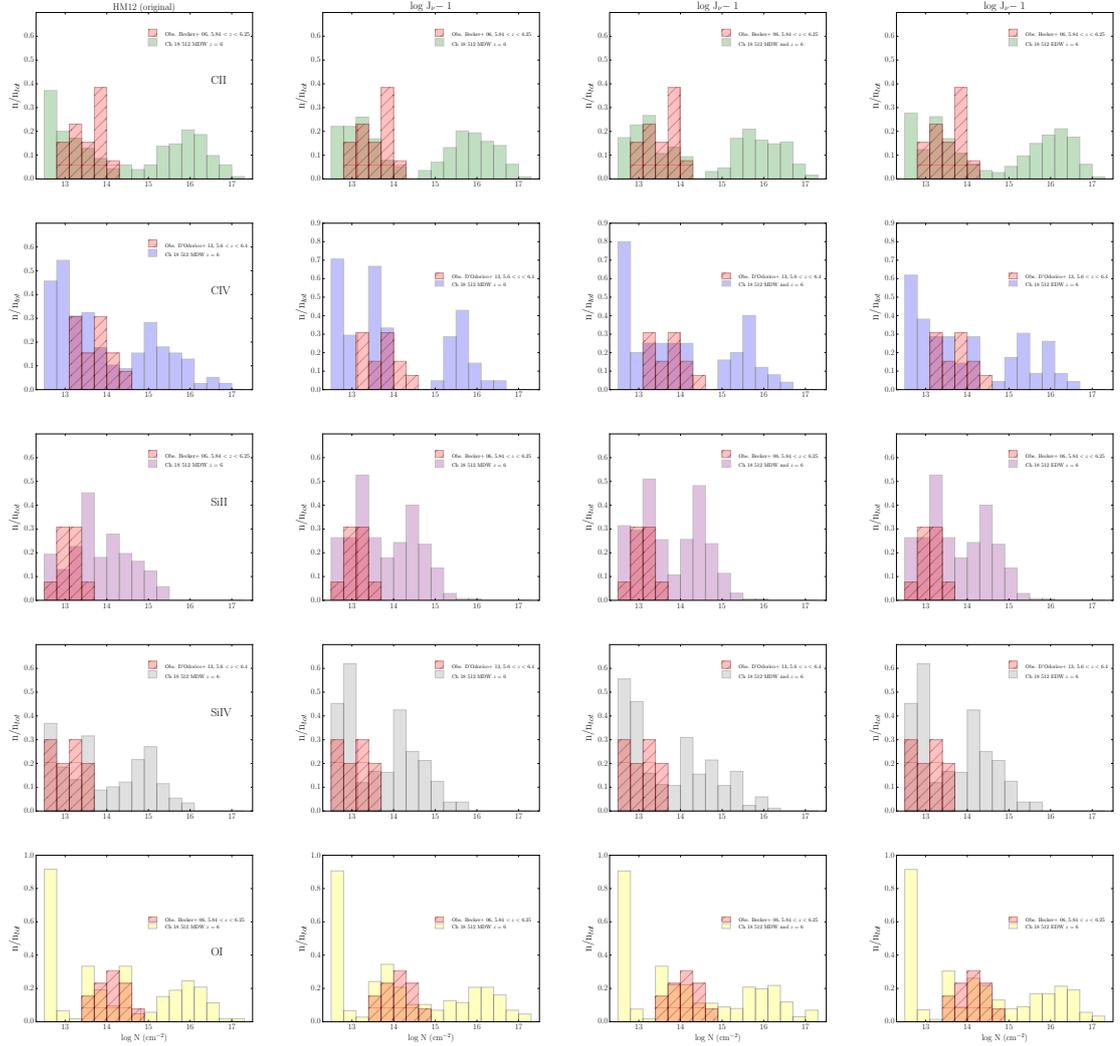


Figure 5.12 Number of absorbers at  $z = 6$  when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd compared with the observations, in the range of column densities defined by the detections of CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). Once the range of the observations is identified (orange hatched histogram), the synthetic absorbers are divided in three ranges: **A**, absorbers below the column density range of the observations; **B**, those which happen to be in the range of the observations and **C**, absorbers above the maximum column density of the detections range. All the synthetic absorbers above the noise threshold are shown ( $12.5 < \log N(\text{cm}^{-2}) < 18$ ). The histograms are normalized by the number of absorbers in the ranges **A**, **B** and **C** -see also Table 5.1- The ions are displayed in rows in the order: CII (green), CIV (blue), SiII (purple), SiIV (grey) and OI (yellow). The columns show from left to right the different cases of the test performed : Ch 18 512 MDW (HM12 original), as a reference, Ch 18 512 MDW ( $\log J_\nu -1$ ), Ch 18 512 MDW mol ( $\log J_\nu -1$ ) and Ch 18 512 EDW ( $\log J_\nu -1$ ). The overlap between the observations and the synthetic absorbers probes the fidelity of the calculation of the number of absorbers in each test in the range of the column densities **B**.

observables described in the previous subsection. On the other hand, SiIV, the other high ionization state in this analysis, shows a moderate improvement in terms of the absorbers in the range of the observations from D’Odorico et al. (2013). This increment is in agreement with findings by Bolton et al. (2011) and it is quite encouraging, because it means that the negative results for CIV should not be generalised to other high ionization states. The number of absorbers of CII, SiII and OI in the range of the observations also rises when the UVB implemented is softer, regardless of the simulation setup. This result is remarkably important because it shows that a moderate reduction in the emissivity of the UVB leads to a better estimation of the column density of low ionization states, and a more compatible incidence rate of the mock spectra with the observations by Becker et al. (2006). The positive result for CII aligns with findings by Finlator et al. (2015). In that work, the authors explore the constraints in the normalization of the UVB and find that CII is well represented with a softer UVB than the HM12 model. On the contrary, our results disagree with conclusions from Finlator et al. (2016), in particular with the idea that SiIV is overproduced by the HM12 UVB.

Finally, it is worth noticing that OI is always overproduced at low column densities by the simulations, regardless of the normalization imposed on HM12. This issue is partially alleviated with the variation of the UVB, and the histograms reveal that the synthetic absorbers incidence rates match the observed ones by Becker et al. (2006). Additional work needs to be done to faithfully reproduce OI, especially when compared with the very large column density systems reported by Becker et al. (2006).

### 5.3.2 Influence of the high energy end in the UVB

Many authors have studied the impact of quasars at the dawn of Reionization and their predominance in the UVB spectrum at  $z \sim 6$ . For instance, Chardin et al. (2017b) proposed an alternative UVB than the HM12 model and confirmed that a large contribution of ionizing photons from quasars can explain the large opacity fluctuations reported by Becker et al. (2015a) and discussed by Davies & Furlanetto (2016).

On the other hand, Finlator et al. (2015) introduced three models of the simulated UVB spectra with galaxies-only, quasars-only and galaxies+quasar contribution. In a subsequent work, Finlator et al. (2016) showed that the quasars-only UVB over-ionizes the gas where metal absorbers live. As a result, the model fails to reproduce the D’Odorico et al. (2013) detections because its spectrum is very hard compared with HM12.

Proposing a model for the UVB with a quasar-only contribution is outside the scope of this thesis. Yet, there is an interesting test that can be made using the uniform HM12 and introducing a very hard contribution from quasars, which has a major impact on the spectrum at high energies. In fact, quasars produce X-ray photons and are much more luminous than galaxies in the UV, but the number density of quasars is significantly lower than that of galaxies at any redshift. In addition, the massive halos that host quasars are quite rare at early times.

Although there is not a physical motivation for the high energy end of the spectrum to be flat (from energy  $\sim 1000$  Ryd) and there is a well defined slope in HM12 as a function of redshift, a very strong variation has been introduced in this test (as shown in Figure 5.13). The purpose of this additional hard UVB is to quantify the variation of the calculated column densities of the ionic species under the influence of this hypothetical spectrum.

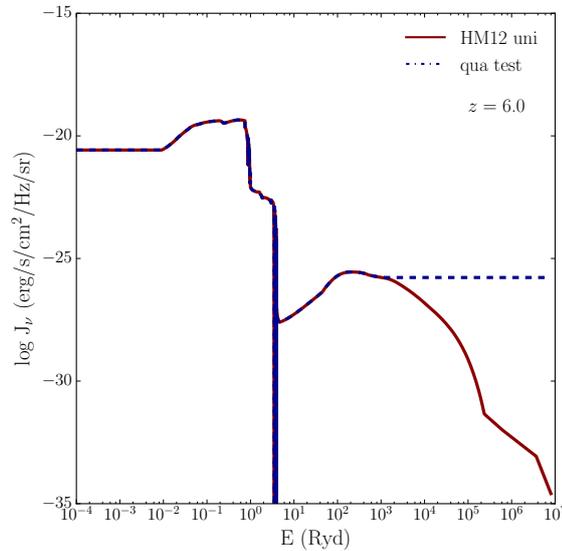


Figure 5.13 UV emissivity vs. energy at  $z = 6$ . The dark red curve shows the uniform HM12 and the blue dashed line a very aggressive contribution from quasars (**qua test**).

This variation of the UVB was performed only at redshift  $z = 6$  with the reference simulation Ch 18 512 MDW. As a sanity check, the column density relationships were calculated and compared with observations from D’Odorico et al. (2013) and Becker et al. (2006). They are presented in Appendix A. However, no significant difference is found in

the relationships. In fact, the distribution of the data resembles the case with the uniform HM12. The first conclusion that can be drawn from the current test is that the variation of the column densities of the ionic species is only appreciable when the emissivity of the UVB varies around the energies/wavelength where the transitions occur.

Keeping that idea in mind, the second observable calculated is the cosmological mass density of CII and CIV, in Table 5.2. The orders of magnitude of both cosmological mass densities with the harder spectrum and the original HM12 are the same, and there is not a remarkable impact on the calculation of the mass densities of these ions.

Additional analysis is done to quantify the number of absorbers with respect to the observations, as shown in Table 5.3 and Figure 5.14. Again, the histograms are normalized with respect to the number of absorbers in each column density range, not the total number of absorbers per ion.

Table 5.2 CII and CIV cosmological mass densities at  $z = 6$  with a modification of the high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. The simulation run used is Ch 18 512 MDW. The mass densities have been calculated with absorbers in the range  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$  for CIV and  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$  for CII.

	$\Omega_{\text{CII}}(z = 6)$ ( $\times 10^{-9}$ )	$\Omega_{\text{CIV}}(z = 6)$ ( $\times 10^{-9}$ )
HM12 uni	3.78	3.45
<b>qua test</b>	2.89	3.66

Table 5.3 Number of absorbers at  $z = 6$  with a modification of the high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. The synthetic absorptions are compared with the observations, in the range of column density defined by the detections of CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). The simulation run used is Ch 18 512 MDW. The first column refers to the ion. In the second column, the model considered in each case. Column **A** corresponds to the number of absorbers that are below the column density range of the observations. Column **B**, the number of absorbers in the column density range defined by the observations and in **C**, the absorbers that are above the maximum column density of the detections range. The final column displays the total number of absorbers in the range of  $12.5 < \log N(\text{cm}^{-2}) < 18$ .

ion		<b>A</b>	<b>B</b>	<b>C</b>	Total
<b>CII</b>	Obs. Becker et al. (2006)		13		14
	HM12 uni	-	70	102	172
	<b>qua test</b>	-	68	89	157
<b>CIV</b>	Obs. D’Odorico et al. (2013)		13		13
	HM12 uni	46	68	39	153
	<b>qua test</b>	46	73	30	149
<b>SiII</b>	Obs. Becker et al. (2006)		13		14
	HM12 uni	-	31	122	153
	<b>qua test</b>	-	22	113	135
<b>SiIV</b>	Obs. D’Odorico et al. (2013)		10		13
	HM12 uni	-	38	148	186
	<b>qua test</b>	-	31	133	164
<b>OI</b>	Obs. Becker et al. (2006)		14		14
	HM12 uni	616	21	53	690
	<b>qua test</b>	603	12	63	678

Table 5.3 and Figure 5.14 reveal little variation of the number of absorbers with the uniform HM12 UVB and the strong contribution from quasars (**qua test**). This has some implications: the ionic species are mostly unaffected by variations of the UVB that occur at energy scales outside the order of a few Ryd, unless the absorptions happen in the vicinity of a quasar itself, in which case, this test is not reliable due to the absence of a proper treatment of the AGN driven ionization. The variations with the quasar contribution can be added in a more realistic way to the spectrum, e.g. applying a radiative transfer treatment in the proximity zone of the quasar. For example, the simulated quasar-only

