Effect of Inter-Galactic Medium on the Observability of Lyman Alpha Emitters around the Epoch of Reionization

Akila Jeeson Daniel

München 2012

Effect of Inter-Galactic Medium on the Observability of Lyman Alpha Emitters around the Epoch of Reionization

Akila Jeeson Daniel

Dissertation
an der Fakultat fur Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Akila Jeeson Daniel
aus Ottapalam, Kerala, India

München, den 13 Januar 2012

Erstgutachter: Prof. Dr. Simon White
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 2 März 2012

Zusammenfassung

Die Reionisierungsepoche („epoch of reionization“, EOR) stellt den Übergang des Universums von einem hauptsächlich neutralen zu einem hoch ionisierten Zustand dar. Lyman Alpha Emitters (LAEs) sind Galaxien mit einer starken Lyα-Emissionslinie aufgrund derer sie bis zu hohen Rotverschiebungen detektiert werden können. Lyα-Strahlung wird sehr leicht an neutralem Gas gestreut und kann daher zur Untersuchung des Ionisationszustands der Region um diese Galaxien und damit der Reionisierungsepoche verwendet werden. Diese Arbeit berichtet über die laufende Forschung zum Verständnis der verschiedenen Aspekte, die für eine präzise Modellierung einer statistisch signifikanten Stichprobe von LAEs bei $z = 7.7$ notwendig sind. Der Vergleich der Stichprobe mit Beobachtungen erlaubt es schließlich die Eigenschaften der EOR einzuschränken.

Im ersten Teil dieser Arbeit wird der Einfluß des intergalaktischen Mediums (IGM) in der Umgebung der Quelle auf die Beobachtbarkeit der LAEs während der EOR studiert. Aus hydrodynamischen Simulationen entnehmen wir Testvolumina des IGM um die Galaxien. Mit Hilfe des 3D-Strahlungstransportcodes CRASHα simulieren wir die durch die Galaxie verursachte Ionisationsregion und den Lyα-Strahlungstransport durch das sich verändernde IGM im Testvolumen. Wir beobachten, dass Inhomogenitäten im IGM zu Strukturen im Oberflächenhelligkeitsprofil der simulierten LAEs führen, besonders im Fall von sehr wenig ionisiertem Gas. Dadurch ergibt sich für eine Quelle eine signifikante Verteilung der Lyα Luminosität in Abhängigkeit der Blickrichtung des Beobachters, besonders für Objekte mit Oberflächenhelligkeiten am Limit der Detektierbarkeit. In Blickrichtungen durch Voids ist die Wahrscheinlichkeit ausströhmender Lyα Photonen höher als in Richtungen durch Filamente. Effekte wie Rotverschiebung der Lyα Photonen durch Interaktion mit dem interstellaren Medium und der höhere Ionisationsgrad des IGM durch ionisierende Hintergrundprozesse und/oder ein Clustern der Quellen erleichtern zusätzlich das Entkommen der Lyα Photonen, was die Beobachtbarkeit von LAEs verbessert. In dieser Arbeit wird gezeigt, dass dadurch die Eigenschaften des IGM in der Nähe der Quelle entscheidenden Einfluss auf die Beobachtbarkeit von LAEs haben und dass deshalb deren angemessene Modellierung und Verständnis unverzichtbar sind um LAE Beobachtungen zur Untersuchung der EOR zu nutzen.

Um den IGM Ionisationsgrad eines Gebiets abzuschätzen muss man die Reionisationsprozesse für ein großes, repräsentatives Volumen des Universums angemessen behandeln. Der „clumping factor“ ist ein Parameter der die echte Anzahl von Rekombinationen einer ionisierten Spezies in einem Zellvolumen mit einer mittleren Gasdichte abschätzt. Er hilft dabei die notwendige Genauigkeit der Ionisationsgeschichte auch bei Simulationen mit niedriger räumlicher Auflösung zu erreichen. Im zweiten Teil der Arbeit untersuchen wir die Entwicklung des clumping factors mit der Rotverschiebung für neutrales und ionisiertes H und He in einer kleinen, aber hochaufgelösten Simulation der EOR. Alle neutralen Spezies zeigen dabei ein ähnliches Verhalten, ebenso die ionisierten. Die Werte und Rotverschiebungsentwicklung des clumping factors hängen dabei von der Definition des Dichteschwellwerts, der Gitterauflösung und der physikalischen Größe des simulierten Volumens ab. Haupteinfluss und Ursache für die Verteilung der clumping factor Werte zwischen den unterschiedlichen Gebieten ist aber die mittlere Dichte.

Abstract

The Epoch of Reionization (EOR) marks the change of the Universe from a mainly neutral to a highly ionized state. Lyman Alpha Emitters (LAEs) are galaxies with a strong Lyα emission line which can be exploited to detect them at high redshifts. Lyα photons, being very sensitive to scattering by neutral gas, can be used as a probe of the ionization structure around these galaxies and thus the EOR. This thesis reports on the ongoing investigation aimed at understanding the different aspects important for the accurate modelling of a statistically significant sample of LAEs at $z = 7.7$. This sample can eventually be compared with observations to constrain EOR.

The first part of this thesis aims to study the effect of the Inter-Galactic Medium (IGM), close to the source, on the observability of LAEs during the EOR. From hydrodynamic simulations, we extract cubes of IGM around galaxies. Using the 3D radiative transfer code CRASHα, we simulate the ionization region produced by the galaxy and the Lyα radiative transfer through the evolving IGM cube. We find that inhomogeneities in the IGM lead to structure in the surface brightness profiles of simulated LAEs, especially for a highly neutral IGM. This leads to a significant scatter in the Lyα luminosities observed from the same object along different lines-of-sight, especially for low surface brightness thresholds. There is a higher probability of Lyα photon escape for lines-of-sight passing through the voids than through the filaments. In addition, effects like redshifting of Lyα photons due to processing in the Inter-Stellar Medium and higher ionization fraction in the IGM due to ionizing background and/or source clustering lead to easier escape of Lyα photons, improving the observability of LAEs. Therefore, this thesis shows that the properties of the IGM close to the source play a significant role in determining the observability of LAEs and thus their proper modelling and understanding is crucial for using LAE observations to study EOR.

Estimating the IGM ionization level of a region needs proper treatment of the reionization process for a large representative volume of the Universe. The clumping factor, a parameter estimating the true number of recombinations of an ionized species in a cell volume given a mean gas density, aids in achieving the required accuracy for the reionization history even in simulations with low spatial resolution. In the second part of this thesis, we study the redshift evolution of clumping factors of different neutral/ionized species of H and He in a small but very high resolution simulation of EOR. All neutral species show similar behaviour and the same is true for the ionized species. The values and redshift trends of clumping factors depend on the definition, overdensity range, grid resolution and box size of the simulations. But the main factor which leads to a spread in the clumping factor values between different regions is the mean gas density.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

" Our whole Universe was in a hot dense state,
Then nearly fourteen billion years ago expansion started Wait
The Earth began to cool,
The autotrophs began to drool,
Neanderthals developed tools,
We built a wall (we built the pyramids),
Math, science, history, unraveling the mysteries,
That all started with the Big Bang!"

- The Big Bang Theory theme song by BNL.

The Hot Big Bang is the currently accepted model for the history of the Universe (e.g. Padmanabhan, 1993; Mo, van den Bosch, & White, 2010). It is derived from the General Theory of Relativity under the assumption that the Universe is homogenous and isotropic. The homogeneity of the Universe was observed in the galaxy counts in deep surveys, while the isotropy was seen in Cosmic Microwave Background (CMB) observations discovered by Penzias and Wilson in 1965. In fact, CMB is one of the predictions of this model (Gamow, 1948; Alpher & Herman, 1948) along with the expansion of the Universe (Hubble, 1929) and the primordial nucleosynthesis (Alpher et al., 1948).

Lambda-Cold Dark Matter (ΛCDM) is the currently accepted cosmology supported, for example, by measurements of the CMB (e.g. Miller et al., 1999; Komatsu et al., 2011), distances using Cepheid variables (e.g. Freedman et al., 2001), Type-I supernovae (e.g. Riess et al., 2004; Astier et al., 2006) and galaxy clustering (e.g. Percival et al., 2007). The energy content of the Universe is dominated by Dark Energy (DE; also known as Λ) which makes for $\sim 73\%$; matter consists of about $\sim 27\%$, of which Cold Dark Matter (CDM; shortened to DM) accounts for $\sim 22.5\%$ and baryons account for the remaining 4.5% (Larson et al., 2011). DE is a form of negative energy which leads to the accelerated expansion of the Universe as seen in Type-I supernovae data (e.g. Riess et al., 1998; Perlmutter et al., 1999).

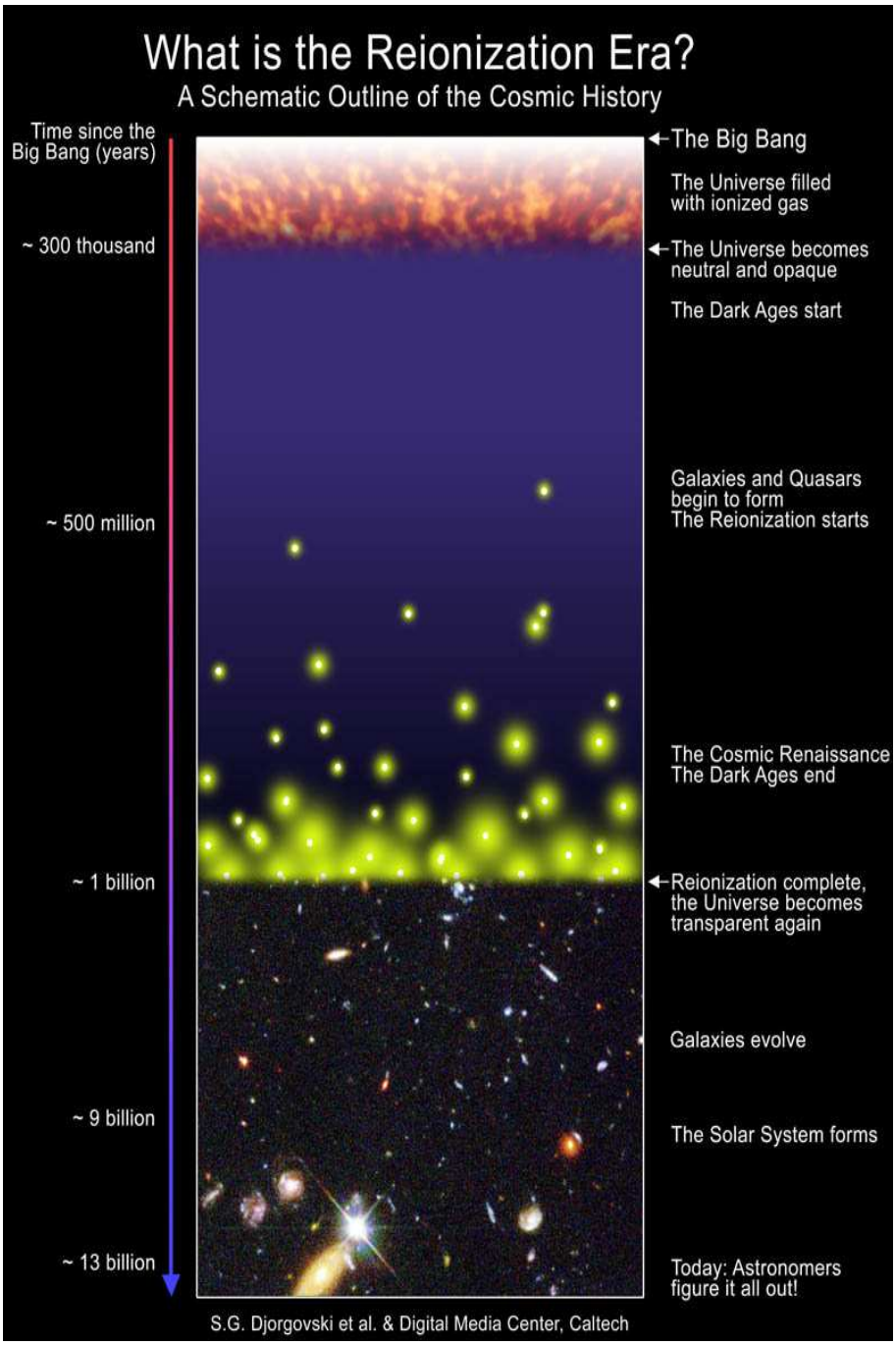

Figure 1.1: Time line of the Universe

Its exact nature is currently unknown (refer to Frieman, Turner, & Huterer, 2008, for a review). DM is the dominant form of matter ($\sim 80\%$); it can only be observed through its gravitational force and it was first inferred by Zwicky (1937). Even though currently ($z \lesssim 0.5$) the dominant component in the Universe is DE, earlier redshifts were dominated by matter ($0.5 \lesssim z \lesssim 3000$) and radiation ($z \gtrsim 3000$).

1.1 History of the Universe

The time line of the history of the Universe is shown in Figure 1.1. The Universe formed instantaneously about 13.8 billion years ago (e.g. Larson et al., 2011) in a singularity of infinite temperature and density which then cooled as the Universe expanded adiabatically. Approximately 10^{-35} seconds after the Big Bang, the Universe went through a rapid phase of expansion, known as the Inflation (Guth, 1981), during which the Universe expanded by about 100 e-folds (i.e. e^{100}). The initial microscopic quantum fluctuations were blown up into macroscopic scale which became the seeds for structure formation in an otherwise homogeneous and isotropic Universe.

Then the Universe continued to expand adiabatically and cooled. Protons, neutrons and electrons were formed when the temperature was $T \sim 3 \times 10^{12}$ K (corresponding to $t \sim 10^{-5}$ seconds). At this stage, the neutron and the protons were in equilibrium (e.g. Dodelson, 2003) through weak interaction conversions to each other till $T \sim 10^{11}$ K (at $t \sim 3$ seconds), when the proton to neutron conversion stopped and the neutron-to-proton ratio 'freezed in' at n/p=1/6. Below this temperature, the neutrons decayed into protons through beta-decay. To form the nuclei of all elements of atomic number of He and higher, the reaction chain needed Deuterium. Being in the radiation dominated era, there existed 1.6×10^9 photons per atom which lead to the high energy photons (in the tail of the Maxwell-Boltzmann distribution) destroying any Deuterium nuclei that was formed, leading to a 'Deuterium bottleneck'. The Universe had to cool further ($T = 10^9$ K at $t \sim 2$ minutes) to form Deuterium which then fused to form heavier elements up to Lithium and Beryllium (e.g. Peacock, 1999). By then, the neutron-to-proton 'froze out' at n/p$\sim$ 1/7) due to beta decay which constrained the production of He in the Universe. The elemental abundance from the 'Big Bang Nucleosynthesis' is H (75%) and ^{4}He (24.8%) with trace amounts of the rest.

As the Universe cooled further, the electrons and the nuclei started recombining to form neutral atoms. But as in the case of Deuterium bottleneck, the Universe had to cool to about $T = 3000$ K (at $t \sim 3 \times 10^5$ years; $z \sim 1100$) for the atoms to stay neutral without getting re-ionized by high energy photons. This is known as the Epoch of Recombination (see review by Sunyaev & Chluba, 2008). The photons decoupled from baryons and started streaming freely through the Universe cooling at a rate of $T \propto (1 + z)$ to about 3 K at $z = 0$. This relic radiation is known as the CMB, which we already mentioned above. The era after the Epoch of Recombination is known as the Dark Ages as there are no photons other than the gradually cooling CMB while the gas in the Universe is neutral and opaque to H ionizing photons.

The structure in Universe formed out of the density fluctuations which originated during the Inflation era. As DM does not interact with radiation, it started to collapse due to gravity, virialised and formed DM haloes. Structures in the Universe were formed "bottom up" by small objects collapsing first and then merging together to form bigger structures, also known as hierarchical structure formation. Baryons fell into these DM potential wells and cooled to eventually make stars and galaxies. Only in about a few hundred million years after the Big Bang ($z = 20 - 30$), the first stars appeared in the Universe, producing H ionizing photons and ending the dark ages (refer to Ciardi & Ferrara, 2005, for a review). The first stars (PopIII) formed out the metal-free primordial gas cooled with the aid of molecular hydrogen (e.g. Bromm, Coppi, & Larson, 2002; Abel, Bryan, & Norman, 2002). They produced metals which helped the gas in DM haloes to cool further and create metal-polluted PopII/I stars. All these different populations of stars generated ionizing photons carving ionized H bubbles in the Inter-Galactic Medium (IGM) which slowly grew in size, eventually to merge and make the Universe transparent again ($z \sim 6$) (e.g. Fan, Carilli, & Keating, 2006). This era is known as the Epoch of Reionization (EOR).

The galaxies continued to evolve till the present day, $z = 0$, through gas accretion, mergers, star formation and feedback - mechanical (e.g. winds, jets), chemical (e.g. metals, dust) and radiative (e.g.UV, X-rays) (e.g. Ciardi & Ferrara, 2005). Also, supermassive blackholes (Active Galactic Nuclei/Quasars) became more common in galaxies producing large amounts of high energy photons which lead the reionization of He in the Universe ($z \sim 3$) (e.g. Jakobsen et al., 1994; Ciardi & Ferrara, 2005). The current epoch is important for us as it is the bench mark to which we compare the rest of the Universe.

1.2 The Inter-Galactic Medium

IGM is commonly referred to as the gas present in the filaments connecting the galaxies and in the voids between them (refer to Meiksin, 2009, and references there in for an excellent review). Since the mean gas density of the Universe changes at each redshift due to the cosmological expansion, gas densities are generally measured as overdensities compared to the mean baryon density at that redshift. From the linear theory of structure formation, matter (mainly DM) is considered to be virialised at densities of approximately 200 times the mean matter density of the Universe at that redshift. Since gas follows DM, we can assume the same overdensity of 200 for virialised gas. The IGM has a typical overdensity the range $\sim 0.1 - 100$. Figure 1.2 (from Zhang et al. 1998) shows the IGM gas in a simulation box of 9.6 comoving Mpc size at $z = 3$. Plotted are the isodensity contours for overdensities of 0.1, 0.3, 0.5, 1, 5 and 10. As we can see, the voids which occupy most of the volume have overdensities < 1, while the filaments are overdensities > 1. The galaxies occupy the regions where the filaments meet. Figure 1.3 shows the evolution of the IGM overdensities with redshift from the same simulation in Zhang et al. (1998). Also over-plotted are the peculiar velocities. We can see that as the redshift decreases, the range of overdensities increases, with the voids becoming less dense and the filaments becoming denser. The peculiar velocities show that the gas flows from the voids into the filaments.

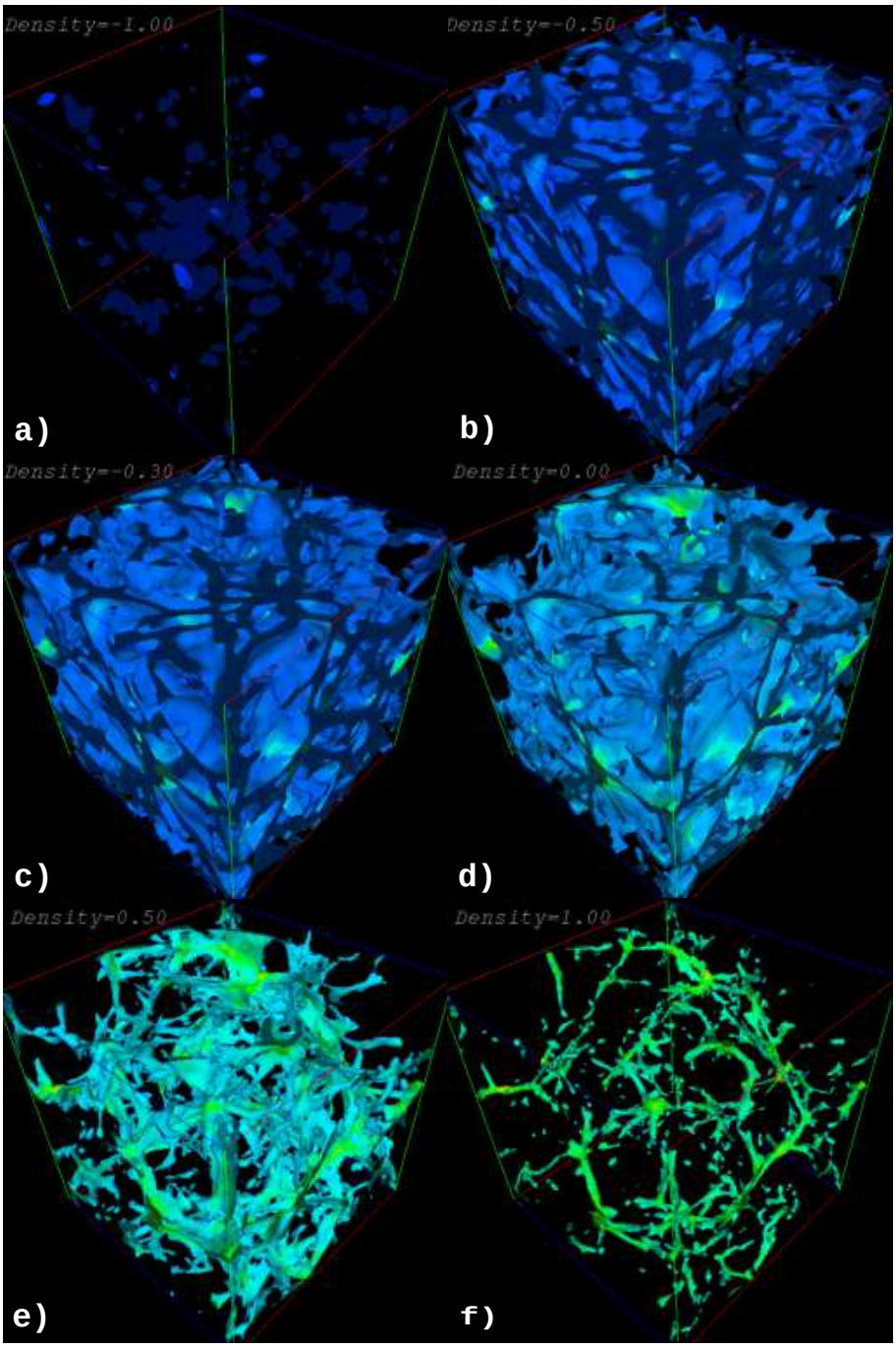

Figure 1.2: From Zhang et al. (1998) - Isodensity contour surfaces of baryon overdensity at z = 3 for a 9.6 comoving Mpc box. The contour levels are for log of the baryon overdensities $\log_{10}(\rho_{\rm b}/\langle\rho_{\rm b}\rangle) = -1.0, -0.5, -0.3, 0.0, 0.5$ and 1.0.

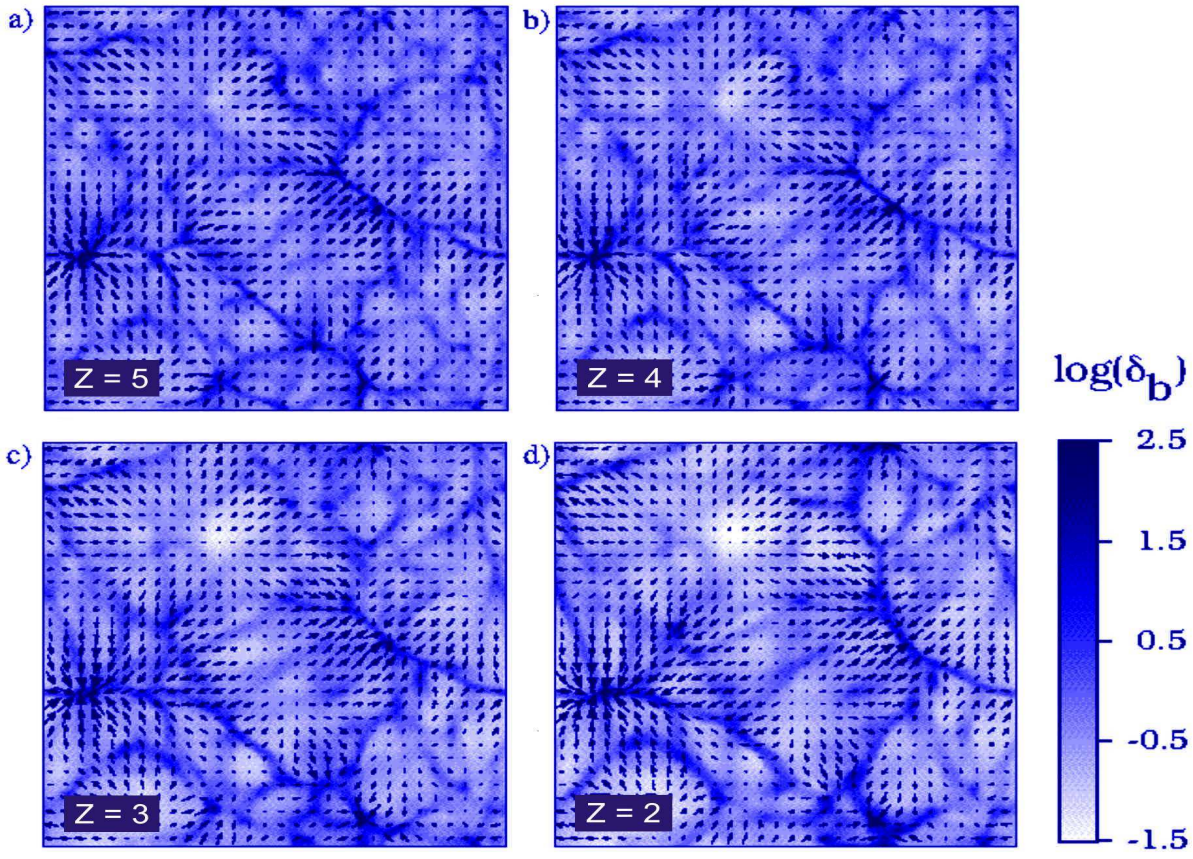

Figure 1.3: From Zhang et al. (1998) - The evolution of baryon overdensity distribution. The baryon overdensities in the box are shown at z = 2, 3, 4, and 5. Superimposed as arrows is the peculiar velocity field.