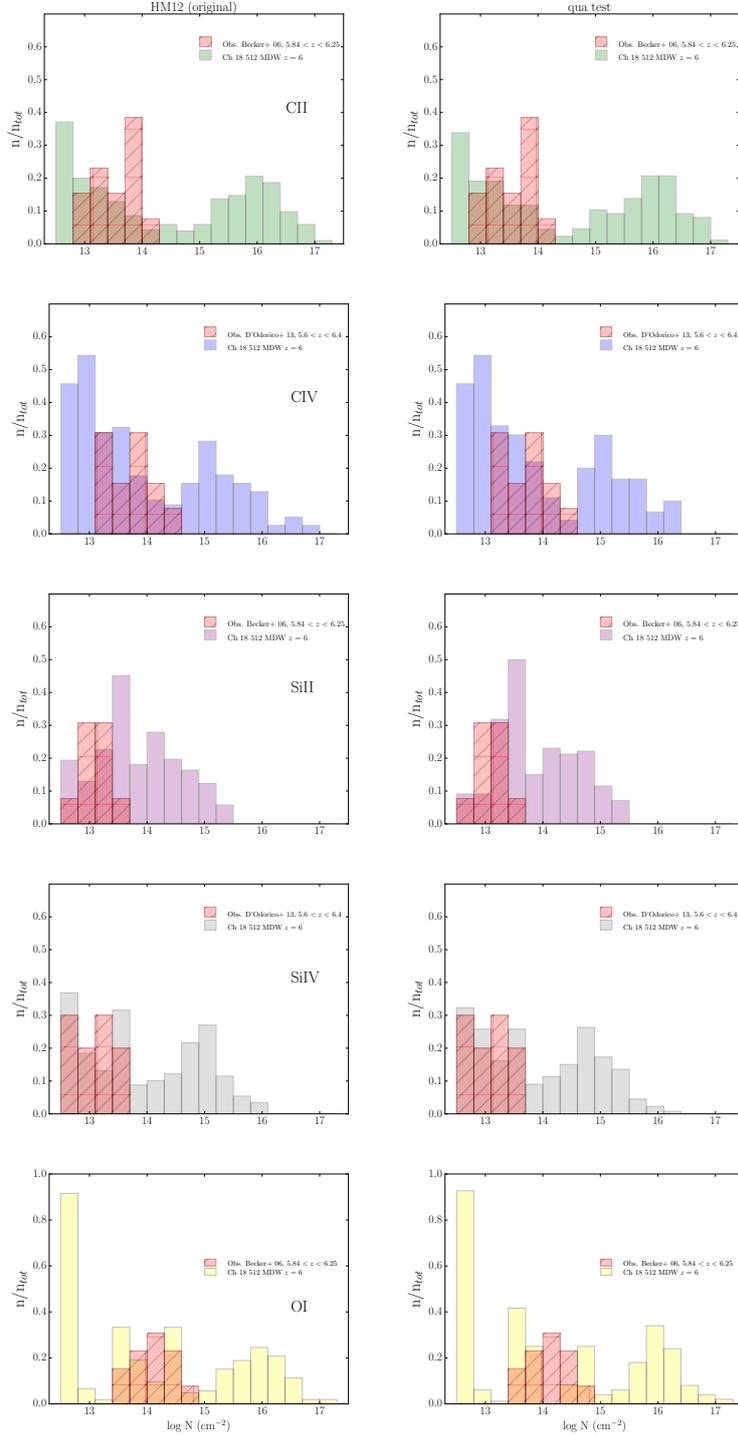


Figure 5.14 Number of absorbers at  $z = 6$  with a modification of the high energy end of the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift. The observational detections define a range for each ion: CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). Once the range of the observations is identified (orange hatched histogram), the synthetic absorbers are divided in three ranges: **A**, absorbers below the column density range of the observations; **B**, those which happen to be in the range of the observations and **C**, absorbers above the maximum column density of the detections range. All the synthetic absorbers above the noise threshold above ( $12.5 < \log N(\text{cm}^{-2}) < 18$ ). The histograms are normalized by the number of absorbers in the ranges mentioned above -see also Table 5.3-. The ions are displayed in rows in the order: CII (green), CIV (blue), SiII (purple), SiIV (grey) and OI (yellow). The left panels show the reference case HM12 uni and the right ones the test with a strong contribution from quasars at  $z = 6$ , **qua test**. The overlap between the observations and the synthetic absorbers probes the fidelity of the calculation of the number of absorbers in each test in the range of the column densities **B**. The simulation run used is Ch 18 512 MDW.

spectrum from Finlator et al. (2016) differs from HM12 in the range of wavelengths where the ionic states occur (100 to 1000 Å). Further modelling could be done in this aspect.

## 5.4 Variation of the metallicity in the photoionization modelling

As mentioned in the motivation of this thesis and Chapter 2, it is common to assume the metallicity to solar at  $z = 6$ , which is not necessary realistic, since the enrichment of the Universe is still in an early stage. Findings from Chapters 3 and 4 reveal that simultaneous to the progression of Reionization at  $z \geq 6$ , the metal pollution of the IGM is taking place, driven by supernovae explosions and the evolution of stars in galaxies. Thereby, it is natural to imagine that the metallicity in the IGM and CGM at early times is significantly lower than the values measured at later times.

Due to the lack of techniques to directly quantify the global metallicity at high redshift, astronomers' best assumption is to impose a solar metallicity at high redshift in theoretical photoionization models. Although the choice is justified and gives a reasonable upper limit for the metallicity content and the ionization states of each chemical element present in the intergalactic medium, it is not the best description of the increasing metallicity of the IGM at early times.

Work by Schaye et al. (2003) with hydrodynamical simulations measured the distribution of Carbon in the IGM as a function of redshift  $z$  and overdensity  $\delta$ , using a HM01 UVB and HI absorptions to derive a median metallicity of Carbon  $[C/H] = -3.47_{-0.06}^{+0.07} + 0.08_{-0.10}^{+0.09}(z - 3) + 0.65_{-0.14}^{+0.10}(\log\delta - 0.5)$  for  $\log \delta = -0.5 - 1.8$  and  $z = 1.8 - 4.1$ . Additionally, they found that the abundance of this metal is spatially inhomogeneous in the IGM at  $z \sim 4$ .

A different study with DLAs at  $z \leq 4.7$  by Rafelski et al. (2014) shows that metallicity in these structures decreases dramatically to a level of  $\log \langle Z/Z_{\odot} \rangle = -2.03_{-0.11}^{+0.09}$  compared with DLA observations at lower redshifts.

Both works show that the metallicity at  $z \sim 4$  in different environments is indeed lower than solar. Motivated by these facts, we run an additional test with CLOUDY at  $z = 6$  where we decrease the metallicity content in the photoionization models by three orders of magnitude below solar. Since the set of simulations is fixed through the course of this work, it is not possible to change the metallicity output resulting from our assumed

stellar yields and IMF. An alternative method is to vary the metallicity below solar in the CLOUDY tables, by changing the input metallicity with respect to the values described in Asplund et al. (2009). Exploring metallicity in the CLOUDY tables is another important test of the underlying models. The new results are compared with the reference case  $Z = Z_{\odot}$  using the fiducial simulation Ch 18 512 MDW and the uniform HM12 radiation field.

Table 5.4 presents the cosmological mass densities of CII and CIV at  $z = 6$ . The mass densities have been calculated with absorbers in the range  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$  for CII and  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$  for CIV. The reference case with solar metallicity  $Z = Z_{\odot}$  is compared with three different metallicities imposed in the photoionization models:  $Z = 0.1Z_{\odot}$ ,  $Z = 0.01Z_{\odot}$  and  $Z = 0.001Z_{\odot}$ .

The cosmological mass densities of CII and CIV have the same order of magnitude in the reference case and when lower metallicities in CLOUDY are used. The trends displayed by the mass densities are counter-intuitive. One would expect that a reduction in the metallicity would lead to a lower mass density of both ions, however, the results from the table show the opposite behaviour for CIV, even when the metallicity is three orders of magnitude below solar.

Table 5.4 CII and CIV cosmological mass densities at  $z = 6$  while varying the metallicity in the CLOUDY modelling. The simulation run used is Ch 18 512 MDW. The mass densities have been calculated with absorbers in the range  $13.8 < \log N_{\text{CIV}}(\text{cm}^{-2}) < 15.0$  for CIV and  $13.0 < \log N_{\text{CII}}(\text{cm}^{-2}) < 15.0$  for CII.

$Z(Z_{\odot})$	$\Omega_{\text{CII}}(z = 6)$ ( $\times 10^{-9}$ )	$\Omega_{\text{CIV}}(z = 6)$ ( $\times 10^{-9}$ )
1	3.78	3.45
0.1	3.50	3.87
0.01	3.42	4.18
0.001	2.94	4.11

Next, the number of absorbers are counted in the same way explained in the previous tests. The results are presented numerically in Table 5.5 and quantitatively in Figure 5.15. In both cases, the absorbers have been split in three categories: **A**, absorbers below the column density range of the observations; **B**, absorbers in the column density range defined by the observations; and **C**, absorbers above the column density range of the detections.

In Figure 5.15 the ions are displayed in the order: CII (green), CIV (blue), SiII (purple), SiIV (grey) and OI (yellow). The columns in the arrangement show (from left to right): the reference case with  $Z = Z_{\odot}$ ,  $Z = 0.1Z_{\odot}$ ,  $Z = 0.01Z_{\odot}$  and  $Z = 0.001Z_{\odot}$  in the CLOUDY tables run with the uniform HM12 UVB.

Besides some minor variations in the number of absorbers of each ion, the test produces a moderate scatter around the number of absorbers found in the photoionization modelling reference case (with solar metallicity).

The conclusion in this case is solid: variations of the metallicity at high redshift should be implemented self-consistently in the simulations, not in the photoionization model in post-process. Results from Tables 5.4 and 5.5 and Figure 5.15 show unimportant differences between solar metallicity and the most severe case  $Z = 0.001Z_{\odot}$  (three orders of magnitude below solar). Such extreme variation should be reflected in the mass density of the ionic species, as well as in the number of absorbers found in the range of the observations.

In order to investigate why these variations have so little effect in our results at  $z = 6$ , the ionization fractions of CII and CIV are plotted as a function of the gas overdensity in Figure 5.16 and the gas temperature in Figure 5.17. In the first case, two temperatures of the gas are considered,  $T = 10^{4.6}$  K and  $T = 10^{5.0}$  K, whereas in the second gas at overdensity  $\Delta \sim 10^{3.6}$  is presented.

Figure 5.16 shows a trend for the ionization fractions of CII and CIV with no departure from the reference case up to  $\log \Delta \sim 10^{1-2}$ , depending on the selected temperature of the gas. This gas overdensity is typical of the transition from the IGM to the CGM. The absence of differences in this regimes explains why the number of absorbers is quite similar in the four cases explored. The range of column densities used to calculate the cosmological mass density of these ions is limited and defined by observations in the IGM/CGM, therefore, it is natural not to obtain significant differences in the number of absorbers found at these intermediate gas densities. Appreciable differences are present at high gas overdensity, but our pipeline is currently not tracing gas in this region for statistical reasons, since we produce random sightlines inside the box. If we follow the procedure explored in section 3.7.1 of Chapter 3, where we deliberately trace gas around galaxies with a given impact parameter, we should retrieve a considerably larger number of high column density absorbers. However, this method would bias our results and mean a departure from the

Table 5.5 Number of absorbers obtained by varying the metallicity in the CLOUDY modelling compared with the observations, in the range of column density defined by the detections of CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). The first column refers to the ion. In the second column, the metallicity content applied in the CLOUDY tables is shown. Column **A** corresponds to the number of absorbers that are below the column density range of the observations. Column **B**, the number of absorbers in the range defined by the observations and in **C**, the absorbers that are above the maximum column density of the detections range. The final column displays the total number of absorbers in the range of  $12.5 < \log N(\text{cm}^{-2}) < 18$ .

ion	$Z(Z_{\odot})$	<b>A</b>	<b>B</b>	<b>C</b>	Total
<b>CII</b>	Obs. Becker et al. (2006)	-	13		14
	1	-	70	102	172
	0.1	-	70	89	159
	0.01	-	69	90	159
	0.001	-	69	89	158
<b>CIV</b>	Obs. D’Odorico et al. (2013)		13		13
	1	46	68	39	153
	0.1	45	70	34	149
	0.01	45	72	32	149
	0.001	46	71	32	149
<b>SiII</b>	Obs. Becker et al. (2006)	-	13		14
	1	-	31	122	153
	0.1	-	23	113	136
	0.01	-	21	111	132
	0.001	-	23	110	133
<b>SiIV</b>	Obs. D’Odorico et al. (2013)		10		13
	1	-	38	148	186
	0.1	-	35	132	167
	0.01	-	34	130	164
	0.001	-	34	129	163
<b>OI</b>	Obs. Becker et al. (2006)		14		14
	1	616	21	53	690
	0.1	601	12	53	662
	0.01	601	12	49	662
	0.001	605	12	48	665