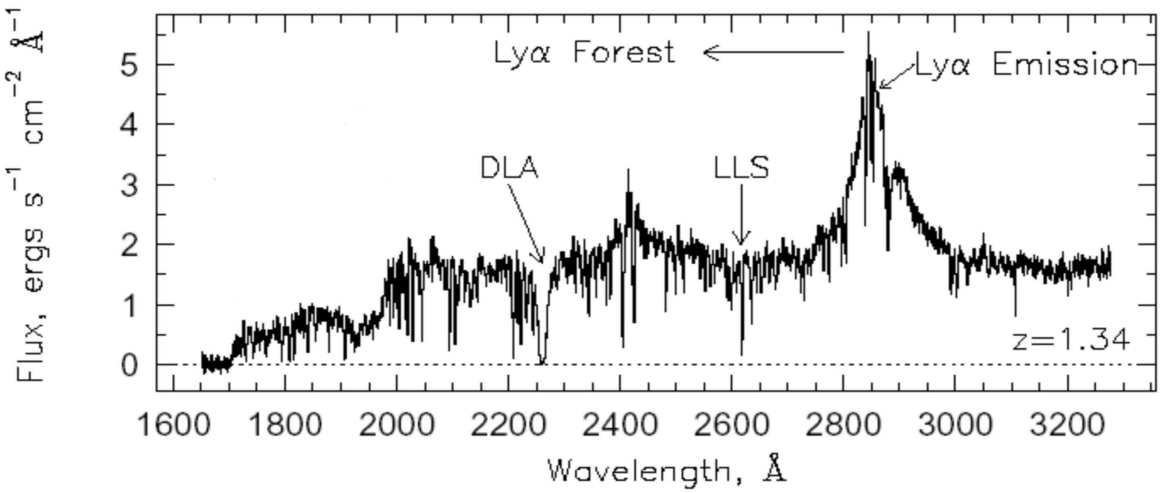

Figure 1.4: The typical spectra of a QSO at $z = 1.34$ with the Lyα emission and absorption lines. From http://www.hs.uni-hamburg.de/EN/For/Exg/Igm/index.html.

The first observations of the IGM came when Gunn & Peterson (1965) observed a slight decrement in flux blue-ward of the Lyman Alpha (Lyα) line (refer to Section 1.4 for details on this line) in the spectra of quasi-stellar objects (QSOs). This was attributed to the neutral (HI) absorption lines due to gas in the IGM.

As a photon of higher wavelength redshifts to Lyα resonance wavelength of 1215.67 Å, the neutral gas present at that redshift along the line-of-sight scatters the Lyα photons away from the line-of-sight to the observer. Therefore, the absorption lines trace the neutral gas along the line-of-sight from the QSO to the observer. Since then, QSO spectra have been the main method of studying the IGM. Figure 1.4 shows the spectrum of a QSO at $z = 1.34$[1]. As we can see, the strong Lyα emission from the QSO is visible in the spectrum. Towards shorter wavelengths, we can see the absorption lines of the Lyα forest, Lyman Limit System (LLS) and Damped Lyα Absorber (DLA). Objects in the IGM are divided according to the column density of neutral gas $N_{\rm HI}$ calculated from the absorption line and they are - Lyα forest ($N_{\rm HI} \leq 10^{17}$ cm^{-2}), LLS (10^{17-19} cm^{-2}), Super Lyman Limit Systems (Super LLS; $10^{19} - 2 \times 10^{20}$ cm^{-2}) and Damped Lyα Absorbers ($\geq 2 \times 10^{20}$ cm^{-2}). The distinction between these objects is not strictly exclusive but helps with their understanding. Also their properties are not well constrained.

The Lyα forest is believed to consist of long filamentary gas clouds (the cosmic web) of size in the range $\sim 15 - 1000$ kpcs, with mean gas particle densities of $10^{-8} - 10^{-3}$ cm^{-3}, temperatures $5000 - 50000$ K and a relatively low metallicity. LLS and Super LLS have a number density of $10^{-3} - 10^{-2}$ cm^{-2} and temperature of $10000 - 30000$ K. Super LLS have higher metallicities than LLS. DLAs are thought to be high density objects (possibly proto-galaxies) of sizes of $15-20$ kpc, gas number density of $10-0.01$ cm^{-3} at temperatures

[1]From http://www.hs.uni-hamburg.de/EN/For/Exg/Igm/index.html.

of $100 - 10000$ K and relatively high metallicity.

The IGM is the reservoir of gas which fuels the star formation in galaxies through hot and cold accretion (e.g van de Voort et al., 2011). It also gets affected by galaxies through winds, AGN jets, metals and ionizing photons. These interactions affect the thermal history of the IGM and control future star formation in galaxies thus playing an important role in shaping the star formation history of the Universe. The most important epoch in the thermal history of the IGM known as the 'Epoch of Reionization' is discussed in the next section.

1.3 The Epoch of Reionization

The Epoch of Reionization is the era in the history of the Universe when the IGM went from a highly neutral state to a mostly ionized one (for reviews, refer to Barkana & Loeb, 2001; Ciardi & Ferrara, 2005; Fan, Carilli, & Keating, 2006; Furlanetto, Oh, & Briggs, 2006; Barkana & Loeb, 2007; Meiksin, 2009; Morales & Wyithe, 2010, and the references there in). Both H and He, the two main elements present in the gas, reionize during the course of time but due to their different ionization energies and recombination coefficients, they have different reionization histories. In this thesis, we focus on the reionization of H.

1.3.1 Observational Evidence

The observational evidence for EOR comes from mainly two directions. The first evidence is the Gunn-Peterson trough seen in QSO spectra (e.g. Becker et al., 2001; Fan et al., 2006). Figure 1.5 shows the spectra of 19 quasars from the Sloan Digital Sky Survey, taken from Fan et al. (2006). We can see that as redshift increases, the flux on the blue side of the QSO Lyα line decreases. The higher the IGM neutral fraction is, the larger is the absorption along a line-of-sight. But the Lyα absorption lines in QSO spectra saturate for values of H neutral fraction of $x_{\rm HI} \sim 10^{-4}$. We need other forms of evidence to probe higher neutral fractions.

The QSO data point towards a reionization that finished by $z > 6$. Recent modelling of high redshift QSO data by Mesinger (2010) found though that an upper limit of $x_{\rm HI} \leq 0.1$ cannot be ruled out at $z = 6$ due to the patchy nature of reionization (refer to Section 4) and the biased distribution of gas around quasars. Recent detection of a QSO at $z \sim 7$ (Mortlock et al., 2011) has given a lower limit of $x_{\rm HI} > 0.1$ (Bolton et al., 2011).

The other major evidence for EOR comes from the CMB observations. Free electrons produced during reionization interact with the CMB photons through Thompson's scattering which softens the CMB temperature anisotropy signal. And at large scales ($> 10°$), the quadrapole anisotropy signal in the CMB generates a polarization signal. The free electrons also add small scale ($< 0.1°$) anisotropies to the CMB signal through effects such as Sunyaev-Zeldovich Effect (Sunyaev & Zeldovich, 1980), Ostriker-Vishniac effect (Ostriker & Vishniac, 1986) and due to patchy reionization (e.g. Santos et al., 2003). The CMB measurements give an the integrated signal with the recent constraints from WMAP7 data

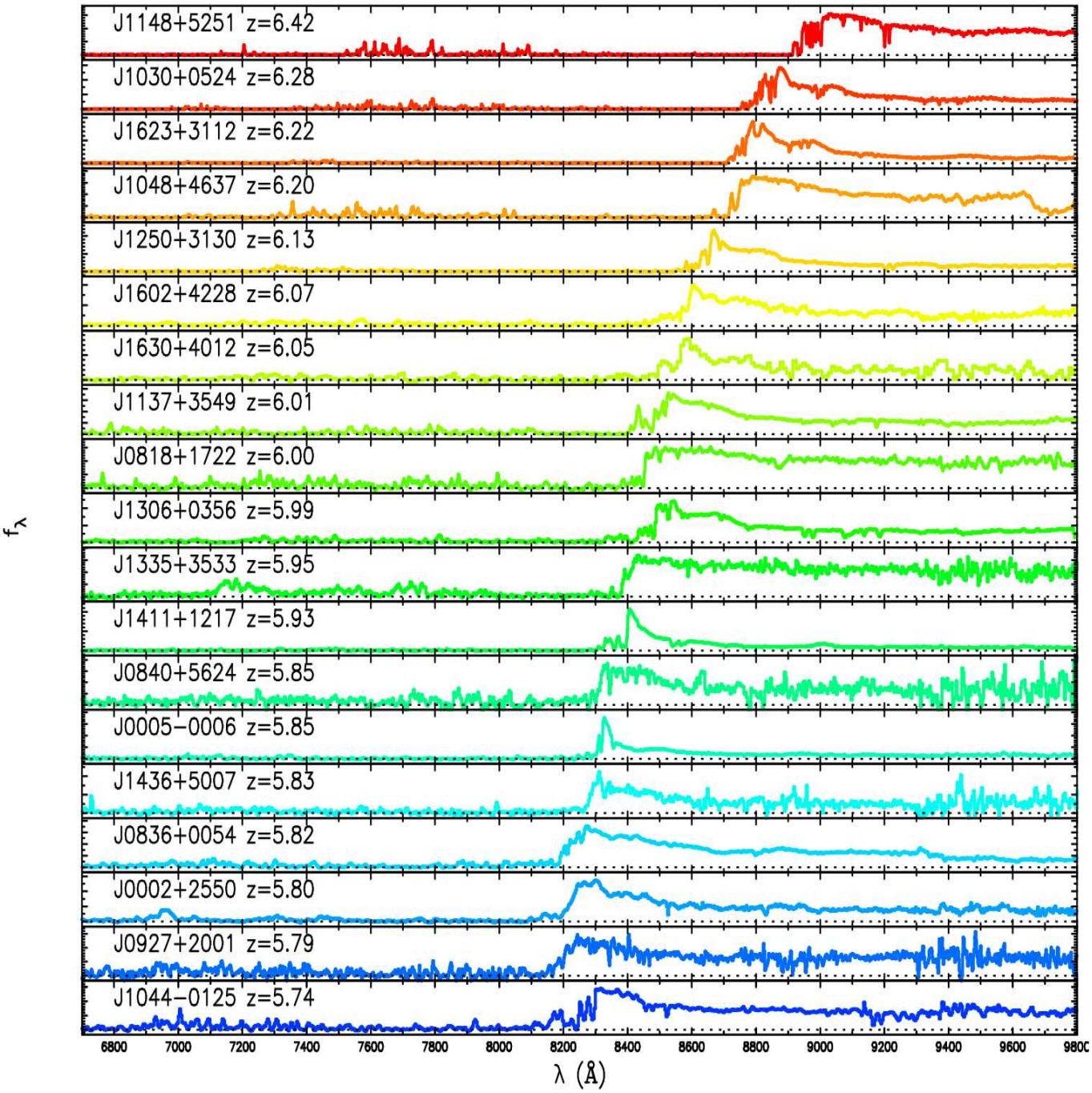

Figure 1.5: From Fan et al. (2006) - Spectra of 19 quasars with redshifts $5.74 < z < 6.42$ from the Sloan Digital Sky Survey.

(Larson et al., 2011) point it to be equivalent to one given by an instantaneous reionization at $z \sim 10.5$.

Other probes of the IGM ionization fraction include Gamma Ray Burst (GRB) spectra, 21 cm observations, Lyman Alpha Emitters (LAEs), among others. GRB afterglow spectra can be used similarly to QSO spectra to study the Gunn-Peterson troughs. Totani et al. (2006) used the GRB 050904 spectra at $z \sim 6.3$ to calculate $x_{\rm HI} < 0.6$. But the difficulty in predicting the location and time of a GRB and the impossibility of repeating the observation, makes it a challenge to study the EOR. Nevertheless GRBs have been observed at redshifts higher than QSOs. The furthest detection of a GRB is at $z \sim 9.4$ (Cucchiara et al., 2011), compared to a QSO at $z \sim 7$ (Mortlock et al., 2011), providing an important tool for studying very high redshifts.

Detection of the 21 cm hyperfine transition from the neutral H gas is one of the main probes of EOR (refer to Morales & Wyithe, 2010; Pritchard & Loeb, 2011, for recent reviews). Intensity mapping and power spectrum analysis of 21 cm observations with the current and upcoming instruments, namely, GMRT[2], PAPER[3], LOFAR[4], MWA[5] and SKA[6], would give an unprecedented view on the ionization history of the IGM. LAEs provide another interesting view on EOR. This is the main topic of this thesis and will be explained in detail in Section 1.4.

1.3.2 Sources of Ionizing Photons

A large number of sources of ionizing photons have been proposed such as PopIII stars, QSOs, galaxies with PopII/I stars, mini-quasars, shock heating and decaying DM particles. Due to their different spectra, lifetimes and spatial distribution, they determine the evolution and topology of EOR. The metal-free PopIII stars are generally massive (~ 100 $M_\odot$ but also refer to recent work by Clark et al., 2011; Greif et al., 2011) and emit at least 50% more ionizing photons than metal enriched stars (e.g. Tumlinson & Shull, 2000; Bromm, Kudritzki, & Loeb, 2001; Bromm & Yoshida, 2011). But their formation in the Universe is regulated by the formation of metals. Thus they are thought to be too few to be the dominant sources for reionization (e.g. Meiksin, 2005). As redshift decreases, other sources of ionizing photons such as galaxies with metal enriched PopII/I stars, QSOs, etc appear in the Universe. QSOs have a harder spectra than galaxies with metal enriched stars and carve out large ionized regions. But QSOs are found only in massive galaxies making them rare at high redshifts. Thus, they are not thought to dominate the H reionization process.

Galaxies with PopII/I stars have a softer spectra but are more numerous in numbers at high redshift than QSOs. Also, they are more evenly distributed in the Universe than QSOs. This leads to a uniform reionization of the region but at a slower rate than due to QSOs. Although stars within galaxies are believed to produce enough ionizing photons to

[2]http://www.gmrt.ncra.tifr.res.in/
[3]http://eor.berkeley.edu/
[4]http://www.lofar.org/
[5]http://www.haystack.mit.edu/ast/arrays/mwa/
[6]http://www.skatelescope.org/

reionize the Universe, whether full reionization can be reached or not depends crucially on the fraction of ionizing photons escaping from the galaxy, $f_{\rm esc}$ (e.g. Meiksin, 2009). A lot of different methods have been employed to estimate/calculate $f_{\rm esc}$ such as observations of Lyman Break Galaxies at low redshifts ($z \sim 3$; e.g. Steidel, Pettini, & Adelberger, 2001; Inoue et al., 2005; Shapley et al., 2006; Iwata et al., 2009), semi-analytic models (e.g. Dove & Shull, 1994; Ricotti & Shull, 2000; Wood & Loeb, 2000; Dove, Shull, & Ferrara, 2000; Ciardi, Bianchi, & Ferrara, 2002; Fujita et al., 2003; Fernandez & Shull, 2011) and radiative transfer simulations (e.g. Razoumov & Sommer-Larsen, 2006; Gnedin, Kravtsov, & Chen, 2008; Wise & Cen, 2009; Razoumov & Sommer-Larsen, 2010; Yajima, Choi, & Nagamine, 2011; Paardekooper et al., 2011). But these methods have provided $f_{\rm esc}$ ranging from $0-1$ with no clear consensus on the trends with halo mass, redshift, age, etc. It remains a very uncertain parameter due to its dependency on a large number of factors.

Some of the alternative class of sources are mini-quasars, accretion shocks and decaying DM particles. Mini-quasars are proposed to be intermediate mass black-holes ($200 - 1000$ $M_\odot$) grown by accretion on to seed black-holes from collapsed PopIII stars (e.g. Madau et al., 2004). These sources could reionize the Universe by $z \simeq 15$ but would over produce the x-ray background. Accretion shock heating of the IGM was proposed as an alternate source of ionizing photons (e.g. Furlanetto & Loeb, 2004; Miniati et al., 2004; Dopita et al., 2011; Wyithe, Mould, & Loeb, 2011). Even though it is not expected to dominate H reionization, it could affect the thermal history of He and luminosity weighted source clustering. Decaying DM particles have been proposed as an early source of ionizing photons at $z \leq 100$ (e.g. Hansen & Haiman, 2004) but the uncertainty regarding the properties of dark matter particles makes it difficult to model.

1.3.3 Modelling of the Reionization Process

The main stages of the reionization process are namely pre-overlap, overlap and post-overlap (Gnedin, 2000). In the pre-overlap stage, the Universe is mostly neutral with small ionized bubbles being formed by early sources. The size of these bubbles and their clustering properties depend on the type of source. The ionization structure in the IGM is very patchy at this stage, with the ionization level in a location depending on the strength and the distance from the nearest source. As time evolves, the ionized bubbles become larger and eventually start merging. This stage is known as overlap. Most of the Universe is ionized and the ionized regions receive photons from a large number of sources. The ionization level through out the Universe is now more uniform, with all the low density regions being ionized and only high density regions far away from sources being (partially) neutral due to self shielding/high recombination rates (inside-out/-middle reionization scenario; e.g. Gnedin, 2000; Ciardi, Stoehr, & White, 2003; Iliev et al., 2006; Trac & Cen, 2007; Choudhury, Haehnelt, & Regan, 2009; Finlator et al., 2009; Petkova & Springel, 2011). Beyond this stage the reionization proceeds at a much slower pace with the ionization front slowly encroaching into the LLSs and DLAs. This is known as the post-overlap phase and continues till the present day with the collapsed systems still retaining some neutral gas.

Due to the complex nature of reionization, semi-analytic models and numerical simulations provide an important tool for the detailed understanding of EOR. For proper modelling of EOR, we need the correct distribution of sources and their properties as well as radiative transfer of ionizing photons from them. Since some galactic properties determining the production of ionizing photons (e.g. star formation efficiency, QSO formation at high redshifts, stellar initial mass function, dust formation and properties, feedback such as winds, jets, etc.) are still very unclear, together with $f_{\rm esc}$ from the ISM of the galaxies to the IGM, this leads to degeneracies in the models (refer to the reviews quoted in Section 1.3). But efforts are being taken to constrain the EOR models using a variety of observations (e.g. Choudhury & Ferrara, 2006).

Numerical simulations is one of the main methods used to study the process of EOR (refer to Trac & Gnedin, 2009, for a recent review). A large number of RT codes have been developed to investigate this problem (e.g. Gnedin, 2000; Razoumov et al., 2002; Ciardi, Stoehr, & White, 2003; Sokasian, Abel, & Hernquist, 2001; Mellema et al., 2006; McQuinn et al., 2007; Semelin, Combes, & Baek, 2007; Trac & Cen, 2007; Altay, Croft, & Pelupessy, 2008; Aubert & Teyssier, 2008; Finlator et al., 2009; Petkova & Springel, 2009; Partl et al., 2011). The main technical challenge in simulating EOR is to achieve high resolution in simulations with large box sizes.

Due to the patchy nature of reionization, simulations of large comoving volumes are needed to capture the topology of reionization. Also, as reionization proceeds the bubble sizes grow up to tens of comoving Mpcs (Furlanetto, Oh, & Briggs, 2006). Therefore, to simulate a representative volume, a comoving box size of ~ 100 Mpcs is needed (e.g. Barkana & Loeb, 2004; Iliev et al., 2006). On the other hand, high spatial resolution is needed to resolve the small atomically-cooled galaxies with dark matter halo masses of $\sim 10^8$ $M_\odot$ and comoving length scales of a few tens of kpcs at $z \sim 7 - 10$, which are thought to dominate the ionizing photon budget (e.g. Bouwens et al., 2008; Finkelstein et al., 2011). In principle, also LLSs and DLAs should be resolved to track the reionization history during the post-overlap phase. But they are of proper sizes of a few kpc and a thousand times the mean density of the Universe at $z \sim 4$ (Kohler & Gnedin, 2007), which is beyond the current computational limit for 100 Mpc boxes.

Many of the RT methods used to simulate EOR have been tested in Iliev et al. (2006) and Iliev et al. (2009). Current state-of-the-art simulations have achieved the required box size 100 h^{-1}Mpc and mass resolution 10^8 $M_\odot$ to resolve the source range (e.g. Iliev et al., 2008; Trac & Cen, 2008; Shin, Trac, & Cen, 2008). Also small box simulations have been performed to study the ionization of small scale structure in the IGM (e.g. Iliev, Scannapieco, & Shapiro, 2005; Kohler & Gnedin, 2007, and Chapter 4 in this thesis). The information from these small simulations can be incorporated into larger simulations to mimic the effect of unresolved small scale structures (e.g Gnedin & Ostriker, 1997; Iliev, Scannapieco, & Shapiro, 2005; Trac & Cen, 2007; Kohler, Gnedin, & Hamilton, 2007; Pawlik, Schaye, & van Scherpenzeel, 2009; Raičević & Theuns, 2010, and Chapter 4 in this thesis).

Semi analytic methods to model the EOR are useful to understand its general trends and to study the parameter space in a faster way (e.g. Zahn et al., 2005; Mesinger &

Furlanetto, 2007; Geil & Wyithe, 2008; Alvarez et al., 2009; Choudhury, Haehnelt, & Regan, 2009; Thomas et al., 2009). They generally use halo mass and distribution functions from simulations and position spherical bubbles around the sources, evolving them in time depending on the source luminosity. Recent work has shown that some of the semi-analytic methods could give results comparable to numerical simulations in predicting 21 cm signal which mainly focuses on large scales (Zahn et al., 2011).

Thus we can see that the different aspects of EOR (both models and observations) are highly uncertain, and future observations (e.g. 21 cm observations) and theoretical (e.g. as mentioned in Chapter 4) work are expected to give better constaints on the reionization history. Since this thesis is motivated by the use of LAEs for the detection of EOR, we will discuss them in more detail in the following.

1.4 Lyman Alpha Emitters

A Lyα line is the transition from levels n=2 to n=1 (ground state) in an excited atom, where n is the principal quantum number. This line was discovered by Theodore Lyman in 1906 while studying the UV spectrum of electrically excited hydrogen atoms and thus the recombination of electrons in an atom to the ground state is named after him. Hydrogen being the most abundant atom in the Universe, Lyα line of Hydrogen at 1215.67 Å rest wavelength is a very important line for Astrophysics. Due to the high recombination coefficient, a large fraction of the H recombinations in a volume of ionized gas, i.e. 45-70 % at 10^4 K, gives rise to this line (e.g. Osterbrock, 1989).

A LAE is defined as an object which emits a strong Lyα line with a narrow band excess (see Ouchi (2008)). A large number of different objects can fall under this criterion, but in this thesis we mainly refer to the Lyα emitting galaxies as LAEs. The main sources of Lyα photons in a galaxy are recombinations in the ISM (e.g. Partridge & Peebles, 1967; Tasitsiomi, 2006). Ionizing photons ($\geq$ 912Å) from young, massive O and B stars ionize the H in the ISM which then recombines to produce Lyα photons. Other sources of Lyα photons are Quasars/AGN and Lyα blobs. Quasars/AGN are believed to be super-massive black-holes in the centres of galaxies with gas falling onto them through an accretion disk. They produce powerful jets along the axis of the disk and tremendous amount of radiation due to the heating of the disk. UV photons from the AGN ionize H which recombines to emit Lyα in two cones perpendicular to the disk (e.g. Haiman & Rees, 2001; Weidinger et al., 2005). They consitute only about 1% of the observed population of LAEs (Dawson et al., 2004; Gawiser et al., 2007; Ouchi, 2008). Lyα blobs are extended Lyα objects spanning 10-100 kpcs (e.g. Steidel et al., 2000; Matsuda et al., 2004; Nilsson et al., 2006). The origin of this emission is thought to be cooling radiation by gravitationally heated gas in DM haloes (e.g. Haiman, Spaans, & Quataert, 2000; Fardal et al., 2001; Dijkstra, Haiman, & Spaans, 2006). But recently it has been suggested that the observed Lyα photons are re-scattered radiation, probably produced by galaxies hosted within the nebulae (Faucher-Giguère et al., 2010; Hayes et al., 2011).

Using multiwavelength observations of LAEs at different redshifts, some of their physi-

cal properties have been inferred even though there are large deviations. From broadband spectral energy distribution (SED) fits, they are found to be either young galaxies (~ 10 Myr; e.g. Gawiser et al., 2007) or relatively old galaxies (1 Gyr; e.g. Nilsson et al., 2007) with stellar masses in the range 10^{8-9} $M_\odot$ (e.g. Gawiser et al., 2006; Nilsson et al., 2007). They have typical star formation rates of 1-10 $M_\odot yr^{-1}$ (e.g. Cowie & Hu, 1998; Gronwall et al., 2007; Nilsson et al., 2007; Blanc et al., 2011). Observations at $z \sim 3$ have set the escape fraction of Lyα photons from the host galaxies at $\sim 30\%$ (e.g. Hayes et al., 2010; Blanc et al., 2011). At low redshifts ($z < 5$) there is evidence for the presence of dust in LAEs, but its abundance shows a large variation (e.g. Gronwall et al., 2007; Lai et al., 2007; Finkelstein et al., 2007; Pirzkal et al., 2007; Blanc et al., 2011). Clustering has been calculated by a large number of authors at different redshifts (e.g. Kovač et al., 2007; Gawiser et al., 2007; Ouchi et al., 2010). Using clustering measurements, Ouchi et al. (2010) found that the LAE dark matter haloes lie in the range 10^{10-11} $M_\odot$. Towards lower redshifts $z < 2$, the fraction of AGN among LAEs seem to increase as expected with the peak of star formation in the Universe (e.g. Wolf et al., 2003). Ouchi (2008) showed that the LAE LF does not evolve much for $z < 5.7$, but it seems to decrease towards higher redshifts, possibly due to increasing IGM neutral fraction before EOR.