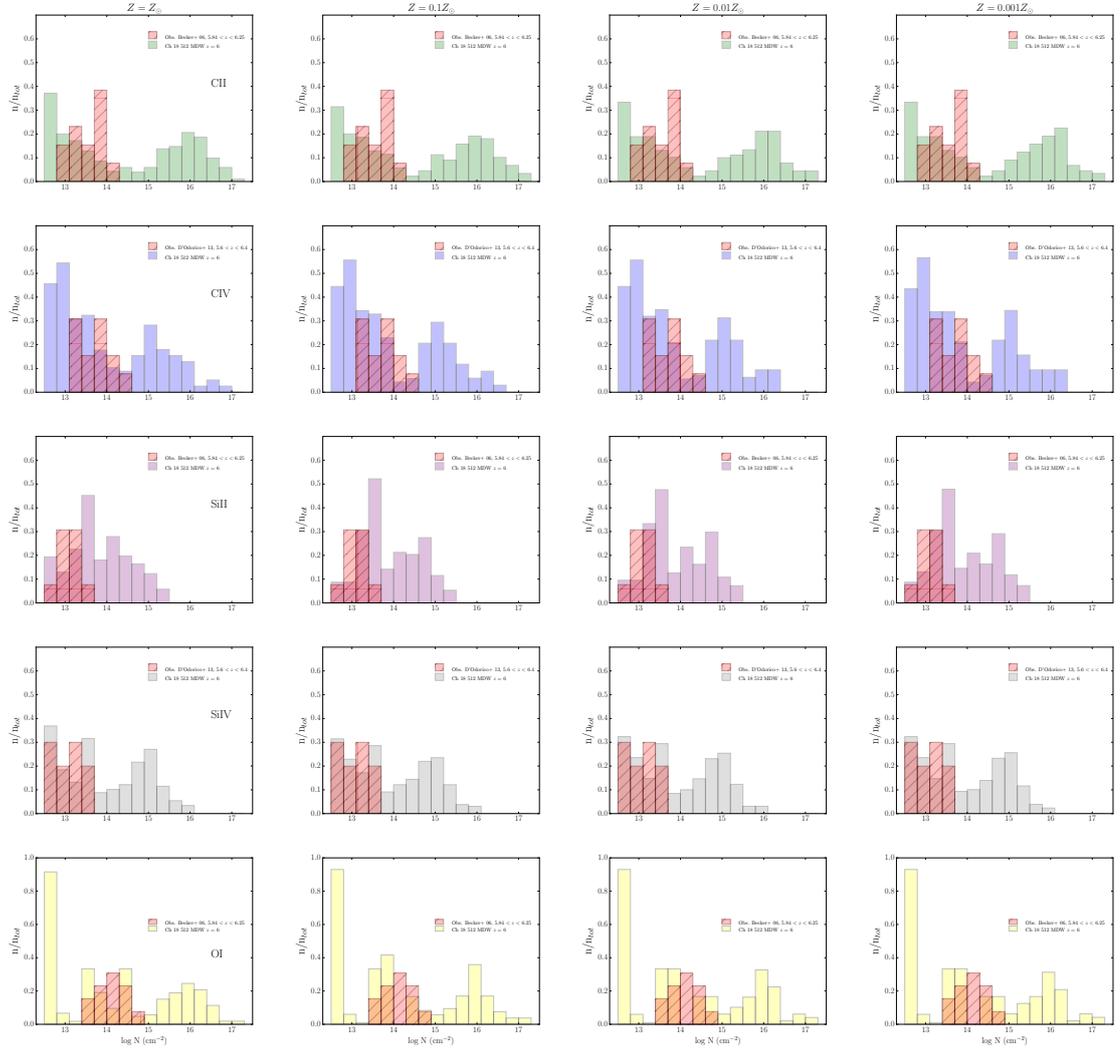


Figure 5.15 Number of absorbers obtained by varying the metallicity in the CLOUDY modelling compared with the observations, in the range of column densities defined by the detections of CIV and SiIV (D’Odorico et al., 2013) and CII, SiII and OI (Becker et al., 2006). Once the range of the observations is identified (orange hatched histogram), the synthetic absorbers are divided in three ranges: **A**, absorbers below the column density range of the observations; **B**, those which happen to be in the range of the observations and **C**, absorbers above the maximum column density of the detections range. All the synthetic absorbers above the noise threshold are shown ( $12.5 < \log N(\text{cm}^{-2}) < 18$ ). The histograms are normalized by the number of absorbers in the ranges mentioned above -see also Table 5.5-. The ions are displayed in rows in the order: CII (green), CIV (blue), SiII (purple), SiIV (grey) and OI (yellow). From left to right, the columns present: the reference case with  $Z = Z_{\odot}$ , the second one  $Z = 0.1Z_{\odot}$ , the third one  $Z = 0.01Z_{\odot}$  and the further right the case with  $Z = 0.001Z_{\odot}$  in the CLOUDY tables. The overlap between the observations and the synthetic absorbers probes the fidelity of the calculation of the latter in each test. The simulation run used is Ch 18 512 MDW.

observational procedure.

Similar information can be inferred from Figure 5.17. CII and CIV ionization fractions have a flat trend with temperature, regardless of the metallicity imposed in the CLOUDY modelling up to  $T \sim 10^{4.6}$  K. From this temperature, both fractions decrease, except for a peak in CIV at  $T \sim 10^{5.0}$  K, where this ion is collisionally ionized. As discussed above, a modified metallicity content in the photoionization modelling causes an important variation in the number of absorbers at very high densities and/or in shock heated gas at  $T \geq 10^{5.0}$  K. As mentioned above, these regions would be traced by our pipeline if we consider lines of sight with a given impact parameter from the centre of galaxies, but we do not do so because the observations of intervening absorbers imprinted on quasar spectra are obtained randomly on the sky.

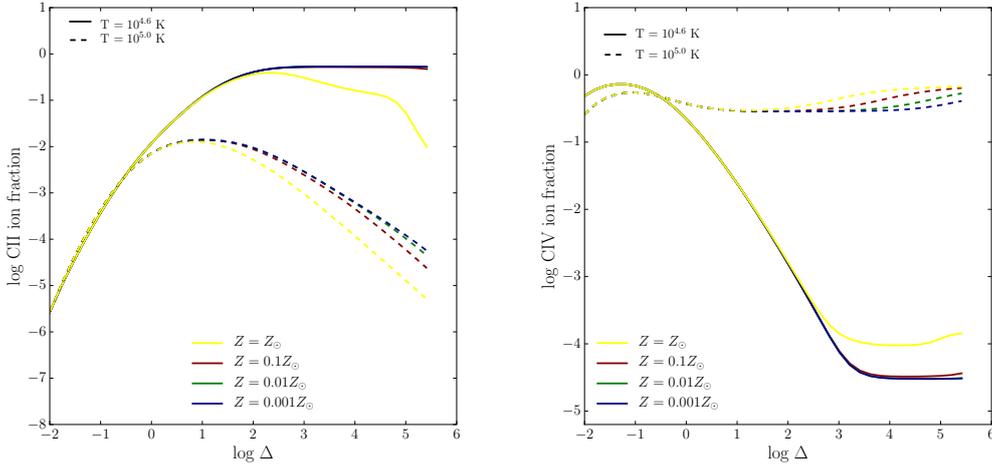


Figure 5.16 Ionization fractions of CII (left panel) and CIV (right panel) as a function of the gas overdensity. Results from CLOUDY modelling at  $z = 6$  comparing gas temperatures of  $T = 10^{4.6}$  K (solid lines) and  $T = 10^{5.0}$  K (dashed lines). Four cases of metallicity have been imposed in the photoionization models:  $Z = Z_{\odot}$  (yellow; reference case),  $Z = 0.1Z_{\odot}$  (dark red),  $Z = 0.01Z_{\odot}$  (green) and  $Z = 0.001Z_{\odot}$  (dark blue). Both ions show a moderate departure from the fiducial case at densities above the one defined for the IGM, the region of interest in this work. A non-negligible difference between the solar and lower cases of metallicity is seen at very high-densities. This explains why the number of absorbers detected at column densities corresponding to the IGM only slightly changes in the histograms discussed above.

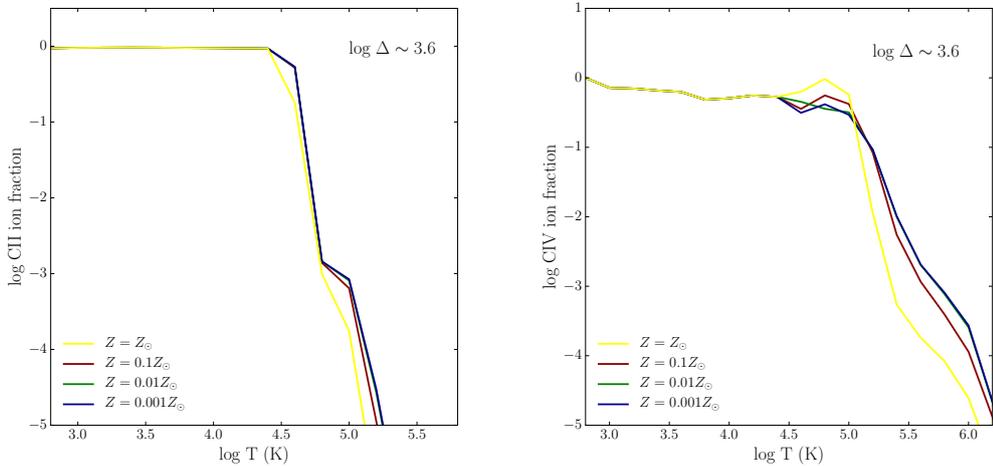


Figure 5.17 Ionization fractions of CII (left panel) and CIV (right panel) as a function of the gas temperature. Results from CLOUDY modelling at  $z = 6$  for a gas overdensity of  $\Delta \sim 10^{3.6}$ . Four cases of metallicity have been imposed in the photoionization models:  $Z = Z_{\odot}$  (yellow; reference case),  $Z = 0.1Z_{\odot}$  (dark red),  $Z = 0.01Z_{\odot}$  (green) and  $Z = 0.001Z_{\odot}$  (dark blue). CII and CIV show similar fractions independently of the temperature and the metallicity imposed in CLOUDY up to  $T \sim 10^{4.6}$  K, where the fractions decrease in a similar fashion. At  $T \sim 10^{5.0}$  K, CIV is collisionally ionized and there is a peak of the fraction around this temperature. Again, different metallicity contents in the photoionization models do not show an appreciable difference at this gas overdensity.

This test is then considered successful as a sanity check for our models and the pipeline developed for this thesis, but it does not allow us to make conclusions in regards of the metallicity content at high redshift.

A possible way to overcome this issue would be to run a new set of simulations with different stellar yields and IMFs, and study the effect of a self-consistent variation of the metallicity in the analysis of the absorption systems.

## 5.5 Perspectives and conclusions

This chapter contains a compilation of different variations to the uniform HM12 ionizing field, that is extensively used in the literature and was assumed in the photoionization models to obtained the results presented in Chapters 3 and 4. However, the large uncertainties on the UVB, especially at high redshift, require better constraints for several probes, and of course, other improved UVB models.

The conclusion that results from the analysis of CIII is that a complete overview of the evolution of the ionization state IGM requires to follow, if not all, at least more ion-

izing states of Carbon. The observational detections of CII and CIV are very few and limited to a short range of column densities, but there is more Carbon in other states in the IGM/CGM. Such states are still undetected or their observational window at high redshift is not available with quasar absorption spectra techniques.

Once the photoionization models are developed and constrained with the observations in hand, further predictions can be done, in particular to quantify how much Carbon is still neutral at  $z \sim 6$  or how much CIV is collisionally ionized. The latter case has been extensively discussed in non-equilibrium by Oppenheimer & Schaye (2013), including photoionization models for Carbon, Silicon and Oxygen, among other metals included in their analysis.

Work from Chen et al. (2017) at  $z \sim 0$  with CLOUDY photoionization models allows the authors to quantify the difference in the inferred CGM metallicity with different UVB spectral slopes. They use observations of ionic species and HI to constrain the fraction of each ionic species as a function of H density. Although a complete test in this direction is not possible at high redshift due to the small sample of detections, future work could be done following the same philosophy.

In a similar vein, observational detections of ionic species could be used to constrain the UVB spectral normalization in the wavelength range where these states occur (100-1000 Å). A careful fine-tuning is required, but most of the steps are clear: from the **quatest**, we learnt that variations in the spectrum have to arise in the wavelength range of the ionic transitions. In addition, a reduction in the normalization of the UV emissivity improves the number of absorbers of CII, SiII, SiIV and OI. A less drastic decrease in the emissivity of the UVB could lead to better results for CIV, keeping the positive findings achieved with the low ionization states and SiIV.

Besides, a more sophisticated model to account for the quasar contribution can be built, taking into account variations at high energies in the UV spectrum, as well as a non-negligible modification in the wavelength range  $100 < \lambda(\text{Å}) < 1000$ . The *toy* model discussed in this Chapter lacks of a variation in the wavelength region where the states take place.

On the other hand, one can speculate about the metallicity content of the IGM at high redshift, and the final test of this Chapter attempted to give some hints in that direction. Nevertheless, variations of the metallicity applied to the photoionization modelling showed

very mild differences in the calculation of the column densities of the ions, due to the fact that our simulations have their own prescribed stellar yields. Furthermore, the procedure followed to produce the lines of sight inside the box has a very low likelihood to trace high density regions, where the high column densities of the low ionization states would be found. The statistical distribution of the synthetic absorbers would change if we trace lines of sight at a given distance from the centre of the galaxies, yet, we prefer to follow the observational pipeline to assure the compatibility of our sample to the observational detections.

A different approach to tackle this issue could be to run a new set of simulations with different stellar yields and IMFs and repeat the methodology followed in previous Chapters to recover the observables for HI and metal ions.