1.4.1 Importance of LAE for EOR

Once produced, the Lyα photons are resonantly scattered by neutral H present in the gas. The Gunn-Peterson optical depth τ due to scattering away from the line-of-sight to the observer (at $z = 0$) for a LAE at $z = z_s$ is given by:

$$\tau \simeq 6.02 \times 10^5 \; x_{\rm HI} \left(\frac{1+z_s}{10}\right)^{3/2}, \tag{1.1}$$

for photons entering the IGM with $\lambda \lesssim \lambda_\alpha = 1215.67$ $\mathring{A}$ (e.g. Gunn & Peterson, 1965; Barkana & Loeb, 2001) and:

$$\tau = 2.9 \; x_{\rm HI} \left(\frac{\Delta v}{600 \text{ km s}^{-1}}\right)^{-1} \left(\frac{1+z_s}{10}\right)^{3/2} \tag{1.2}$$

for photons entering the IGM with $\lambda = \lambda_\alpha \, (1 + \Delta v/c)$ (e.g. Miralda-Escude & Rees, 1998; Dijkstra & Wyithe, 2010). We can see that the optical depth τ at Lyα rest wavelength for a source at $z = 7$ embedded in a gas of neutral hydrogen fraction $x_{\rm HI} = 0.1$ is $\tau = 43000$, which is very high. This effect would lead to an effective dimming of flux from LAEs at high redshifts. Only when the photon has redshifted to a wavelength equivalent to a velocity shift of 125 km s^{-1} does the optical depth drop down to ~ 1. Thus only sources embedded in ionized bubbles with a radial distance large enough to allow the photons to redshift by > 125 km s^{-1} due to cosmic expansion, will have their flux reaching the observer without scattering. This effect is applied to determine the ionization history of the Universe using LAE observations (e.g. Miralda-Escude & Rees, 1998; Haiman & Spaans, 1999; Malhotra & Rhoads, 2004; McQuinn et al., 2007).

LAEs LFs, number density and clustering are the main methods used to study EOR history. Haiman & Spaans (1999) showed that the evolution of the LF of LAEs can be used to constrain the epoch of reionization. They showed that an increment in the neutral fraction of the Universe leads to a reduction of the observed number of sources in each luminosity bin, thus supressing the LFs towards higher redshifts. Recently, Ouchi et al. (2010) used the drop in LAE LF at $z = 5.7 - 6.6$ to constraint $x_{\rm HI} \lesssim 0.2 \pm 0.2$ at $z = 6.6$. Malhotra & Rhoads (2006) proposed a method to set a lower limit to the IGM volume averaged ionization fraction. They estimate $x_{\rm HI}$ considering the volume occupied by fully ionized spherical bubbles around observed sources such that 50% of the intrinsic Lyα photons from these sources escape towards the observer. They find that at least 20-50% of the volume of the Universe needs to be ionized to account for the observed number of LAEs at $z = 6.5$ in Taniguchi et al. (2005). McQuinn et al. (2007) showed that clustering of LAEs increases with the neutral fraction in the universe and thus can be used to constrain EOR (see also Furlanetto, Oh, & Briggs 2006). Ouchi et al. (2010) used LAE clustering measurements at $z = 6.6$ to give a constraint of $x_{\rm HI} \lesssim 0.5$.

Due to their strong Lyα emission lines, LAEs are the natural targets of high redshift galaxy surveys (explained in the next sub-section) using narrow band detection techniques. The other main method of detecting high redshift galaxies is through searches for the Lyman break in the galaxy spectra (e.g. Steidel et al., 1996). The galaxies detected by this method are known as 'Lyman Break Galaxies' (LBGs). These have recently been observed up to very high redshifts (e.g. Schenker et al., 2011; Ono et al., 2011; Pentericci et al., 2011). A fraction of these galaxies also have a strong Lyα emission. Using the sample, they find that the fraction of LAEs in LBGs increases with redshift for $z < 6$, but decreases strongly between $z = 6 - 7$ which is thought to be due to the above explained effect. The results are consistent with $x_{\rm HI} \sim 0.6 - 0.9$ at $z \sim 7$.

Thus we can see that LAEs provide a useful tool to study EOR which is further explored in Section 3.3 of this thesis. In addition to being one of the main methods of detecting high redshift galaxies to study the evolution of galaxy properties with redshift and studying the EOR (e.g. Miralda-Escude & Rees, 1998; Malhotra & Rhoads, 2004; McQuinn et al., 2007), LAEs can also be used to constrain dark energy properties using Baryonic Acoustic Oscillations in the LAE clustering signal (e.g. Hill et al., 2008; Eisenstein et al., 2005).

1.4.2 Observations

Since the prediction of the Lyα photon production in galaxies by Partridge & Peebles (1967), a large number of efforts were aimed at observing these objects especially at high redshifts. The strength of the Lyα line makes it an ideal candidate to search for high redshift galaxies (e.g. Partridge, 1974; Davis & Wilkinson, 1974; Meier, 1976), especially using Narrow-Band (NB) surveys (e.g. Pritchet & Hartwick, 1989; Rhee, Webb, & Katgert, 1989). But early observations lead to non-detections which were explained using models such as absorption of Lyα by dust (e.g. Spitzer, 1978; Meier & Terlevich, 1981; Hartmann et al., 1988; Charlot & Fall, 1993) and lower number of massive stars due to aging of stellar populations (Valls-Gabaud, 1993).

First claims of LAE detections were made by Lowenthal et al. (1991) but it was not clear as the object was close to an AGN which is also another strong source of Lyα photons. The first detections of a high redshift LAE sufficiently far away from an AGN was reported by Hu & McMahon (1996). Since then a large number of LAEs have been observed mainly using narrow band photometry and spectroscopy with some of the notable surveys being the Large Area Lyman Alpha (LALA) survey (e.g. Rhoads et al., 2003; Dawson et al., 2007) and Subaru Deep field (SDF; e.g. Taniguchi et al., 2005; Kashikawa et al., 2011). Candidates selected through NB surveys are confirmed using spectroscopy to remove contaminants such as Hα/β or OII/OIII lines from a lower redshift galaxy (e.g. Kashikawa et al., 2011). Some detections have also been made using weak lensing measurements which amplify the signal from low luminosity LAEs at high redshifts (e.g. Hu et al., 2002; Santos, 2004; Stark et al., 2007).

The redshift distribution for the current observations are as follows. At low redshifts ($z < 1.5$) the observations are made from space based missions like International Ultraviolet Explorer satellite (IUE; e.g. Meier & Terlevich, 1981), Galaxy Evolution Explorer (GALEX; e.g. Deharveng et al., 2008) and Hubble Space Telescope (HST; e.g. Hayes et al., 2007; Atek et al., 2008; Östlin et al., 2009). The lowest redshift where we currently have a LF of LAE is $0.2 < z < 0.35$ (Deharveng et al., 2008). As we move towards higher redshifts, we have LFs, for example, at $0.195 < z < 0.44$ and $0.65 < z < 1.25$ (Cowie, Barger, & Hu, 2010), $z \sim 2.1$ (Guaita et al., 2010), $z \sim 2.3$ (e.g. Nilsson et al., 2011), $z \sim 3$ (Nilsson et al., 2007; Gronwall et al., 2007; Blanc et al., 2011), $z = 4.5$ (Rhoads & Malhotra, 2001; Wang et al., 2009), $z = 5.7$ (Rhoads & Malhotra, 2001; Taniguchi et al., 2005), $z = 6.6$ (Ouchi et al., 2009; Kashikawa et al., 2011), $z \sim 7$ (Iye et al., 2006; Fontana et al., 2010; Vanzella et al., 2011; Schenker et al., 2011; Pentericci et al., 2011; Ono et al., 2011), $z = 7.3$ (Shibuya et al., 2011) and $z \sim 8.6$ (Lehnert et al., 2010). Some of the major surveys which are ongoing or planned include the one with Hyper Supreme Cam[7] on Subaru Telescope, Observations with JWST[8], studies at $z = 7.7$ (e.g. Tilvi et al., 2010; Clément et al., 2011) and HETDEX[9] at $z \sim 3$, which will give us a much larger and deeper sample of LAEs at different redshifts.

1.4.3 Modelling

Modelling LAEs is a complex task as one needs to take into account the properties of the galaxies such as star formation, dust production, Lyα production in the ISM, outflows from the galaxy due to winds, etc along with the changing IGM properties due to the redshift evolution of gas due to gravity and by radiative feedback from the UV background and other ionizing sources in the region. On top of these one needs to properly model the 3D Lyα RT through the gas and dust distribution in the ISM, circum-galactic medium of the galaxy and the IGM around it till Lyα photons redshift out of resonance and stream freely towards the observer. The spatial scales and the physics involved cannot currently be modelled with the required accuracy in a single simulation. But the availability of large

[7] http://www.naoj.org/Projects/HSC/HSCProject.html
[8] http://www.jwst.nasa.gov/
[9] http://hetdex.org/

LAE datasets and the curiosity to understand the EOR compels us to attack the problem from different angles.

Analytic solutions were calculated for the Lyα RT through different relatively simple gas and dust configurations (e.g. Ambarzumian, 1932; Henyey, 1940; Zanstra, 1949; Unno, 1952; Osterbrock, 1962; Gunn & Peterson, 1965; Neufeld, 1990, 1991; Loeb & Rybicki, 1999; Dijkstra, Haiman, & Spaans, 2006; Tasitsiomi, 2006) [Refer to Laursen (2010) Chapter 3 for a good review]. These were then used in both semi-analytic works (e.g. McQuinn et al., 2007; Iliev et al., 2008; Dayal et al., 2009) and Lyα RT simulation codes (e.g. Dijkstra, Haiman, & Spaans, 2006; Tasitsiomi, 2006; Verhamme, Schaerer, & Maselli, 2006; Pierleoni, Maselli, & Ciardi, 2009; Laursen, 2010; Faucher-Giguère et al., 2010; Zheng et al., 2010; Yajima et al., 2011).

Semi analytic methods and simulations try to model a coherent picture of the observed LAEs at different redshifts which are then compared to observations to obtain fits to the free parameters in the model. A large number of efforts have gone into modelling different aspects of this problem with most of them focussing on the evolution of LFs. Some of the papers are as follows. Santos (2004) simulated a single LAE and studied how its shape changes for different parameters in the model such as halo properties, IGM ionization and velocity fields, redshift, etc. He used analytic models for halo and IGM properties and the RT through the IGM was calculated using the analytic formula of the optical depth due to scattering of the photons off the line-of-sight. It was shown that galactic winds reduces the strength of the IGM ionization fraction constraints from the LAE line profile data. Dijkstra, Lidz, & Wyithe (2007) showed the importance of initial spectrum of the line from the galaxy, clumping of gas and the effect of clustering of nearby sources in improving the observability of LAEs. In the next paper, Dijkstra, Wyithe, & Haiman (2007) showed that the neutral gas due to recombinations in a highly ionized IGM can still scatter Lyα photons reducing the observability. The effects due to the expansion of the IGM with redshift was enough to explain the joint evolution of UV and LAE LFs between $z = 5.7$ and 6.6.

Kobayashi, Totani, & Nagashima (2007) showed that LAE LFs at different redshifts can be fit with changing values of Lyα photon escape fraction from the galaxy and changing transmission through the IGM. The escape fraction of Lyα photons in a galaxy depends on the dust distribution and outflow properties in the galaxy which is a topic of a large number of simulations as would be explained later. Samui, Srianand, & Subramanian (2009) used models where the LAE LF is fit to observations considering that only a fraction of the galaxies are LAEs at a specific time with the fraction increasing with redshift. They use dust clumping in galaxies as a method to change the fraction of galaxies which could be observed as LAEs. Tilvi et al. (2009) based the fits of DM halo mass function to LAE LF at each redshift on the star formation efficiency parameter which depends on the gas accretion onto the galaxy at that redshift. Dayal et al. (2009) included dust in the galaxy, cooling radiation, clustering of sources and evolving IGM to fit for escape fraction of Lyα photons. They showed that ionization fraction estimated from LAE LF is degenerate with escape fraction calculations. Other than fitting LFs, McQuinn et al. (2007) showed that clustering of LAEs can also be used to infer IGM ionization and galaxy properties (see also

Furlanetto, Oh, & Briggs (2006); Mesinger & Furlanetto (2008); Iliev et al. (2008)).

Zheng et al. (2010) showed that Lyα RT effects in the IGM along with observational thresholds can lead to a low effective escape fraction of Lyα photons (also refer to Laursen (2010)). Stacked results of Lyα observations at $z \sim 3$ by Steidel et al. (2011) showed that as one observed lower surface brightness thresholds (see also Hayashino et al. (2004)), all the Lyα flux from the galaxy is recovered leading to an effective Lyα escape fraction of 1. The Lyα photons from the galaxy were modelled to be scattered by the neutral gas in the galactic outflows present in the circum-galactic medium of the galaxy. Dijkstra & Wyithe (2010) have argued that galactic outflows can help Lyα photons escape a galaxy even in highly neutral IGM. Also, Schaerer et al. (2011) has generated a suite of models which show how the Lyα line profile changes for different outflow profiles by moving it to redder wavelengths thus easing escape from neutral gas in the IGM. Zheng et al. (2010) has showed that Lyα RT effects also affect clustering signal from LAEs.

On galaxy scales, Tasitsiomi (2006) showed that the Lyα RT in a galaxy leads to more extended surface brightness profile and affects the line shape. They also showed that the source of Lyα photon production in recombination is more than collisional cooling and supernovae remnants. Recently, Yajima et al. (2011) showed the effects of Lyα RT including dust absorption and scattering effects and the time evolution effects on the Lyα observations. They also showed the coevolution with UV and Infrared emission from the galaxy.

Thus we see that a lot of parameters, effects and spatial scales can affect LAE observability. In this thesis, we focus on the IGM close to the source and its effects on LAE observability.

1.5 Motivation for this Thesis

As we have discussed, LAEs provide a useful tool to study the epoch of reionization and is currently the main focus of a large number of observational campaigns. But understanding the observations has been a hard task due to the large number of factors and spatial scales affecting it. A lot of effort has gone into modelling the LAEs with RT schemes focussing on galaxy scales and IGM scales separately. But these simulations either lack the resolution or the necessary physics needed for the proper modelling of the LAEs. Also most of them show a dependence on a specific reionization model. In this thesis, we try to minimize the uncertainties and study their effects while specifically looking into the effect of the structure in the IGM close to the galaxy on the observability of LAEs. IGM close to the galaxy is important as it has the highest impact on the scattering of photons if we are trying to use them to study the epoch of reionization. Further away from the source, due to the redshift of the photons, IGM has lesser effect on Lyα photon RT unless it has high infall velocities, which is very unlikely.

Previous works assume either very low ionization levels or has semianalytic models for the IGM. Semianalytic models do not accurately represent the spatial and dynamical properties of the IGM. Recent work by Zheng et al. (2010) focussed on the effect of IGM

on LAE observations but it was done at a lower resolution. This would lead to not enough resolution of the density and velocity field close to the object. Also, the ionization level of their simulation is very high $x_{HI} = 10^{-4}$. But they showed the importance of RT for studying the observability of LAEs. In this work, we simulate a large sample > 100 LAEs at a high redshift $z = 7.7$ in a initially neutral IGM where the ionized bubble is created by the source. We study the effects of IGM structure on the surface brightness maps of the objects and the statistical trends they give. We also do a parameter study of the different factors in our simulations which could significantly affect our results.

We find that detailed initial ionization structure due to neighbouring/background sources play an important role in determining the observability of LAEs. Therefore, to accurately obtain the ionization history of the region due to background sources and the ones clustered in the region, we need simulations of reionization. To simulate reionization accurately, one needs large comoving volumes to sample a representative volume of the Universe containing the patchy nature of reionization and the large bubbles while high spatial/density resolution to resolve the high density LLSs which control the rate of evolution of reionization history towards low redshifts. Since simulating very large volumes with very high resolution is a difficult task, the usual method is to simulate large volumes with lower spatial resolution but using 'clumping factors' to evaluate the correct ionization level in the cell if we were to have the required high resolution. Therefore, we analyse a suite of simulations to understand the important factors which govern reionization simulations at different redshifts. We also calculate clumping factors from the highest resolution simulation we have of the IGM and study the different factors which affect clumping factor calculations.

The work in this thesis is part of a larger project aimed at simulating LAEs in a large comoving volume and comparing to upcoming observations. Chapter 2 explains the pipelines developed for modelling LAEs. The next chapter shows the statistical results and the parameter study. Chapter 4 deals with the analysis of the reionization simulations and clumping factor calculations. The conclusions of this thesis are presented in Chapter 5.

Chapter 2

LAE Modelling

In this thesis, we study the effect of structure in the IGM close to the object on the escape of Lyα photons from the source. This region is important because the inhomogeneities in the gas density play an important role in determining the size of the ionized region, especially at very low ionization fractions. The residual neutral hydrogen close to the source is very efficient in scattering the Lyα photons as the probability of scattering decreases as the photons redshift to longer wavelengths. In addition, the velocity field in this region strongly affects the Lyα escape through gas inflows and outflows. This has never been studied in detail for a large (i.e. > 100) sample of objects. The study is done on objects at $z = 7.7$ assuming an initially neutral IGM. Recent observational constraints place the expected mean ionization levels at $x_{\rm HI} > 0.1$ for $z > 7$ while the exact predicted values are model dependent (e.g. Mesinger, 2010; Schenker et al., 2011; Bolton et al., 2011). Confirmed detections of LAEs have been made up to $z \sim 7$ (e.g. Iye et al., 2006; Ono et al., 2011) with more studies aiming for $z \sim 7.7$ (e.g. Tilvi et al., 2010; Clément et al., 2011) which makes it an interesting epoch for the current study.

Zheng et al. (2010) previously simulated a large sample of LAEs with Lyα RT at moderate spatial resolution in a highly ionized IGM at $z = 5.7$. Because they studied the impact of IGM radiative transfer on much larger spatial scales in a highly ionized IGM, their analysis is complementary to this thesis. More recently, Laursen, Sommer-Larsen, & Razoumov (2011) studied scattering in the IGM using hydrodynamical simulations that had much higher spatial resolution, but again focused on the post-reionization IGM.

In this thesis we simulate LAEs using galaxies from hydrodynamical cosmological simulations and radiative transfer using CRASHα (Pierleoni, Maselli, & Ciardi, 2009; Pierleoni et al., 2012). We make surface brightness maps of the objects and study the effect of IGM structure on the surface brightness profiles. In the following Section, we describe the cosmological simulations and Section 2.2 briefly describes the radiative transfer code CRASHα. Section 2.3 explains the pipeline used to simulate a statistically significant sample of LAEs.

Model	L [h^{-1}Mpc]	number of particles	$m_{\rm DM}$ [$M_\odot$]	$m_{\rm gas}$ [$M_\odot$]	η [h^{-1}kpc]
L05	5	2×320^3	3.93×10^5	6.04×10^4	0.78
L10	10	2×320^3	3.14×10^6	4.83×10^5	1.56
L20	20	2×320^3	2.52×10^7	3.87×10^6	3.13
L30	30	2×320^3	8.49×10^7	1.30×10^7	4.69

Table 2.1: Simulation properties. From let to right: model name; comoving box size, L; total number of particles (DM and gas); mass of DM particles, $m_{\rm DM}$; mass of gas particles, $m_{\rm gas}$; softening length, η.

2.1 Simulations of Galaxy Formation

To study high redshift galaxies and their surrounding IGM, we need simulations with a large box size to provide a statistical sample of halos in a wide mass range in different environments. At the same time, we need high resolution to resolve the halo and the structure in the surrounding IGM. A compromise was achieved by using medium range box sizes 5-30 h^{-1}Mpc. The simulations used in this thesis are described in Maio et al. (2010), although additional ones have been run for different box sizes.

The simulations were run using the TREE-PM SPH code Gadget-2 (Springel, 2005) modified to include primordial Hydrogen, Helium and Deuterium based chemistry [e^-, H, H^+, He, He^+, He^{++}, H_2, H_2^+, H^-, HeH^+, D, D^+, HD] (Yoshida et al., 2003; Maio et al., 2007), stellar evolution and metal pollution (Tornatore et al., 2007) and fine structure metal transition cooling [O, C^+, Si^+, Fe^+] at $T < 10^4$ K (Maio et al., 2007, 2009). Metals produced by AGB stars and supernovae [SNII, SNIa] are spread by supernovae/wind feedback. The Initial Mass Function (IMF) is chosen according to the metallicity of the stellar particles, Z. We assume that a transition from metal-free/very metal-poor Population III (PopIII) stars to metal-enriched, more standard PopulationII/I (PopII/I) stars takes place whenever the gas reaches a critical metalicity $Z_{\rm crit}$, which is determined by the cooling and fragmentation properties of the gas. $Z_{\rm crit}$ quoted by different authors range between 10^{-6} to $10^{-3.5}$ $Z_\odot$ (e.g. Schneider et al., 2003; Bromm & Loeb, 2003; Schneider et al., 2006). In these simulations we choose $Z_{\rm crit} = 10^{-4}$ $Z_\odot$. Below $Z_{\rm crit}$ we assume a Salpeter IMF in the mass range [100,500] $M_\odot$, while for higher metallicities, a Salpeter IMF in the mass range [0.1,100] $M_\odot$ is taken.

The simulations analyzed in this thesis have box sizes L = 5, 10, 20 and 30 h^{-1}Mpc with 320^3 particles each in dark matter and gas. The mass of dark matter particles is 3.93 $\times$ 10^5 $M_\odot$ (L/5 h^{-1}Mpc)^3 and the gas particle mass is 6.04 $\times$ 10^4 $M_\odot$ (L/5 h^{-1}Mpc)^3. The comoving softening length is 0.78 h^{-1}kpc (L/5 h^{-1}Mpc)^3 which is $\sim 1/20$ of the mean inter-particle distance. The details of the simulations are given in Table 2.1.

2.2 The Radiative Transfer Code CRASHα

CRASHα (Pierleoni, Maselli, & Ciardi, 2009; Pierleoni et al., 2012) is the first radiative transfer code for cosmological application where the propagation of Lyα and ionizing photons are coupled (refer to Yajima et al. 2011 for a similar code). The ionizing part is based on CRASH (Ciardi et al., 2001; Maselli, Ferrara, & Ciardi, 2003; Maselli, Ciardi, & Kanekar, 2009) which is a 3D ray-tracing grid-based RT code using Monte-Carlo (MC) techniques to sample the probability distribution of several parameters like spectrum of sources, emission direction and optical depth. The ionizing radiation is propagated through an arbitrary static H/He gas density field. Both radiation from point sources and diffuse radiation from ultraviolet background or recombination of H/He gas can contribute to the ionizing flux. The total ionizing photon count, $E_{\rm s}$ [phot], from each source s of luminosity $L_{\rm s}$ [phot s^{-1}] over the whole length $t_{\rm sim}$ of the RT simulation, is distributed among $N_{\rm p}$ photon packets containing ionizing photons of different frequencies depending on the source/background spectrum. The emission of photon packets from each source happens at equally spaced time intervals $dt = t_{\rm sim}/N_{\rm p}$. The direction of emission is assigned by the MC sampling of the angular characteristic of the source. As the packets are propagated through each cell i, photons are absorbed according to the cell optical depth $\tau_{\rm c}^{\rm i}$ calculated combining the contributions from HI, HeI and HeII. The probability that a photon in a packet is absorbed is calculated using $P(\tau_{\rm c}^{\rm i}) = 1 - e^{\tau_{\rm c}^{\rm i}}$. This changes the temperature and ionization conditions of the cell and the photon distribution in a packet. The trajectory of a packet is followed till all the photons are absorbed or in case of non-periodic boundary conditions, till it reaches the edge of the box.

CRASH is modified to follow the time evolution of Lyα photons through an evolving ionization configuration of gas. Lyα scattering being a random-walk process for each photon is very computationally intensive. Thus a statistical approach to the Lyα RT is adopted using pre-compiled tables from MCLyα (Verhamme, Schaerer, & Maselli, 2006). Lyα radiation is emitted by both point sources and recombination of ionized gas. Just like in the ionizing case, the Lyα photons produced by each point source s of luminosity $L_{\rm Ly\alpha,s}$ over the entire simulation time $t_{\rm sim}$ is divided between Lyα photon packets. $N_{\rm p,Ly\alpha}$ Lyα photon packets are emitted by each source s at regular time intervals $dt_{\rm Ly\alpha} = t_{\rm sim}/N_{\rm em,l}$ in random directions. $N_{\rm em,l}$ is the number of times the Lyα photons are propagated during the RT simulation. When a packet enters a cell, depending on the conditions in the cell, the photons are either absorbed or scattered. In case of scattering, a new wavelength and the time of escape from this cell is obtained from the pre-calculated tables. Photons from the recombination in the gas are added to the packets.

Dust is an important factor which determines Lyα radiation transport. Dust optical depth is calculated as a fraction f_τ of the Lyα optical depth. f_τ is determined as $f_\tau = m_{\rm p/dust} \times f_{\rm H/dust} \times \sigma_{\rm dust} \times \rho_{cell} \times d_{\rm cell}$ where proton-to-dust mass ratio $m_{\rm p/dust} = 5.0 \times 10^{-8}$, gas-to-dust ratio $f_{\rm H/dust} = 5 \times 10^{-3}$ (Verhamme, Schaerer, & Maselli, 2006; Pierleoni et al., 2012), dust absorption cross-section $\sigma_{\rm dust} = 2 \times \pi \times r_{\rm dust}^2$ for a dust grain of radius $r_{\rm dust}$ $= 2.0 \times 10^{-6}$cm. ρ_{cell} is the gas density in the cell and $d_{\rm cell}$ is the distance traveled within the cell by the photon. Because dust scattering is negligible compared to dust absorption

Model	$M_{\rm DM}$ [10^{10} $M_\odot$]	M_* [10^7 $M_\odot$]	$M_{\rm gas}$ [10^8 $M_\odot$]	Z [10^{-2} $Z_\odot$]
L05	0.12 - 2.58	6.61 - 256	0.03 - 3.63	1.20 - 6.87
L10	1.33 - 9.70	9.06 - 108	15 - 110	2.02 - 4.80
L20	1.84 - 8.57	4.59 - 64	26 - 110	0.76 - 3.52
L30	3.58 - 16.6	3.54 - 134	47.2 - 213	0.34 - 2.85

Table 2.2: Properties of haloes at z=7.7. From left to right: name of the model; DM mass, $M_{\rm DM}$; stellar mass, M_*; gas mass, $M_{\rm gas}$; gas metallicity, Z.

in high density regions, here we neglect its effect.

For more details, we refer the reader to the original papers.