An important ingredient missing in this Chapter are calculations of Hydrogen statistics. HI observables have been ignored for two reasons: 1) at high redshift a uniform UVB does not provide a good representation of the rapid evolution of HI during Reionization; 2) the self-shielding prescription from Rahmati et al. (2013) is calibrated with the photoionization rates of HM12. The analysis presented in Chapter 3 could be repeated with a brand new model for HI self-shielding in the IGM. Chardin et al. (2017a) recently proposed a self-shielding prescription with best-fit parameters up to  $z = 10$ . Our set of simulations could be tested at all redshifts and results from Chapter 3 could be compared for consistency.

Finally, in this chapter we introduced histograms that compared single synthetic absorbers directly with observations. The method provides a straightforward comparison of the incidence rate of simulated absorbers in the range where detections are available, but also discriminates different absorptions below/above the observational column density ranges. This allows us to predict absorptions occurring in the IGM and CGM, currently out of the observational window probed with quasar spectra techniques.

It is worth noting that the inclusion of this method is an alternative to the column density comparisons shown in Chapter 3 and Appendix A, and also studied by Oppenheimer et al. (2009), Becker et al. (2011) and Finlator et al. (2016), among others.



# 6

## Conclusions

This thesis studies the properties of gas in the redshift range of  $4 < z < 8$ , while the Epoch of Reionization is progressing and concluding. We use a set of high resolution cosmological hydrodynamical simulations based on a customized version of the GADGET3 code, called P-GADGET3 (XXL). This work is an extension at high redshift of the ANGUS project (Tescari et al., 2014; Maio & Tescari, 2015; Katsianis et al., 2015, 2016, 2017). The numerical runs satisfactorily describe the global star formation rate (SFR) history in the Universe and the chemical enrichment of the IGM from stars and supernovae. SFR and chemical enrichment are regulated by different feedback prescriptions: energy (EDW) and momentum driven winds (MDW).

In addition, we study the effect of molecular and low-temperature metal cooling, that we found is quite important to achieve an accurate prediction of the HI observables in gas at high density.

The aim of this work is to reproduce the physical and chemical state of the baryonic gas in the IGM at high redshift, in regions ionized by a uniform HM12 UV background introduced in post-process. The radiation field accounts for the contribution of ionizing photons produced by galaxies and quasars. Ultimately, the thesis goal is to provide an accurate description of the evolution of observed ionic transitions that take place in the IGM/CGM: CIV, CII, SiII, SiIV and OI, as well as some observables associated with HI up to  $z \leq 6$ . All the theoretical predictions derived from our simulations have been compared with available observations and the vast majority of them show a remarkable agreement with the observational archive and other simulated results.

Different numerical works that can be compared with the content of this thesis include Oppenheimer & Davé (2006); Oppenheimer et al. (2009); Tescari et al. (2011); Cen & Chis-

ari (2011); Finlator et al. (2013); Pallottini et al. (2014); Keating et al. (2014); Finlator et al. (2015); Rahmati et al. (2016); Keating et al. (2016). In these studies, metal absorption lines have been investigated taking into account different feedback prescriptions, photoionization modelling and variations in the UV ionizing background at high redshift. The methods employed by each of these works, as well as the type of hydrodynamical simulations, show to have their advantages and this thesis builds on their findings and merits.

HI calculations up to  $z = 6$  build on the work and analysis of on DLAs carried out by Tescari et al. (2009); Nagamine et al. (2004); Pontzen et al. (2008); Barnes & Haehnelt (2009); Bird et al. (2014); Rahmati et al. (2015); Maio & Tescari (2015).

The thesis has two main components: in the first part, the simulations are tested against observations of HI and metal ions at high redshift with a uniform HM12 UVB, as shown in Chapters 3 and 4. The second part was devoted to explore variations in the assumed UVB and the metallicity content in CLOUDY photoionization modelling. The results of this second part are discussed in Chapter 5.

In summary, 1000 lines of sight are randomly created inside the cosmological box at redshifts  $4 < z < 8$  (with  $\Delta z = 0.4$ ), to recreate the observations of high redshift quasars and recover synthetic spectra for each ion along each sightline. The pipeline derived for this purpose relies on the physical conditions of the gas (reproduced with the hydrodynamical simulations), a uniform ionizing UV background HM12, a CLOUDY photoionization model and a HI self-shielding prescription by Rahmati et al. (2013). Once the synthetic spectra are computed, they are convolved with instrumental noise, following the observational procedure. Furthermore, individual absorption features in the synthetic spectra are fitted with the VPFIT code. In this way, the pipeline closely mimics the methodology used by observers to recover the column densities of the ions. Finally, a noise threshold has been imposed to the simulated absorbers to distinguish real absorption features from pure instrumental noise, introduced to make the synthetic spectra comparable with the observations, at a column density of  $10^{12.5} \text{ cm}^{-2}$ . The minimal column density fitted by VPFIT is  $10^8 \text{ cm}^{-2}$ , however this order of column density in the simulated spectra is just continuum noise.

The triply ionized state of Carbon, CIV, receives special attention in this work. Due to the large observational archive of data available for this ion, several observables are

compared with results from our simulations. In the framework of the uniform HM12 UV background, the observed CIV column density distribution function (CDDF) at  $z = 4.8$  and  $5.6$  (D’Odorico et al., 2013) is used to constrain our models. The calculated CDDF at  $z = 6.4$  is compatible with upper limits from Bosman et al. (2017). Additionally, when the cosmological mass density of CIV,  $\Omega_{\text{CIV}}$ , is computed, there is a remarkable agreement between the observations at redshifts 4 to 7 and the simulated results. There is a decrease in  $\Omega_{\text{CIV}}$  when moving from  $z = 4$  to  $z = 8$ , caused by the change in the ionization state of the gas driven by the UVB and the drop in the metallicity content in the IGM predicted by the models (see Chapter 3, section 3.6.4).

The excellent agreement of the simulated  $\Omega_{\text{CIV}}$  with the observations is in part due to restricting our calculation to the observed range of column densities in CIV. Using the range probed by D’Odorico et al. (2013),  $13.8 \leq \log N_{\text{CIV}} (\text{cm}^{-2}) \leq 15$ , the calculated mass density of CIV is compatible with a range of observations: Pettini et al. (2003) and Ryan-Weber et al. (2009), Songaila (2001, 2005), Simcoe et al. (2011), D’Odorico et al. (2013), Boksenberg & Sargent (2015), Bosman et al. (2017) and Díaz et al. (in prep). If synthetic absorbers with column density  $15 < \log N_{\text{CIV}} (\text{cm}^{-2}) < 16$  are included, the calculated CIV cosmological mass density increases by more than an order of magnitude. This prediction from our model can be tested with future observations that detect rarer high column density systems.

In addition, an appreciable difference in the number of high column density absorbers would be detected if we trace high gas overdensity, but our pipeline is currently not tracing gas in this region for statistical reasons, since we produce random sightlines inside the box. If we follow the procedure explored in section 3.7.1 of Chapter 3, where we deliberately trace gas around galaxies with a given impact parameter, we should retrieve a considerably larger number of high column density absorbers. However, this method would bias our results and mean a departure from the observational procedure.

We also find that the interplay between low and high ionization states of Carbon, CII and CIV (respectively), are well represented by the simulations. When comparing the mass density of these ions, there is a crossover in  $\Omega_{\text{CIV}}$  and  $\Omega_{\text{CII}}$  at  $z \sim 6\text{--}6.5$ , which indicates that the IGM is transitioning to a more neutral state at high redshift, and therefore low ionization states, as CII, are more predominant at early times. Although the introduction of the UVB in post-process is quite restrictive in the conclusions about the progress of Reionization, we can infer from this result that at  $z \sim 6$  the EoR is reaching its end. This

finding is aligned with independent observations that date the tail of Reionization around this redshift.

Interestingly, results for CIV in the presence of the uniform HM12 background show that, differently than works from Finlator et al. (2015) and Keating et al. (2016), our simulations produce CIV absorbers in the IGM with large column densities. This is mostly due to a combination of efficient galactic winds and a post-processing pipeline that mimics observational methods and, in particular, accounts for individual features to calculate the column densities from Voigt profile fits to the absorption lines. However, when the UV emissivity of HM12 is reduced by one dex, the number of synthetic absorbers significantly decreases and CIV is underrepresented by the simulations, as well as other CIV observables, as discussed in Chapter 5.

Another remarkable prediction related to CIV absorbers from our models is the theoretical confirmation of the observed LAE – CIV absorption pair detected by Díaz et al. (2015). We studied the physical processes that produced the metal enrichment of the IGM at  $z \sim 5.6$  in Chapter 4, and explored two scenarios to explain the enrichment of the region where Diaz’s absorber is observed. In the first one, an outflow wind expelled by the observed LAE seems unlikely, due to the distance of  $1384 \text{ ckpc}/h$  to the absorption. In the context of the feedback models implemented in this thesis, the most powerful outflow would have had to have been emitted very early in the Universe, at redshift  $z \sim 30$ , when no galaxies have been detected yet. Instead, the models favour a scenario in which a dwarf galaxy, below the detection limits of current observations of the field, very close to the absorber ( $119 \text{ ckpc}/h$ ), with  $M_\star = 1.87 \times 10^9 M_\odot$ ,  $M_h = 9.67 \times 10^9 M_\odot$  and  $\text{SFR} = 0.07 M_\odot/\text{yr}$ , is responsible for the CIV absorption.

In addition, we find that the main drivers of the chemical enrichment of the IGM and CIV strong absorptions at  $z = 5.6$  are dwarf galaxies, which are mostly observed at a mean distance of  $700 \text{ ckpc}/h$  ( $106.1 \text{ pkpc}/h$ ) from the metal features. These galaxies have stellar masses  $M_\star$  in the range  $10^{8.46-9.42} M_\odot$  and  $\text{SFR}=0.01-2.5 M_\odot/\text{yr}$ . From these parameters, we infer an absolute magnitude  $M_{\text{UV}}$  in the range of  $(-20.5,-18.8)$ , derived by combining results of Song et al. (2016) with the Kennicutt relation (Kennicutt, 1998). We stress that the investigation with future surveys of these faint galaxies is fundamental, since they are believed to be the largest contributors of photons to the Reionization of Hydrogen.

Moreover, the study of the galaxy – absorber connection at high redshift shows that there is no correlation between the mass of the closest galaxy to the absorber and the distance

between them. On the other hand, we encounter a weak negative correlation between the column density of the CIV absorbers and the separation of the galaxy – absorber pairs. The largest column densities are preferentially seen when separations are smallest.

The column density relationships among different ionic species taught us many things. Initially, these diagrams were introduced in the analysis to compare the synthetic absorbers with the observations in hand. In addition, there was the intention to use low-to-high ionization states column density relationships to derive useful information on the neutral-to-ionized state of Hydrogen. This idea was dismissed in the project for different reasons. Firstly, the method used to recover the column density of Hydrogen is different from the one used for the ionic species. Secondly, there is a high scatter in the column density of the low-to-high ionization states considered in this work: i) Carbon states show a large dispersion in their column densities; ii) OVI is not observable at high redshift, therefore, it cannot be contrasted with OI; iii) there is not a dataset of quasar absorption lines with simultaneous detections of SiII and SiIV to compare with the synthetic absorptions. Instead, the column density relationships are used in this work to understand the statistical distribution of synthetic absorbers above the noise threshold ( $\log N \text{ (cm}^{-2}\text{)} \geq 12.5$ ) in comparison with the observational data sets by D’Odorico et al. (2013), Becker et al. (2006) and a recent detection from Bosman et al. (2017).

With the HM12 UVB, the high ionization states are relatively well represented, especially the distribution of  $N_{\text{CIV}}$ . The low ionization species incidence rates from our simulations are produced in a similar number of the observed detections, but the distribution of column densities is not well represented with respect to the observations, and column densities are overestimated for OI and SiII.

Detections of SiII associated to OI by Becker et al. (2006) show low column densities (even below the numerical noise threshold), therefore the distribution is very difficult to reproduce with mock spectra. Instead, OI observations by Becker et al. (2006) are in the range  $10^{13-14.7} \text{ cm}^{-2}$  (considered high column densities for quasar absorption features), and our simulations reproduce these absorptions, but also display a large number of absorbers with either low column density (around the noise threshold) or above  $10^{15} \text{ cm}^{-2}$ , not detected with Becker et al. (2006) sample.

CII shows some discrepancy with the observational data, and the column density correlations evidence a bimodality in the distribution of column densities, which was analyzed in Section 3.7.1. The results from that particular test reveal that current observations

are just detecting the CII absorbers in the IGM with low column densities, however, if observations in the CGM or in the proximities of galaxies -via either a larger statistical sample of random lines of sight or targeting galaxies in absorption will be done in the future-, large CII column densities should be detected.

The simulated column density relationships struggle to reproduce the observed trends, regardless of the feedback model used, possibly due to the UV background implemented. We inferred that the current HM12 normalization is not appropriate to study the low ionization states and that is why we decided to perform a test with a softer UVB in Chapter 5. Theoretical work from Bolton et al. (2011) and Finlator et al. (2015) showed that the resulting ion column densities are sensitive to variations in the normalization and hardness of the HM12 UVB.