2.3 Simulating Lyα Emitters

Our plan is to study how the observability of LAEs is affected by transmission through the IGM. For this, we calculate Lyα surface brightness profiles of a large number of simulated LAEs at $z = 7.7$ covering a wide range of dark matter halo masses, thus sampling an equally wide range in IGM environments. The redshift was chosen as observational efforts are underway to investigate LAEs at $z = 7.7$ (e.g. Tilvi et al., 2010; Clément et al., 2011). As at this relatively high z the IGM is expected to be still substantially neutral, a full 3D RT approach is necessary to investigate the observability of LAEs.

The method adopted to simulate LAEs is to extract a cube around each dark matter halo from the snapshot of the simulation of galaxy formation, grid the density and velocity fields, get the spectrum produced by its stellar population from STARBURST99 (Leitherer et al., 1999) using the halo properties, input all the details to CRASHα along with the default values for temperature and ionization fields and run the RT simulation to obtain Lyα photons escaping the simulation cube. We can use the details of the photons exiting each of the six faces of the cube to make surface brightness maps. Each step is explained in detail in the following Sections.

2.3.1 Extracting Halos and Gridding

A snapshot of the hydrodynamical simulation is taken and the properties of the dark matter haloes are obtained using the Friends-of-Friends (FoF) routine (Davis et al., 1985). The centre-of-mass (COM) of the dark matter (DM) halo, the DM, gas and stellar mass, and mean gas metallicity are used in the rest of the analysis. A summary of the DM halo properties for the four different simulations used in this thesis are given in Table 2.2. In our study, we choose a subset of the resolved haloes (i.e. with mass $M_{\rm DM} > 100\ m_{\rm DM}$)

from L05-30 to get a fair sampling of the DM mass range 10^{9-11}M$_\odot$. This results in ~ 130 haloes.

In our simulations, we have the resolution to resolve the structure in the IGM, but this is not enough to properly resolve the ISM of the individual galaxies. To mimic the absorption/scattering of radiation within the ISM, we use the commonly adopted prescription based on the escape fraction of ionizing photons, $f_{\rm esc,ion,ISM}$, and the escape fraction of Lyα photons, $f_{\rm esc,Ly\alpha,ISM}$ (see Sec. 3.2.1). Thus we need to exclude from our RT calculations those cells which represent the ISM, to avoid accounting for their effect twice (refer to Appendix A). We obtain this by removing all cells with a density larger than 0.03 cm^{-3}. Even values up to < 0.05 cm^{-3} are acceptable but might show slight dependence on the gridded density resolution especially for high resolution gridding in the case of small haloes (explanation in the next paragraph). For the case of 0.03 cm^{-3}, the effect of density resolution is negligible in the range of simulations we deal with in this work. Values < 0.03 cm^{-3} give a similar ionization structure but lead to loss of some of the IGM gas. As a precaution to avoid removal of gas from high density regions not associated with this specific dark matter halo, we restrict the removal of cells to a distance of $\sim 0.7 \times r_{200}$ from the source. The precise choice of this radial distance does not affect our main results as long as it is less than the nearest massive dark matter halo in the cube.

A cube around the COM of the halo is extracted with a side of $35 \times r_{200}$, where r_{200} is the radius in comoving units at which the mean DM density inside the sphere is 200 times the critical density of the universe at that redshift. This is the smallest box containing the HII region produced by the galaxy at its equilibrium, so that we get the highest possible spatial resolution for the gridding of density and velocity fields for our RT simulations (refer to Appendix B). A cube size in the range $25 - 50 \times r_{200}$ gives similar results, but the lower bound is too close to the edges of the ionized region, while larger boxes would have lower spatial resolution at fixed grid size, which is undesirable. The exact value of the cube dimension though does not affect our main results.

Once the cube is extracted, the gas density and velocity fields are gridded as input for CRASHα. For gridding, we use GadgettoGrid (Pakmor, 2010). The code distributes the particle mass and velocities using an SPH kernel to 64 neighboring particles and then grids the fields. The default grid size used is 256^3, which is set by memory and runtime constraints of CRASHα, although different grid sizes have been used for testing purposes.

Figure 2.1 shows a slice through the simulation box. The top panel refers to the gas number density in the box with the to-be-extracted cube around a halo marked by a solid red line and the COM of the halo marked by a green dot. The extracted density field is gridded and shown in the bottom panel. Note that the gas density in the IGM is not uniform and the gridded density field has a range in values spanning many orders of magnitude.

2.3.2 Luminosity of Stellar Sources

As mentioned in the previous Section, we obtain the halo properties using the FOF and use the associated IMF, stellar mass and metallicity to calculate the corresponding ionizing

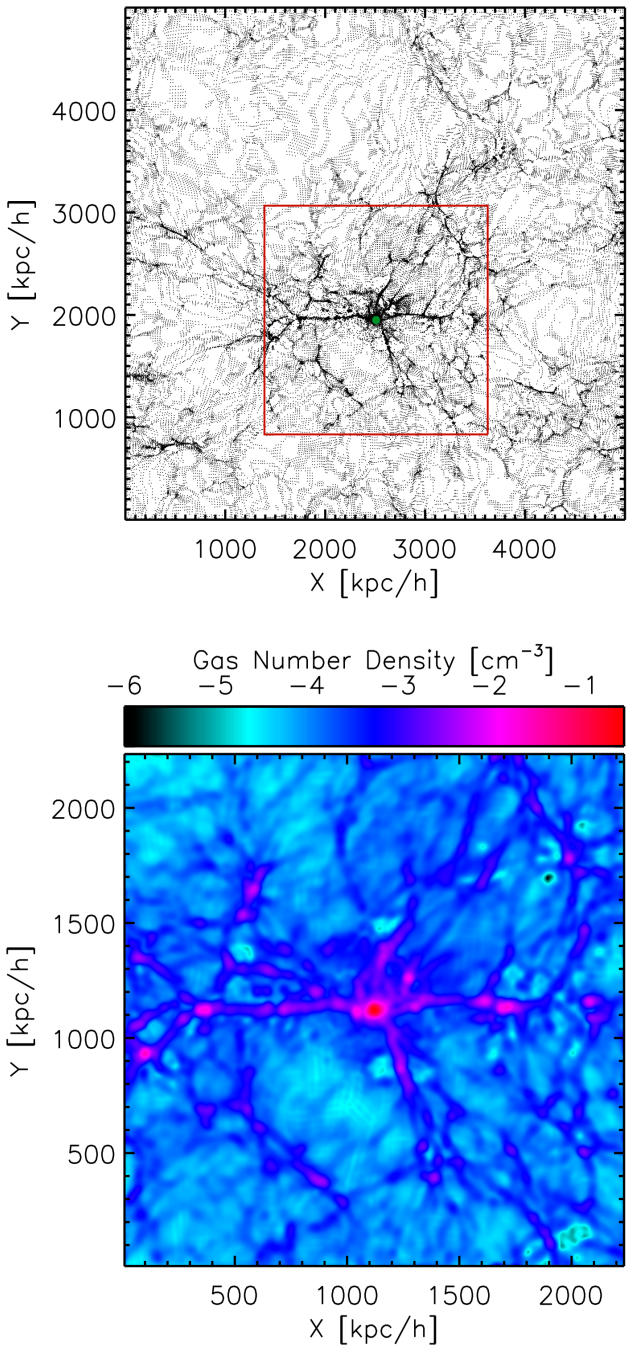

Figure 2.1: Example of the extraction of a cube around a dark matter halo and the gridding of the gas density and velocity fields for CRASHα. *Top*: The gas distribution in a slice of the simulation box L05 centred on the COM (shown as a green dot) of the most massive halo at $z = 7.7$ in comoving units of length. The comoving width of the slice is 10 h^{-1}kpc. The cube to be extracted around the object is shown with a solid red square. *Bottom*: The gridded gas number density field (in logarithmic scale) corresponding to the cube in the top panel.

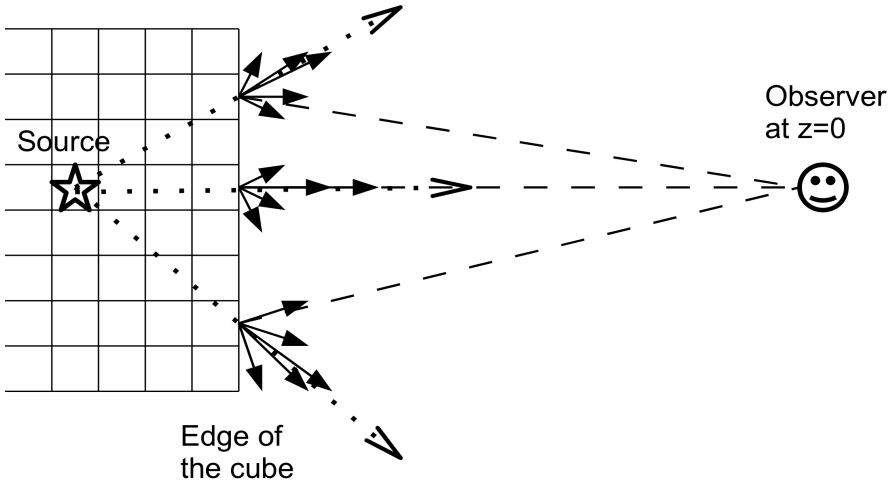

Figure 2.2: Sketch of the method used to calculate the surface brightness. The star represents the pixel with the stellar source. The dotted lines represent the radial direction connecting each pixel on the edge of the box to the source. The solid lines with arrows represent the directions of photon packets within such pixels. Dashed lines connect the pixels to the observer at $z = 0$.

and Lyα photon rate. To this purpose, we use STARBURST99 (Leitherer et al., 1999), which is a population synthesis code that provides the spectrum of a stellar population given an IMF, stellar mass, metallicity and age.

The lowest metallicity available in STARBURST99 is $Z = 0.0004 \sim 0.02$ Z$_\odot$ for Padova tracks. We choose the Padova original stellar evolution tracks instead of Padova AGB tracks as the Hubble age at $z = 7.7$ is less than 1 Gyr, allowing us to ignore the contribution of AGB stars to the spectrum. The average ionizing flux for the same stellar mass is mildly higher for lower metallicities, i.e. $\sim 17\%$ higher for $Z = 3 \times 10^{-3}$ Z$_\odot$ (refer to Schaerer 2003 for low metallicity values). Since our metallicities are above $Z_{\rm crit}$, we use a Salpeter IMF in the mass range [0.1,100] M$_\odot$ in our STARBURST99 models, consistently with our cosmological simulations. The spectrum is scalable with respect to the stellar mass of the halo. Therefore, we choose $M_* = 10^6$ M$_\odot$ as the default stellar mass for the STARBURST99 model and normalize the photon count to this stellar mass. We also choose instantaneous star formation mode for computing the spectrum. The mean ionizing photon rate $\dot{N}_{\rm ion}$ was obtained by averaging the number of photons emitted over $t = 2 \times 10^8$ years in the wavelength range $[91, 912]$Å. The time t was chosen as the cumulative number of ionizing photons $\dot{N}_{\rm ion}(< t)$ reaches convergency. Thus, we obtain a normalized ionizing luminosity $\dot{N}_{\rm ion} = 7.77 \times\ 10^{50}$ $(M_*/10^6$ M$_\odot)$ phot s^{-1}. Only a fraction of these ionizing photons reach the IGM, the rest is converted to Lyα photons in the ISM. Therefore, the ionizing luminosity reaching the IGM, $\dot{N}^{\rm esc}_{\rm ion}$, is:

$$\dot{N}^{\rm esc}_{\rm ion} = f_{\rm esc,ion,ISM} \times \dot{N}_{\rm ion}, \tag{2.1}$$

where $f_{\rm esc,ion,ISM}$ is the escape fraction of ionizing photons from the ISM. Our reference

value is $f_{\rm esc,ion,ISM} = 0.02$ (Gnedin, Kravtsov, & Chen, 2008). In Section 3.2.1 we discuss more extensively this choice.

The calculation of the Lyα photon rate $\dot{N}_{\rm Ly\alpha}$ is more complicated. There are three different components considered in our work - stellar continuum $\dot{N}_{\rm Ly\alpha,\star}$, nebular continuum $\dot{N}_{\rm Ly\alpha,\ neb}$ and recombination in the ISM $\dot{N}_{\rm Ly\alpha,\ ISM}$:

$$\dot{N}_{\rm Ly\alpha} = \dot{N}_{\rm Ly\alpha,\star} + \dot{N}_{\rm Ly\alpha,\ neb} + \dot{N}_{\rm Ly\alpha,\ ISM}. \tag{2.2}$$

$\dot{N}_{\rm Ly\alpha,\star}$ and $\dot{N}_{\rm Ly\alpha,\ neb}$ are calculated from the spectrum by averaging over t at 1215.67Å with a line width of 2 Å, which is the typical intrinsic line width of LAEs (Partridge & Peebles, 1967).