Furthermore, our simulations are in good agreement with observations of the HI column density distribution function at  $z = 4$  and the HI cosmological mass density,  $\Omega_{\text{HI}}$ , at  $4 < z < 6$  (see Chapter 3). Notably, models that include molecular and low-temperature metal cooling are in better agreement with the observed HI cosmological mass density, because the chemistry of the module warrants that some HI at high density ( $N_{\text{HI}} > 2 \times 10^{20} \text{ cm}^{-2}$ ) is converted into H molecules ( $\text{H}_2$ , HD), and prevent the overproduction of neutral Hydrogen. Yet, all our models are compatible with the HI-CDDF reported in the literature up to  $z = 6$ , when the slow evolution of the HM12 is not a good representation of the variations of the H content in the IGM during the EoR. In addition, our results validate the estimate made in the observational work of Crighton et al. (2015) that DLA systems contribute  $\sim 80\%$  of the  $\Omega_{\text{HI}}$ .

In the second and final part of this thesis, variations of the UVB and metallicity in the photoionization modelling are presented. We explore the sensitivity of the metal ion observables at  $z = 6$  with different tests introduced in post-processing in Chapter 5. HI calculations are not presented in Chapter 5 for two reasons: 1) at high redshift a uniform UVB does not provide a good representation of the rapid evolution of HI during Reionization; 2) the self-shielding prescription from Rahmati et al. (2013) is calibrated with the photoionization rates of HM12. In the future, HI observables could be recalculated with the new HI self-shielding model from Chardin et al. (2017a), with best-fit parameters up to  $z = 10$ .

Due to difficulty in discriminating different models (feedback prescription and molec-

ular content) with the column density correlations, we have introduced an alternative method to count single ionic absorbers and compare them with the observational sample. The absorbers are compared directly in the range of column densities defined by the detections from D’Odorico et al. (2013) and Becker et al. (2006). This method shows a more fair comparison between the simulated and detected absorbers in each ion, and allows us to study and constrain the variations implemented in Chapter 5.

When the emissivity of the assumed UVB is reduced by one dex, we found that the total number of CIV absorbers significantly decreases compared with the original HM12 and absorbers in the range of the detections by D’Odorico et al. (2013) are definitely not well represented (e.g. there is a discrepancy of the CIV comoving mass density and CIV–CDDF with the observations). Instead, there is a moderate improvement in the calculated number of absorbers for SiIV with a softer UVB, in agreement with findings by Bolton et al. (2011).

It is particularly encouraging that the number of absorbers of CII, SiII and OI in the range of the observations also rises when the UVB implemented is softer, independently of the simulation used, and therefore the estimate of the column densities of these low ionization states improves, justifying the hypothesis that led us to do this test.

Additional work can be done in this direction, with a more moderated reduction in the emissivity of the UVB that would give better predictions for the column density of low ionization states, and be more compatible with the CIV absorbers incidence rate inferred with D’Odorico et al. (2013) observations and our mock spectra for that ion.

When comparing results at  $z = 6$  with the uniform HM12 UVB and a toy model with a flat high energy end in the UV emissivity (that accounts for a very aggressive contribution from quasars at this redshift), we found no change in the column density relationships and a very mild variation in the cosmological mass density of CII and CIV. These results indicate that variations in the UVB spectra from hard ionizing sources have a negligible effect on the considered ion states at wavelengths around or greater than  $912 \text{ \AA}$  (or energies about or below 1 Ryd). The results discussed in Chapter 5 are very close to the values computed with a uniform HM12. The conclusion in this case is clear: in order to have an effective improvement in the column density of the ionic states, it is necessary to change the spectrum at wavelengths where these states take place ( $192 \text{ \AA}$  for CIV  $\rightarrow$  CV to  $910 \text{ \AA}$  for OI  $\rightarrow$  OII). Our results show that not even an extreme contribution from quasars at  $z = 6$  alters results of metal transitions in an appreciable way.

Although AGN/black holes do not play a major role at high redshift (findings in Tescari et al. (2014) show that AGN–feedback is not overall efficient at high redshift, since the massive halos that host quasars are quite rare at  $z \geq 6$ ), if the absorptions happen in the vicinity of a quasar itself, the AGN driven ionization would have to be treated differently, for instance, applying a radiative transfer model in the proximity zone of the quasar.

On the other hand, variations of the metallicity introduced in the photoionization modelling in Chapter 5 do not reflect the true change due to a lower metallicity in the IGM/CGM at high redshift. The change in the metallicity in CLOUDY show very mild differences in the calculation of the column densities of the ions, due to the fact that our simulations are particularly robust describing observables in the IGM, but the procedure followed to produce the lines of sight inside the box has a very low likelihood to trace high density regions, where the effect of this metallicity change would be noticeable. The statistical distribution of the synthetic absorbers would change if we trace lines of sight at a given distance from the centre of the galaxies, yet, we prefer to mimic the observational selection method to assure the compatibility of our sample to the observational detections. A different approach to study the metallicity content at high redshift will be through new simulations with different stellar yields and IMFs and a posterior analysis to recover the observables for HI and metal ions with the pipeline proposed in Chapters 3 and 5.

The set of simulations was run and analysed while the pipeline was being tested and improved. Although there was some previous work on CIV at lower redshift (Tescari et al., 2011), setting up the pipeline at high redshift was quite demanding and allowed us to explore the calibration of simulated spectra at  $z > 5$ . As discussed in Chapter 3, the lack of reliable measurements of the mean normalized flux of the Lyman- $\alpha$  forest at  $z \geq 5$ , when the QSOs spectra display very large Gunn–Peterson troughs, leads to the absence of an effective parametrization of the optical depth of HI to re-scale the simulated Hydrogen and ion spectra. Different approaches have been used in the literature to reproduce the evolution of the metal ions independently from HI. We calculated the CIV column density distribution function at  $z = 4.8$  and  $5.6$  (see section 3.6.1), and found a scaling factor,  $A_{\text{CIV}}$ , which provided the best chi-by-eye agreement between the simulated and observed distribution functions.

This method is one of the possible alternatives to calibrate the synthetic spectra at high redshift. We chose to compare the simulated absorbers to the CIV-CDDF fitting functions kindly provided by D’Odorico et al. (2013), because these functions include observations

from  $4.3 \leq z \leq 6.4$  and cover a large redshift range within our simulations set.

A remarkable finding from Chapters 3 and 5 is that the choice of a UVB plays a major role in the number of ionic absorbers predicted by the simulations. The UVB implemented has more impact in the overall results than feedback or the molecular and low-temperature cooling model. We stress that this conclusion just holds for the metal ions. In the framework of this thesis, it is not possible to study variations in the HI content when the UVB is modified if the HI self-shielding prescription is not changed accordingly.

The feedback prescription in our models determines mostly the chemical enrichment history, with energy driven winds predicting less mass density for metals than the MDW model, because the first one is more efficient at quenching the star formation rate and, as a consequence, the production of metals at  $z < 6$ .

On the other hand, the implementation of low-temperature metal and molecular cooling affects mostly neutral Hydrogen statistics. The molecular cooling suppresses HI in regions where the majority of it is self-shielded, due to the conversion of HI to H<sub>2</sub> at high densities. This effect is relevant when computing the HI column density distribution function and cosmological mass density. It is important to mention that the introduction of this module has a little, but not negligible effect on the calculation of the cosmological mass densities of CII and CIV. Simulations with molecular and low-temperature cooling predict a slightly higher CII and CIV mass densities, independently of the hardness of the UVB or the feedback model considered.

Besides, we found that what drives the relative mass density of metal ion species, and indeed the EoR, in reality is the UVB: the grand sum of ionizing flux from quasars and galaxies. Although results from Chapter 5 show that the contribution of quasars to the very hard ionizing end of the UVB has little effect on the observed metals, it may well be that a transition from a galaxy-dominated UVB to a quasar-dominated one is occurring around the tail end of the Reionization, as evidenced by the observed transition from  $\Omega_{\text{CII}}$  to  $\Omega_{\text{CIV}}$ , also seen in the simulations presented. This assumption will be tested once JWST<sup>1</sup> measures the faint end of both the galaxy and quasar luminosity function out to  $z \sim 10$ .

On a different note, we stress that radiative transfer effects are not included in our simulations, and therefore, we do not follow the progression of Reionization, nor the evolu-

---

<sup>1</sup><https://www.jwst.nasa.gov/>

tion of the HII bubbles or their topology. The implicit assumption is that our boxes (that are small compared to the size of the HII bubbles at the redshifts of interest) represent a region of the Universe already reionized at a level given by the HM12 UVB. At  $6 < z < 8$ , chemical enrichment occurs mostly inside and in close proximity of galaxies (interstellar medium, CGM and high density IGM) where, assuming an inside-out progression of Reionization, the gas in which metals lie should be ionized. Although proper RT calculations would be more accurate, they are extremely expensive from the computational point of view and they could be done in the future. A first approach in this direction has been done by Finlator et al. (2015).

A fundamental issue persists in the field from the numerical point of view: it is extremely challenging to model the low ionization states present in the gas and provide a good description of the environment where they lie, mainly due to insufficient resolution and a proper self-shielding treatment for the ions. As discussed in Chapter 3, we have considered only the effect of HI self-shielding (Rahmati et al., 2013), but did not introduce any self-shielding of low ionization absorbers (which lie in clumpy structures). A first attempt has been proposed for DLA systems at low redshift by Bird et al. (2015). However, current works miss this component at high redshift because it is still not well understood how high density regions self-shield the gas during the progression of Reionization.

Future work will be focused on understanding the impact of the UVB, since it has been shown to have more influence on metal absorbers than other components of our models. Using the performed tests described in Chapter 5 with a modified UVB, it will be possible to make a sensible fine-tuning of the UVB with the observational constraints from metal absorption lines at  $z \sim 6$  at wavelengths  $> 912 \text{ \AA}$ .

Once observations of other ionic species and metals at high redshift become available, we can introduce them self-consistently to our pipeline and derive compelling conclusions about their physical properties and evolution. The pipeline has been designed in such way that is easily adaptable. There are high expectations of detecting metal absorption lines towards the highest redshift quasars ( $z \sim 7$ ). These metal absorbers will be also observed with the new generation of 25-40 m optical/IR ground-based telescopes (e.g. GMT<sup>2</sup>, TMT<sup>3</sup>, E-ELT<sup>4</sup>). Moreover, fainter galaxies can be detected by their absorption

---

<sup>2</sup><https://www.gmto.org/>

<sup>3</sup><http://www.tmt.org/>

<sup>4</sup><https://www.eso.org/sci/facilities/eelt/>

---

lines compared to those seen in emission. Studying the faint galaxy population is critical to the census of ionizing photons required to keeping the IGM ionized.

Simulations of these metal absorbing galaxies are critical as they can provide more than a “snapshot” in time. By studying the statistics of quasar absorber–galaxy pairs at high redshift and tracking the metals that give rise to an absorption feature back to the parent galaxy, one can in principle derive the percentage contribution of different galaxies to a given metal absorption system. The galaxy–absorber pair analysis presented in Chapter 4 is an example of the first step towards a much more robust analysis of many pairs. Simulations will provide critical interpretations of IFU galaxy observations around high redshift lines of sight, such as those conducted with Keck/KCWI<sup>5</sup> and VLT/MUSE<sup>6</sup>.

Finally, it is reassuring to note that theoretical predictions made from the simulated spectra have been borne true over the duration of the thesis. A reliable calibration of the theoretical models with the observations in hand allowed us to have a good prediction of the  $6.4 < z < 7$  CIV-CCDF before Bosman et al. (2017) results were published. Detections of DLAs from Bird et al. (2017) were also compatible with our preceding HI predictions. We rely on our theoretical models presented in this thesis, since they provide a successful description of the state of the IGM at the tail of Reionization (through the introduction of a UVB in post-process) and the pollution of the IGM with metals released by supernovae and stars. We have compared HI and ions observables available at high redshift and found that most of the results discussed are compatible in the redshift range  $4 < z < 8$ . It is worth noting that all results from mock spectra will be improved in the future with more observational detections of ion absorbers in the high redshift quasar spectra.

---

<sup>5</sup><https://www2.keck.hawaii.edu/inst/kcwi/>

<sup>6</sup><https://www.eso.org/sci/facilities/develop/instruments/muse.html>



# Bibliography

- Aarseth S. J., 1985, in IAU Symposium, Vol. 113, Dynamics of Star Clusters, Goodman J., Hut P., eds., pp. 251–258
- Abel T., Bryan G. L., Norman M. L., 2002, *Science*, 295, 93
- Alam S. et al., 2015, *ApJS*, 219, 12
- Appel A. W., 1985, *SIAM Journal on Scientific and Statistical Computing*, vol. 6, no. 1, January 1985, p. 85-103., 6, 85
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Aubourg É. et al., 2015, *Phys. Rev. D*, 92, 123516
- Barkana R., 2002, , 7, 85
- Barkana R., Loeb A., 2001, *Phys. Rep.*, 349, 125
- Barnes J. E., 1986, in *Lecture Notes in Physics*, Berlin Springer Verlag, Vol. 267, The Use of Supercomputers in Stellar Dynamics, Hut P., McMillan S. L. W., eds., p. 175
- Barnes L. A., Haehnelt M. G., 2009, *MNRAS*, 397, 511
- Bauer A., Springel V., Vogelsberger M., Genel S., Torrey P., Sijacki D., Nelson D., Hernquist L., 2015, *MNRAS*, 453, 3593
- Baugh C. M., 2006, *Reports on Progress in Physics*, 69, 3101
- Becker G. D., Bolton J. S., Lidz A., 2015a, , 32, e045
- Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015b, *MNRAS*, 447, 3402
- Becker G. D., Sargent W. L. W., Rauch M., Calverley A. P., 2011, *ApJ*, 735, 93
- Becker G. D., Sargent W. L. W., Rauch M., Simcoe R. A., 2006, *ApJ*, 640, 69
- Becker R. H. et al., 2001, *AJ*, 122, 2850
- Bennett C. L. et al., 2013, *ApJS*, 208, 20
- Bird S., Garnett R., Ho S., 2017, *MNRAS*, 466, 2111
- Bird S., Haehnelt M., Neeleman M., Genel S., Vogelsberger M., Hernquist L., 2015, *MNRAS*, 447, 1834

- Bird S., Vogelsberger M., Haehnelt M., Sijacki D., Genel S., Torrey P., Springel V., Hernquist L., 2014, *MNRAS*, 445, 2313
- Blake C. et al., 2011, *MNRAS*, 418, 1725
- Boksenberg A., Sargent W. L. W., 2015, *ApJS*, 218, 7
- Bolton J. S., Haehnelt M. G., Viel M., Springel V., 2005, in *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, Williams P., Shu C.-G., Menard B., eds., pp. 219–224
- Bolton J. S., Haehnelt M. G., Warren S. J., Hewett P. C., Mortlock D. J., Venemans B. P., McMahon R. G., Simpson C., 2011, *MNRAS*, 416, L70
- Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, *MNRAS*, 464, 897
- Bondi H., 1952, *MNRAS*, 112, 195
- Bongiorno A. et al., 2007, *A&A*, 472, 443
- Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53
- Borgani S. et al., 2004, *MNRAS*, 348, 1078
- Bosman S. E. I., Becker G. D., Haehnelt M. G., Hewett P. C., McMahon R. G., Mortlock D. J., Simpson C., Venemans B. P., 2017, *ArXiv e-prints*
- Bouwens R. J. et al., 2009, *ApJ*, 705, 936
- Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, *ApJ*, 811, 140
- Bromm V., 2003, in *Bulletin of the American Astronomical Society*, Vol. 35, American Astronomical Society Meeting Abstracts, p. 1377
- Bromm V., Coppi P. S., Larson R. B., 2002, *ApJ*, 564, 23
- Bromm V., Loeb A., 2003, in *American Institute of Physics Conference Series*, Vol. 666, *The Emergence of Cosmic Structure*, Holt S. H., Reynolds C. S., eds., pp. 73–84
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000

- Burchett J., Tripp T. M., Bordoloi R., Willmer C., 2016, in American Astronomical Society Meeting Abstracts, Vol. 227, American Astronomical Society Meeting Abstracts, p. 109.08
- Bürzle F., Clark P. C., Stasyszyn F., Greif T., Dolag K., Klessen R. S., Nielaba P., 2011, MNRAS, 412, 171
- Carswell R. F., Webb J. K., 2014, VPFIT: Voigt profile fitting program. Astrophysics Source Code Library
- Cen R., Chisari N. E., 2011, ApJ, 731, 11
- Cen R., Ostriker J., 1992, ApJ, 393, 22
- Cen R., Ostriker J. P., 1993, ApJ, 417, 415
- Chabrier G., 2003, PASP, 115, 763
- Chardin J., Kulkarni G., Haehnelt M. G., 2017a, ArXiv e-prints
- Chardin J., Puchwein E., Haehnelt M. G., 2017b, MNRAS, 465, 3429
- Chen H.-W., Johnson S. D., Zahedy F. S., Rauch M., Mulchaey J. S., 2017, ApJ, 842, L19
- Choudhury T. R., 2009, Current Science, 97, 841
- Choudhury T. R., Ferrara A., 2005, MNRAS, 361, 577
- Choudhury T. R., Haehnelt M. G., Regan J., 2009, MNRAS, 394, 960
- Chuzhoy L., Alvarez M. A., Shapiro P. R., 2006, ApJ, 648, L1
- Ciardi B., Ferrara A., 2005, Space Sci. Rev., 116, 625
- Cole S. et al., 2005, MNRAS, 362, 505
- Cooke J., Ryan-Weber E. V., Garel T., Díaz C. G., 2014, MNRAS, 441, 837
- Cooksey K. L., Thom C., Prochaska J. X., Chen H.-W., 2010, ApJ, 708, 868
- Couchman H. M. P., 1991, ApJ, 368, L23
- Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, ApJ, 452, 797
- Cowie L. L., Barger A. J., Trouille L., 2009, ApJ, 692, 1476

- Crichton N. H. M. et al., 2015, MNRAS, 452, 217
- Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1997, ApJ, 488, 532
- Croton D. J., 2006, MNRAS, 369, 1808
- Cucciati O. et al., 2012, A&A, 539, A31
- Danforth C. W., Shull J. M., 2008, ApJ, 679, 194
- Davies F. B., Furlanetto S. R., 2016, Monthly Notices of the Royal Astronomical Society, 460, 1328
- Díaz C. G., Ryan-Weber E. V., Cooke J., Koyama Y., Ouchi M., 2015, MNRAS, 448, 1240
- Dijkstra M., Haiman Z., Loeb A., 2004, ApJ, 613, 646
- Dodelson S., 2003, Modern Cosmology. Academic Press, Elsevier Science
- D'Odorico V., Calura F., Cristiani S., Viel M., 2010, MNRAS, 401, 2715
- D'Odorico V. et al., 2016, MNRAS, 463, 2690
- D'Odorico V. et al., 2013, MNRAS, 435, 1198
- Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
- Dolag K., Vazza F., Brunetti G., Tormen G., 2005, MNRAS, 364, 753
- Efstathiou G., Bond J. R., White S. D. M., 1992, MNRAS, 258, 1P
- Efstathiou G., Eastwood J. W., 1981, MNRAS, 194, 503
- Efstathiou G., Jones B. J. T., 1980, Comments on Astrophysics, 8, 169
- Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11
- Evrard A. E., 1988, in Post-Recombination Universe, Kaiser N., Lasenby A. N., eds., pp. 363–366
- Evrard A. E., 1990, ApJ, 363, 349
- Fabjan D., Borgani S., Tornatore L., Saro A., Murante G., Dolag K., 2010, MNRAS, 401, 1670
- Fan X., 2006, , 50, 665

- Fan X., 2012, *Research in Astronomy and Astrophysics*, 12, 865
- Fan X., Carilli C. L., Keating B., 2006, *ARA&A*, 44, 415
- Fardal M. A., Giroux M. L., Shull J. M., 1998, *AJ*, 115, 2206
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Ferland G., Williams R., 2016, *Spectral Modeling in Astrophysics - The Physics of Non-equilibrium Clouds*
- Ferland G. J. et al., 2017, *ArXiv e-prints*
- Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761
- Ferland G. J. et al., 2013, , 49, 137
- Field G. B., 1958, *Proceedings of the IRE*, 46, 240
- Finlator K., Muñoz J. A., Oppenheimer B. D., Oh S. P., Özel F., Davé R., 2013, *MNRAS*, 436, 1818
- Finlator K., Oppenheimer B. D., Davé R., Zackrisson E., Thompson R., Huang S., 2016, *MNRAS*, 459, 2299
- Finlator K., Thompson R., Huang S., Davé R., Zackrisson E., Oppenheimer B. D., 2015, *MNRAS*, 447, 2526
- Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, *ApJ*, 473, 576
- Fontanot F., Cristiani S., Pfrommer C., Cupani G., Vanzella E., 2014, *MNRAS*, 438, 2097
- Freeman K. C., 1970, *ApJ*, 160, 811
- Furlanetto S. R., Mesinger A., 2009, *MNRAS*, 394, 1667
- Furlanetto S. R., Oh S. P., 2008, *ApJ*, 681, 1
- Furlanetto S. R., Oh S. P., Briggs F. H., 2006, *Phys. Rep.*, 433, 181
- Furlanetto S. R., Schaye J., Springel V., Hernquist L., 2005, *ApJ*, 622, 7
- Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, *ApJ*, 613, 1

- García L. A., Tescari E., Ryan-Weber E. V., Wyithe J. S. B., 2017a, *MNRAS*, 470, 2494
- García L. A., Tescari E., Ryan-Weber E. V., Wyithe J. S. B., 2017b, *MNRAS*, 469, L53
- Gingold R. A., Monaghan J. J., 1982, *Journal of Computational Physics*, 46, 429
- Gnat O., Ferland G. J., 2012, *ApJS*, 199, 20
- González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, 735, L34
- Gott, III J. R., 1977, *ARA&A*, 15, 235
- Grazian A. et al., 2017, *A&A*, 602, A18
- Gunn J. E., 1982, in *Astrophysical Cosmology Proceedings*, Brueck H. A., Coyne G. V., Longair M. S., eds.
- Gunn J. E., Peterson B. A., 1965, *ApJ*, 142, 1633
- Haardt F., 1999, *Mem. Soc. Astron. Italiana*, 70, 261
- Haardt F., Madau P., 2001, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, Neumann D. M., Tran J. T. V., eds., p. 64
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Haiman Z., 2016, *Understanding the Epoch of Cosmic Reionization: Challenges and Progress*, 423, 1
- Haiman Z., Loeb A., 1998, *ApJ*, 503, 505
- Haiman Z., Rees M. J., 2001, *ApJ*, 556, 87
- Hassan S., Davé R., Mitra S., Finlator K., Ciardi B., Santos M. G., 2017, *ArXiv e-prints*
- Hernquist L., Katz N., 1989, *ApJS*, 70, 419
- Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J., 1996, *ApJ*, 457, L51
- Hildebrandt H. et al., 2010, *A&A*, 523, A31
- Hui L., Gnedin N. Y., Zhang Y., 1997, *ApJ*, 486, 599
- Hutter A., Dayal P., Partl A. M., Müller V., 2014, *MNRAS*, 441, 2861

- Jeans J. H., 1902, *Philosophical Transactions of the Royal Society of London Series A*, 199, 1
- Jenkins A. et al., 1998, *ApJ*, 499, 20
- Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, *MNRAS*, 321, 372
- Jiang L. et al., 2013, *ApJ*, 772, 99
- Kang H., Ostriker J. P., Cen R., Ryu D., Hernquist L., Evrard A. E., Bryan G. L., Norman M. L., 1994, *ApJ*, 430, 83
- Kashikawa N. et al., 2011, *ApJ*, 734, 119
- Katsianis A., Tescari E., Blanc G., Sargent M., 2017, *MNRAS*, 464, 4977
- Katsianis A., Tescari E., Wyithe J. S. B., 2015, *MNRAS*, 448, 3001
- Katsianis A., Tescari E., Wyithe J. S. B., 2016, , 33, e029
- Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
- Keating B., Miller N., 2006, , 50, 184
- Keating L. C., Haehnelt M. G., Becker G. D., Bolton J. S., 2014, *MNRAS*, 438, 1820
- Keating L. C., Puchwein E., Haehnelt M. G., Bird S., Bolton J. S., 2016, *MNRAS*, 461, 606
- Kennicutt, Jr. R. C., 1998, *ApJ*, 498, 541
- Klypin A. A., Shandarin S. F., 1983, *MNRAS*, 204, 891
- Kotarba H., Lesch H., Dolag K., Naab T., Johansson P. H., Stasyszyn F. A., 2009, *MNRAS*, 397, 733
- Kravtsov A. V., Klypin A., Hoffman Y., 2002, *ApJ*, 571, 563
- Kroupa P., Tout C. A., Gilmore G., 1993, *MNRAS*, 262, 545
- Lacey C., 2001, in *Astronomical Society of the Pacific Conference Series*, Vol. 222, *The Physics of Galaxy Formation*, Umemura M., Susa H., eds., p. 273
- Larson D. et al., 2011, *ApJS*, 192, 16

- Liddle A., Lyth D., 2000, *Cosmological Inflation and Large-Scale Structure*, *Cosmological Inflation and Large-scale Structure*. Cambridge University Press
- Lidz A., 2016, *Understanding the Epoch of Cosmic Reionization: Challenges and Progress*, 423, 23
- Lidz A., Oh S. P., Furlanetto S. R., 2006, *ApJ*, 639, L47
- Linde A. D., 1980a, *Physics Letters B*, 96, 293
- Linde A. D., 1980b, *Physics Letters B*, 93, 394
- Linde A. D., 1980c, *Physics Letters B*, 92, 119
- Linde A. D., 1980d, *Physics Letters B*, 93, 327
- Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, 462, 235
- Loeb A., 2006, *Scientific American*, 295
- Loeb A., 2010, *How Did the First Stars and Galaxies Form?*
- Loeb A., Rasio F. A., 1994, *ApJ*, 432, 52
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Madau P., Dickinson M., 2014, *Ann. Rev. Astron. Astrophys.*, 52, 415
- Madau P., Haardt F., 2015, *ArXiv e-prints*
- Maio U., Ciardi B., Dolag K., Tornatore L., Khochfar S., 2010, *MNRAS*, 407, 1003
- Maio U., Dolag K., Ciardi B., Tornatore L., 2007, *MNRAS*, 379, 963
- Maio U., Tescari E., 2015, *MNRAS*, 453, 3798
- Martin C. L., 2005, *ApJ*, 621, 227
- Matteucci F., 2003, *The Chemical Evolution of the Galaxy*
- McQuinn M., 2016, *ARA&A*, 54, 313
- McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, *MNRAS*, 377, 1043

- Meiksin A., 2005, MNRAS, 356, 596
- Mellema G., Iliiev I. T., Pen U.-L., Shapiro P. R., 2006, MNRAS, 372, 679
- Mesinger A., Furlanetto S., 2007, ApJ, 669, 663
- Mestel L., 1963, MNRAS, 126, 553
- Miralda-Escudé J., 2003, ApJ, 597, 66
- Miralda-Escudé J., Cen R., Ostriker J. P., Rauch M., 1996, ApJ, 471, 582
- Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1
- Mitra S., Choudhury T. R., Ferrara A., 2015, ArXiv e-prints
- Monaghan J. J., Lattanzio J. C., 1985, A&A, 149, 135
- Mortlock D. J. et al., 2011, Nature, 474, 616
- Nagamine K., Springel V., Hernquist L., 2004, MNRAS, 348, 421
- Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
- Navarro J. F., White S. D. M., 1993, MNRAS, 265, 271
- Noterdaeme P., Petitjean P., Srianand R., Ledoux C., López S., 2011, A&A, 526, L7
- Nusser A., 2005, MNRAS, 359, 183
- O'Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666
- Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
- Oppenheimer B. D., Davé R., Finlator K., 2009, MNRAS, 396, 729
- Oppenheimer B. D., Schaye J., 2013, MNRAS, 434, 1043
- Osterbrock D., Ferland G., 2006, *Astrophysics Of Gas Nebulae and Active Galactic Nuclei*.  
University Science Books
- Ouchi M. et al., 2010, ApJ, 723, 869
- Ouchi M. et al., 2004, ApJ, 611, 685
- Padmanabhan T., 1993, *Structure Formation in the Universe*. p. 499

- Padovani P., Matteucci F., 1993, *ApJ*, 416, 26
- Pakmor R., Röpke F. K., Weiss A., Hillebrandt W., 2008, *A&A*, 489, 943
- Pallottini A., Ferrara A., Gallerani S., Salvadori S., D’Odorico V., 2014, *MNRAS*, 440, 2498
- Pawlik A. H., Schaye J., 2008, *MNRAS*, 389, 651
- Pawlik A. H., Schaye J., 2011, *MNRAS*, 412, 1943
- Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, *MNRAS*, 394, 1812
- Peacock J. A., 1999, *Cosmological Physics*. p. 704
- Peebles P. J. E., 1993, *Principles of Physical Cosmology*
- Péroux C., McMahon R. G., Storrie-Lombardi L. J., Irwin M. J., 2003, *MNRAS*, 346, 1103
- Pettini M., Madau P., Bolte M., Prochaska J. X., Ellison S. L., Fan X., 2003, *ApJ*, 594, 695
- Planck Collaboration et al., 2015, *ArXiv e-prints*
- Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, *MNRAS*, 431, 1487
- Pontzen A. et al., 2008, *MNRAS*, 390, 1349
- Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, *ApJ*, 635, 123
- Prochaska J. X., Wolfe A. M., 2009, *ApJ*, 696, 1543
- Puchwein E., Springel V., 2013, *MNRAS*, 428, 2966
- Quinn T., Katz N., Stadel J., Lake G., 1997, *ArXiv Astrophysics e-prints*
- Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, *ApJ*, 782, L29
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, 430, 2427
- Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, *MNRAS*, 452, 2034

- Rahmati A., Schaye J., Crain R. A., Oppenheimer B. D., Schaller M., Theuns T., 2016, MNRAS
- Rees M. J., 2000, Phys. Rep., 333, 203
- Refregier A., Teyssier R., 2002, Phys. Rev. D, 66, 043002
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
- Robertson B. E. et al., 2013, ApJ, 768, 71
- Ryan-Weber E. V., Pettini M., Madau P., 2006, MNRAS, 371, L78
- Ryan-Weber E. V., Pettini M., Madau P., Zych B. J., 2009, MNRAS, 395, 1476
- Salpeter E. E., 1955, ApJ, 121, 161
- Salvaterra R., Haardt F., Ferrara A., 2005, MNRAS, 362, L50
- Schaye J., 2001, ApJ, 559, 507
- Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768
- Sellwood J. A., 1987, ARA&A, 25, 151
- Shapiro P. R., Giroux M. L., Babul A., 1994, ApJ, 427, 25
- Shen S., Madau P., Guedes J., Mayer L., Prochaska J. X., Wadsley J., 2013, ApJ, 765, 89
- Siana B. et al., 2008, ApJ, 675, 49
- Simcoe R. A., 2006, ApJ, 653, 977
- Simcoe R. A., 2011, ApJ, 738, 159
- Simcoe R. A. et al., 2011, ApJ, 743, 21
- Simcoe R. A., Sullivan P. W., Cooksey K. L., Kao M. M., Matejek M. S., Burgasser A. J., 2012, Nature, 492, 79
- Smith A., Bromm V., Loeb A., 2017, ArXiv e-prints
- Sokasian A., Abel T., Hernquist L., 2002, MNRAS, 332, 601
- Song M. et al., 2016, ApJ, 825, 5

- Songaila A., 2001, *ApJ*, 561, L153
- Songaila A., 2005, *AJ*, 130, 1996
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., 2010, *ARA&A*, 48, 391
- Springel V., Di Matteo T., Hernquist L., 2005a, *MNRAS*, 361, 776
- Springel V., Frenk C. S., White S. D. M., 2006, *Nature*, 440, 1137
- Springel V., Hernquist L., 2002, *MNRAS*, 333, 649
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V. et al., 2005b, *Nature*, 435, 629
- Springel V. et al., 2005c, *Nature*, 435, 629
- Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, *ApJ*, 519, 1
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289
- Stiavelli M. et al., 2005, *ApJ*, 622, L1
- Sunyaev R. A., Zeldovich I. B., 1980, *ARA&A*, 18, 537
- Tegmark M. et al., 2006, *Phys. Rev. D*, 74, 123507
- Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, *ApJ*, 474, 1
- Telfer R. C., Zheng W., Kriss G. A., Davidsen A. F., 2002, *ApJ*, 565, 773
- Tescari E., Katsianis A., Wyithe J. S. B., Dolag K., Tornatore L., Barai P., Viel M., Borgani S., 2014, *MNRAS*, 438, 3490
- Tescari E., Viel M., D'Odorico V., Cristiani S., Calura F., Borgani S., Tornatore L., 2011, *MNRAS*, 411, 826
- Tescari E., Viel M., Tornatore L., Borgani S., 2009, *MNRAS*, 397, 411
- Theuns T., Leonard A., Efstathiou G., 1998, *MNRAS*, 297, L49

- Thielemann F.-K. et al., 2003, *Nuclear Physics A*, 718, 139
- Thomas R. M., Zaroubi S., 2008, *MNRAS*, 384, 1080
- Tormen G., Bouchet F. R., White S. D. M., 1997, *MNRAS*, 286, 865
- Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, *MNRAS*, 382, 1050
- Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, *MNRAS*, 445, 794
- Umemura M., Loeb A., Turner E. L., 1993, *ApJ*, 419, 459
- van den Hoek L. B., Groenewegen M. A. T., 1997, *A&AS*, 123, 305
- van der Kruit P. C., 1987, *A&A*, 173, 59
- van der Kruit P. C., Freeman K. C., 2011, *ARA&A*, 49, 301
- Vanzella E. et al., 2015, *A&A*, 576, A116
- Vanzella E. et al., 2016, *ApJ*, 825, 41
- Volonteri M., Gnedin N. Y., 2009, *ApJ*, 703, 2113
- Weinberg S., 2008, *Cosmology*, *Cosmology*. OUP Oxford
- White S. D. M., 1983, in *IAU Symposium, Vol. 100, Internal Kinematics and Dynamics of Galaxies*, Athanassoula E., ed., pp. 337–344
- White S. D. M., Frenk C. S., 1991, *ApJ*, 379, 52
- Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, *MNRAS*, 399, 574
- Willott C. J. et al., 2010, *AJ*, 139, 906
- Wolfe A. M., Gawiser E., Prochaska J. X., 2005, *ARA&A*, 43, 861
- Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D., 1986, *ApJS*, 61, 249
- Wood K., Loeb A., 2000, *ApJ*, 545, 86
- Woosley S. E., Weaver T. A., 1995, *ApJS*, 101, 181
- Wurster J., Thacker R. J., 2013, *MNRAS*, 431, 2513
- Wyithe J. S. B., Loeb A., 2003, *ApJ*, 586, 693

- Wyithe J. S. B., Loeb A., 2004, *Nature*, 427, 815
- Wyithe J. S. B., Morales M. F., 2007, *MNRAS*, 379, 1647
- Xu G., 1995, *ApJS*, 98, 355
- Yepes C., Kates R., Klypin A., Khokhlov A., 1996, in *Astronomical Society of the Pacific Conference Series*, Vol. 94, Mapping, Measuring, and Modelling the Universe, Coles P., Martinez V., Pons-Borderia M.-J., eds., p. 125
- Yoshida N., 2010, in *COSPAR Meeting*, Vol. 38, 38th COSPAR Scientific Assembly, p. 2
- Yoshida N., Sokasian A., Hernquist L., Springel V., 2003, *ApJ*, 598, 73
- Zafar T., Péroux C., Popping A., Milliard B., Deharveng J.-M., Frank S., 2013, *A&A*, 556, A141
- Zahn O. et al., 2012, *ApJ*, 756, 65
- Zaldarriaga M., 1997, *Phys. Rev. D*, 55, 1822
- Zaroubi S., 2013, in *Astrophysics and Space Science Library*, Vol. 396, The First Galaxies, Wiklind T., Mobasher B., Bromm V., eds., p. 45
- Zaroubi S., Thomas R. M., Sugiyama N., Silk J., 2007, *MNRAS*, 375, 1269
- Zwaan M. A., Prochaska J. X., 2006, *ApJ*, 643, 675

# A

## Appendix A

As discussed in Chapter 5, the ion column density relationships are quite insensitive to variations of the UVB. Although the incidence rate of synthetic absorbers is well represented by the simulations, the comparison of single ions with detections of D’Odorico et al. (2013) and Becker et al. (2006) at  $z = 6$  seems more robust.

In this Appendix, the column density relationships of the tests developed in Chapter 5 are presented for completeness. Figures A.1, A.2 and A.3 display the effect of a change in the normalization of HM12 in the column density relations at  $z = 6$  for the runs Ch 18 512 MDW, Ch 18 512 MDW mol and Ch 18 512 EDW, respectively.

Figure A.4 shows the result when a variation of HM12 is introduced in the high energy range (that accounts for a hard contribution of the quasars in the UVB spectrum) at  $z = 6$ . This final test was only made with the reference simulation Ch 18 512 MDW. The observational detections are very few at high redshift and limited to a short range of column density, most likely associated to the IGM. Therefore, the column density relationships are not the optimal method to distinguish the effects of different feedback prescriptions and/or molecular cooling content.

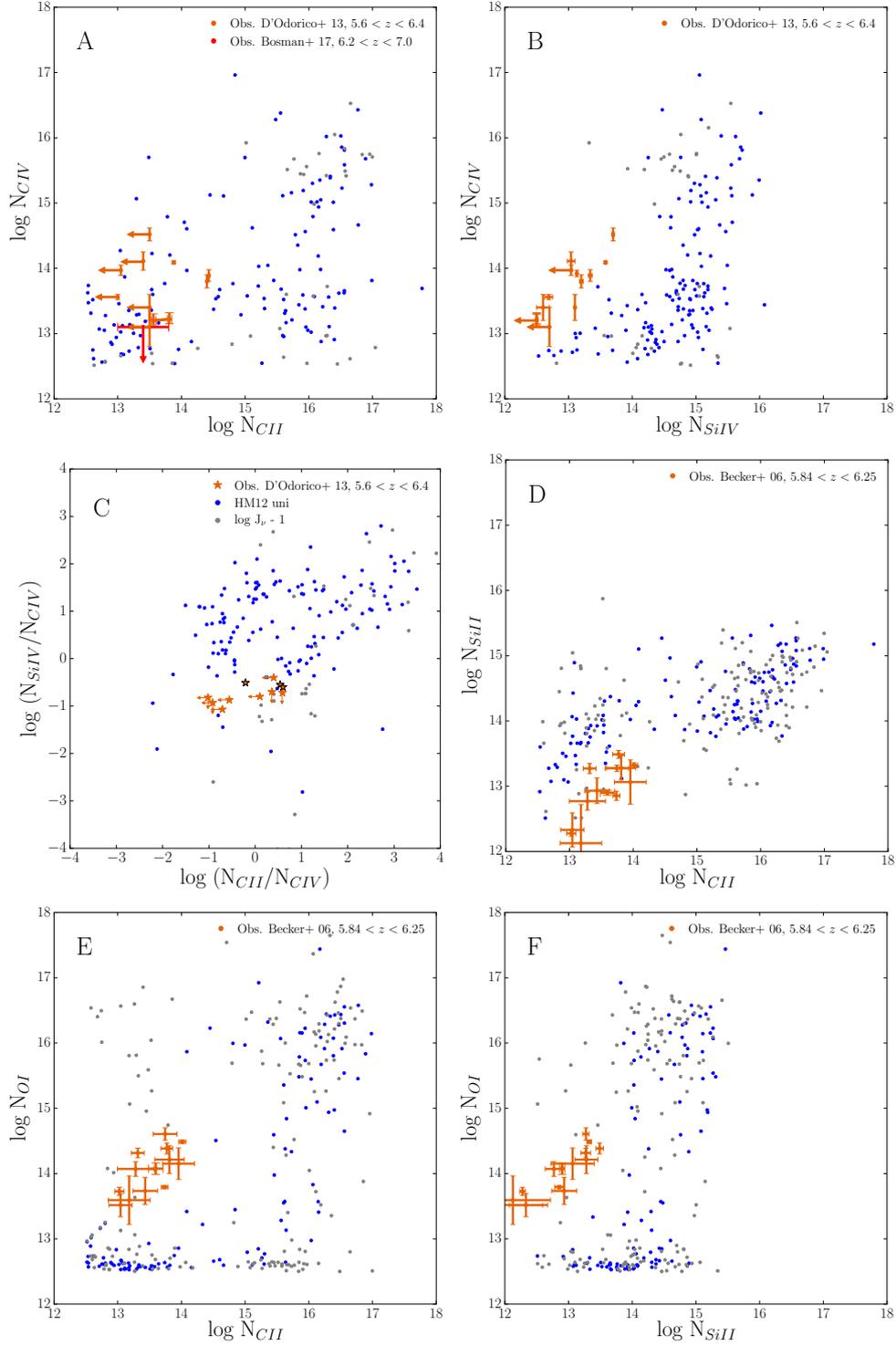


Figure A.1 Column density relationships among metal absorption lines at  $z = 6$  when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd. The simulation run used is Ch 18 512 MDW. Panels A, B and C compare with systems observed by D'Odorico et al. (2013) in the redshift range of (5.6–6.4) and D, E and F with absorbers detected by Becker et al. (2006) in the redshift range of (5.84–6.25), both in orange. An additional data point from Bosman et al. (2017) has been introduced in panel A, in red. The upper limit in CIV is found in the redshift range of (6.2–7.0). From the synthetic spectra of each ion, the column densities are calculated and displayed. The simulated data shown are above the threshold noise at  $\log N_{th} = 12.5$  ( $\text{cm}^{-2}$ ). In panel C, the observations by D'Odorico et al. (2013) produced 3 systems with column densities reported (represented by the black stars) and 8 with upper limits in  $N_{CII}$  and/or  $N_{SiIV}$ , represented by arrows.



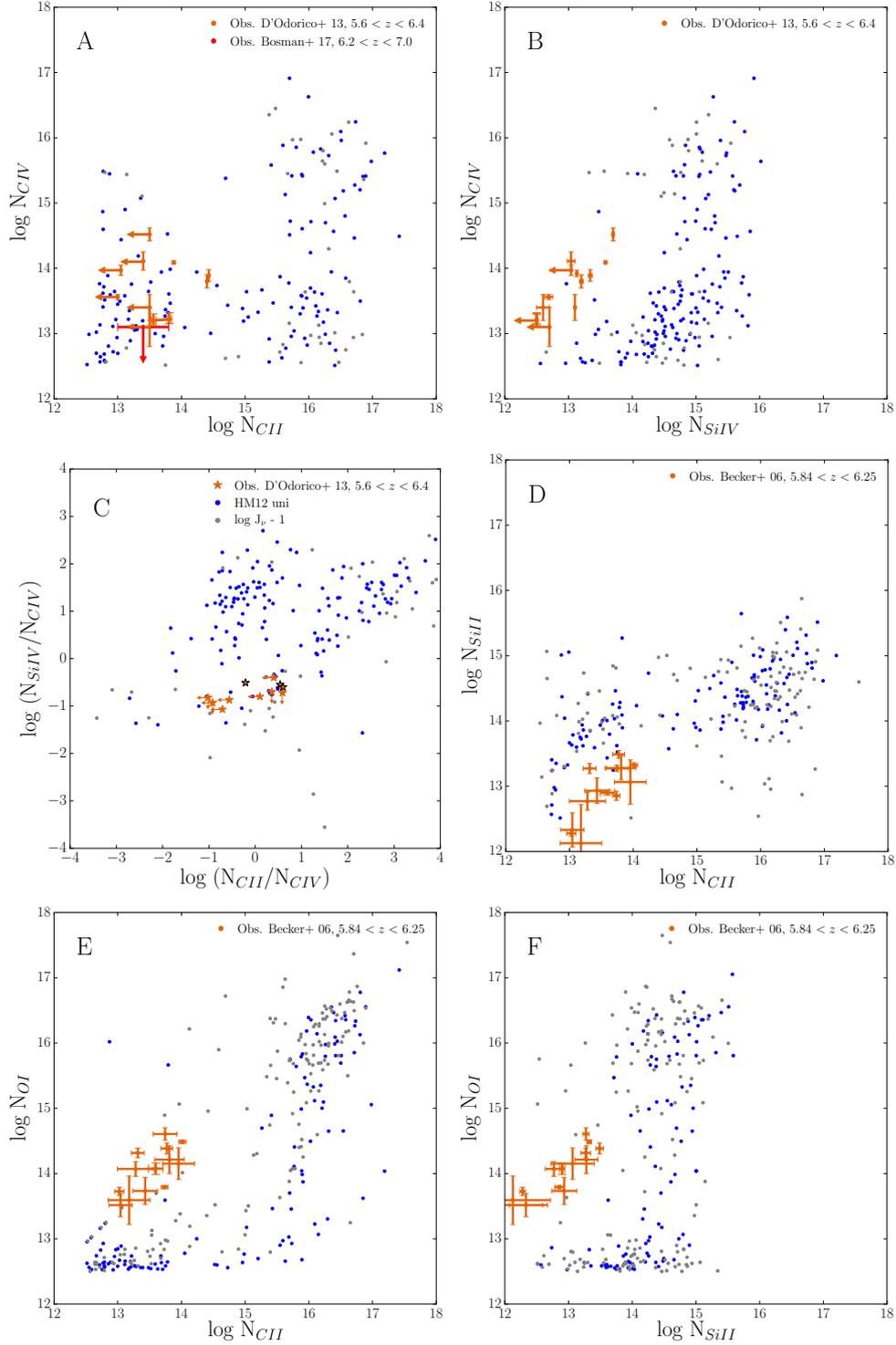


Figure A.3 Column density relationships among metal absorption lines at  $z = 6$  when the normalization of the uniform HM12 UVB is reduced by one dex at 1 Ryd. The simulation run used is Ch 18 512 EDW. Panels A, B and C compare with systems observed by D'Odorico et al. (2013) in the redshift range of (5.6–6.4) and D, E and F with absorbers detected by Becker et al. (2006) in the redshift range of (5.84–6.25), both in orange. An additional data point from Bosman et al. (2017) has been introduced in panel A, in red. The upper limit in CIV is found in the redshift range of (6.2–7.0). From the synthetic spectra of each ion, the column densities are calculated and displayed. The simulated data shown are above the threshold noise at  $\log N_{th} = 12.5$  ( $\text{cm}^{-2}$ ). In panel C, the observations by D'Odorico et al. (2013) produced 3 systems with column densities reported (represented by the black stars) and 8 with upper limits in  $N_{CII}$  and/or  $N_{SiIV}$ , represented by arrows.

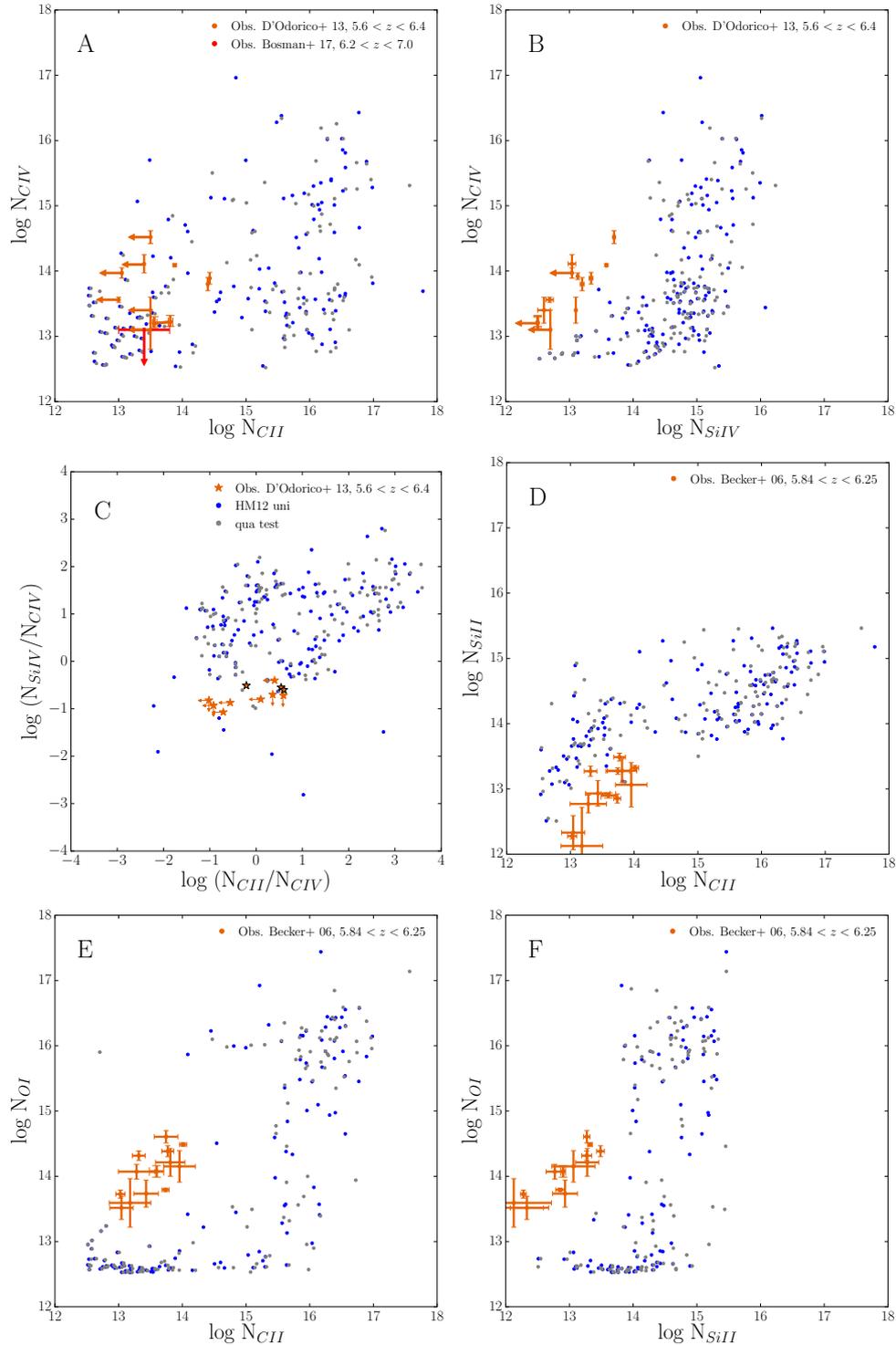


Figure A.4 Column density relationships among metal absorption lines at  $z = 6$  with a modification of the high energy end in the uniform HM12 UVB spectrum, as a result of a strong contribution of quasars at this redshift (**qua test**). The simulation run used is Ch 18 512 MDW. Panels A, B and C compare with systems observed by D'Odorico et al. (2013) in the redshift range of (5.6–6.4) and D, E and F with absorbers detected by Becker et al. (2006) in the redshift range of (5.84–6.25), both in orange. An additional data point from Bosman et al. (2017) has been introduced in panel A, in red. The upper limit in CIV is found in the redshift range of (6.2–7.0). From the synthetic spectra of each ion, the column densities are calculated and displayed. The simulated data shown are above the threshold noise at  $\log N_{th} = 12.5$  ( $\text{cm}^{-2}$ ). In panel C, the observations by D'Odorico et al. (2013) produced 3 systems with column densities reported (represented by the black stars) and 8 with upper limits in  $N_{CII}$  and/or  $N_{SiIV}$ , represented by arrows.

## List of publications

**García** L. A., Tescari E., Ryan-Weber E. V., Wyithe J. S. B. 2017, “Simulated metal and HI absorption lines at the conclusion of Reionization”, *Monthly Notices of the Royal Astronomical Society*, vol. 470, issue 2, pp. 2494-2509.

**García** L. A., Tescari E., Ryan-Weber E. V., Wyithe J. S. B. 2017, “Theoretical study of an LAE–CIV absorption pair at  $z = 5.7$ ” in *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 469, issue 1, pp. L53-L57.