$\dot{N}_{\rm Ly\alpha,\ ISM}$ is instead computed from the ionizing photon rate as (also see e.g. Schaerer, 2003):

$$\dot{N}_{\rm Ly\alpha,\ ISM} = 0.68 \times \dot{N}_{\rm ion} \times (1 - f_{\rm esc,ion,ISM}). \tag{2.3}$$

Assuming our reference value of $f_{\rm esc,ion,ISM} = 0.02$, the Ly$\alpha$ luminosities are $\dot{N}_{\rm Ly\alpha,\star} = 6.94 \times\ 10^{48}\ (M_*/10^6\ {\rm M}_\odot)$ phot s^{-1}, $\dot{N}_{\rm Ly\alpha,\ neb} = 2.83 \times\ 10^{47}\ (M_*/10^6\ {\rm M}_\odot)$ phot s^{-1} and $\dot{N}_{\rm Ly\alpha,\ ISM} = 5.07 \times\ 10^{50}\ (M_*/10^6\ {\rm M}_\odot)$ phot s^{-1}. Thus, $\dot{N}_{\rm Ly\alpha}$ is $5.15 \times\ 10^{50}\ (M_*/10^6\ {\rm M}_\odot)$ phot s^{-1}, and our total Lyα luminosity is completely dominated by photons that were emitted as recombination radiation in the ISM.

Finally, the $L^{\rm esc}_{\rm Ly\alpha}$ which escapes the galaxy into the IGM after absorption by dust can be defined as:

$$L^{\rm esc}_{\rm Ly\alpha} = f_{\rm esc,Ly\alpha,ISM} \times \dot{N}_{\rm Ly\alpha} \times E_\alpha, \tag{2.4}$$

where $f_{\rm esc,Ly\alpha,ISM}$ is the escape fraction of Lyα due to absorption by dust and $E_\alpha = 1.63 \times 10^{-11}$ ergs is the energy of the Lyα photon. Our fiducial value is $f_{\rm esc,Ly\alpha,ISM} = 0.3$ (Dayal et al., 2009). In Section 3.2.1 we discuss more extensively this choice.

Our value of $\dot{N}_{\rm ion}$ is a factor of two lower than the numbers quoted by e.g. Dayal et al. (2009) and Zheng et al. (2010) due to differences in the calculation of the ionizing photon rate. For example, Dayal et al. (2009) uses the IMF mass range of [1,100] ${\rm M}_\odot$ rather than [0.1,100] ${\rm M}_\odot$ employed in our simulations. Thus, their values would be comparable to ours for $f_{\rm esc,ion,ISM} = 0.04$ and $f_{\rm esc,Ly\alpha,ISM} = 0.6$.

2.3.3 CRASHα Input

The inputs for a CRASHα run are the gridded density, velocity and temperature fields (Sec. 2.3.1) and a source distribution with luminosity and spectrum (Sec. 2.3.2). In our reference simulations, we adopt uniform initial fields of ionization, ${\rm x_{ion}} = n_{\rm HII}/n_{\rm H} = 0$, and temperature, $T = 100$ K. A discussion on the different choices for these values is given in Section 3.2.3. The length of the RT simulation is set to 2×10^8 years. The source locations and luminosity were calculated as explained in the previous Section. Other simulation parameters, $N_{\rm p} = 10^6$, $N_{\rm p,Ly\alpha} = 10^4$ and $N_{\rm em,l} = 500$ were determined using resolution tests.

Throughout the simulation, CRASHα collects photon packets exiting the box, recording their frequency, time, position and directional information, which can be used to quantify

observed properties of the source such as surface brightness maps and spectra. Each cell on a side of the simulation cube emits photon packets in different directions. Just like the photon counts, the true distribution of angular directions of the photons exiting from the cells are also sampled by the photon packets. Each direction points towards a different observer thus making each cell (3D) act as a pixel (2D) in the plane perpendicular to the direction of the observer. Here we define the direction of the observer as the one perpendicular to the plane of the side of the cube. By this definition we get six different observer directions for each source. Both surface brightness maps and spectra can be calculated for each observer direction.

The surface brightness $SB_{\rm em}$ of a pixel of size p_s due to a source of luminosity L at a distance d (in Euclidian geometry) is given by:

$$SB_{\rm em} = \frac{L}{4\pi d^2} \times \frac{1}{(p_s/d)^2} = \frac{L}{4\pi p_s^2}. \tag{2.5}$$

The photons, after travelling through the expanding universe, arrive at the observer at $z = 0$. In a cosmological context, the observed surface brightness $SB_{\rm obs}$ of a source of luminosity L at redshift $z = z_s$ observed in a pixel of proper length p_s is (Tolman, 1934):

$$SB_{\rm obs} = \frac{L}{4\pi D_L^2} \times \frac{1}{(p_s/D_A)^2} = \frac{L}{4\pi p_s^2 (1+z_s)^4} = \frac{SB_{\rm em}}{(1+z_s)^4}, \tag{2.6}$$

where D_L is the luminosity distance of the source and D_A is its angular diameter distance. Therefore, to calculate the observed surface brightness in a pixel for a specific observer, we need the number of photons in each cell travelling in the direction of the observer. The sampling of the angular distribution of photons depends on the number of photon packets in each cell. But the average number of photon packets in each cell of our simulation is too low (~ 12) for a smooth sampling of the underlying angular directional distribution. Thus we calculate the angular directional distribution using all the photon packets on a side, by assuming that all cells on each side of the cube sample the same angular directional distribution with respect to the true north for each cell, which is defined as the radial direction from the source to the cell.

Figure 2.2 shows the sketch of the method. The star represents the cell with the stellar source. The dotted lines represent the radial direction connecting each cell on the edge of the box to the source (i.e. true north for each cell). The solid lines with arrows represent the directions of photon packets exiting such cells. The dashed lines from each cell to the observer represent the observer direction which for $z_s = 7.7$ are perpendicular to the side. The packets which are not scattered by the IGM follow the radial direction from the source (i.e. true north), while the scattered photon packets sample random directions. To calculate the $SB_{\rm obs}$ we need to estimate the number of photons reaching the observer at $z = 0$, i.e. traveling along the dashed lines. To calculate the underlying angular direction distribution, we grid the directions of all photon packets (short solid arrows) in each cell, correcting for the different true north direction of each cell (dotted arrow). The gridded distribution of angles with respect to their true north in all the cells on a side are added

and normalised using the total number of photon packets exiting the side. This normalized gridded angular distribution is then imposed on the photon count in each cell to obtain the number of photons in each pixel going in the direction of the observer.

The luminosity, $L_{\rm pixel}$, in a given direction within a pixel is given by:

$$L_{\rm pixel} = \frac{L}{4\pi d^2} \times p_s^2, \tag{2.7}$$

which is calculated by adding up the energy of all the photons in the pixel in the direction $\theta \leq p_s/d$. d is the proper distance of the cell to the source (at the centre of the simulation box). $L_{\rm pixel}$ is converted to the units of surface brightness, i.e. ergs s^{-1} cm^{-2} rad^{-2}, using the formula:

$$SB_{\rm em} = L_{\rm pixel} \times \frac{d^2}{p_s^4}. \tag{2.8}$$

This is then converted to the observed surface brightness $SB_{\rm obs}$ dividing by $(1+z)^4$ as in Equation 2.6. The observed surface brightness $SB_{\rm obs}$ was converted to flux, F, as:

$$F = SB_{\rm obs} \times \frac{p_s^2}{D_A^2}, \tag{2.9}$$

where p_s/D_A is the pixel size of the observer in radians and D_A is the angular diameter distance of the source. Adding up the flux in all pixels, we can obtain the observed source luminosity $L_{\rm obs}$ as:

$$L_{\rm obs} = 4\pi D_L^2 \times F, \tag{2.10}$$

where D_L is the luminosity distance of the source.

This method has been tested against the direct estimation of surface brightness profile using the 1D Lyα RT code of Dijkstra & Wyithe (2010). Our method gives the same surface brightness profiles as the Rybicki-Loeb halo for the case of uniform sphere of neutral gas in an ionized medium (Loeb & Rybicki, 1999; Dijkstra & Wyithe, 2010) and it works well in test cases similar to the configurations we are studying in our simulations.

Thus we have a pipeline to simulate large samples of LAEs. In the next chapter we use this pipeline to simulate > 100 galaxies and study their observability under different conditions.

Chapter 3

Simulated LAEs

In the previous chapter, we set up a pipeline to simulate LAEs at $z = 7.7$ in a neutral IGM. In this chapter, we use that pipeline to simulate a sample of LAEs large enough to obtain meaningful statistical properties and trends at that redshift. We simulate ~ 130 LAEs, evenly sampling the DM mass range 10^{9-11} $M_\odot$ using objects extracted from the boxes L05-30. Table 2.2 contains the properties of these haloes. The simulations were run with our default parameters (see Sec. 2.3.3). Section 3.1 studies the effect of IGM on the observability of an individual LAE and the trends it leads to in a sample of 130 objects. Dependence of the results on the choice of the parameters are discussed in Section 3.2. In Section 3.3, we discuss the effect of IGM structure on the methods used in estimating the mean ionization fraction of the universe. Finally a discussion of the results and the parameter study are done in Section 3.4.

3.1 Results

Before trying to understand the properties and statistical trends of 130 galaxies, we focus on understanding how the IGM close to the objects affects the appearance of a single LAE. We investigate the ionization structure in IGM and how it affects the surface brightness maps. Having understood the effects on an individual galaxies, we study the behaviour of the whole sample and the trends shown in observability and Lyα escape fractions due to the IGM.

3.1.1 Behaviour of an Individual Galaxy

To understand how the density field in the IGM around the source affects the shape of the ionization region, which in turn will reflect on the propagation of the Lyα photons, we look at the ionization structure around the most massive dark matter halo in L05 (same galaxy as in Fig. 2.1). Figure 3.1 shows the mid-plane of the ionization structure at the end of the RT simulation. We can see that the ionized region is not spherical and its edges are shaped by the high density regions (e.g. filaments) in the IGM. Lyα photons

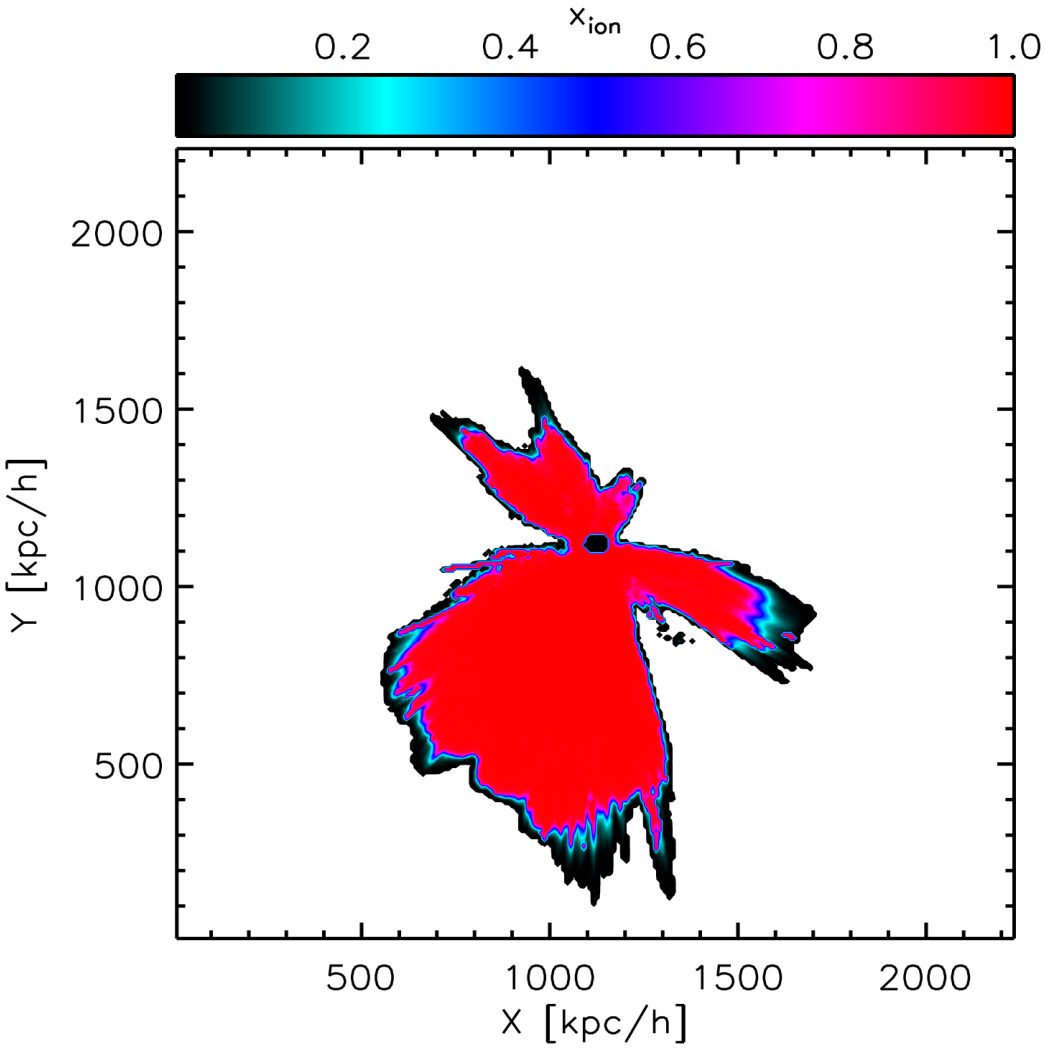

Figure 3.1: The ionized region around the most massive object in L05 shaped by the density structure of the IGM close to the source. The degree of ionization of the gas is shown by the color bar at the top.

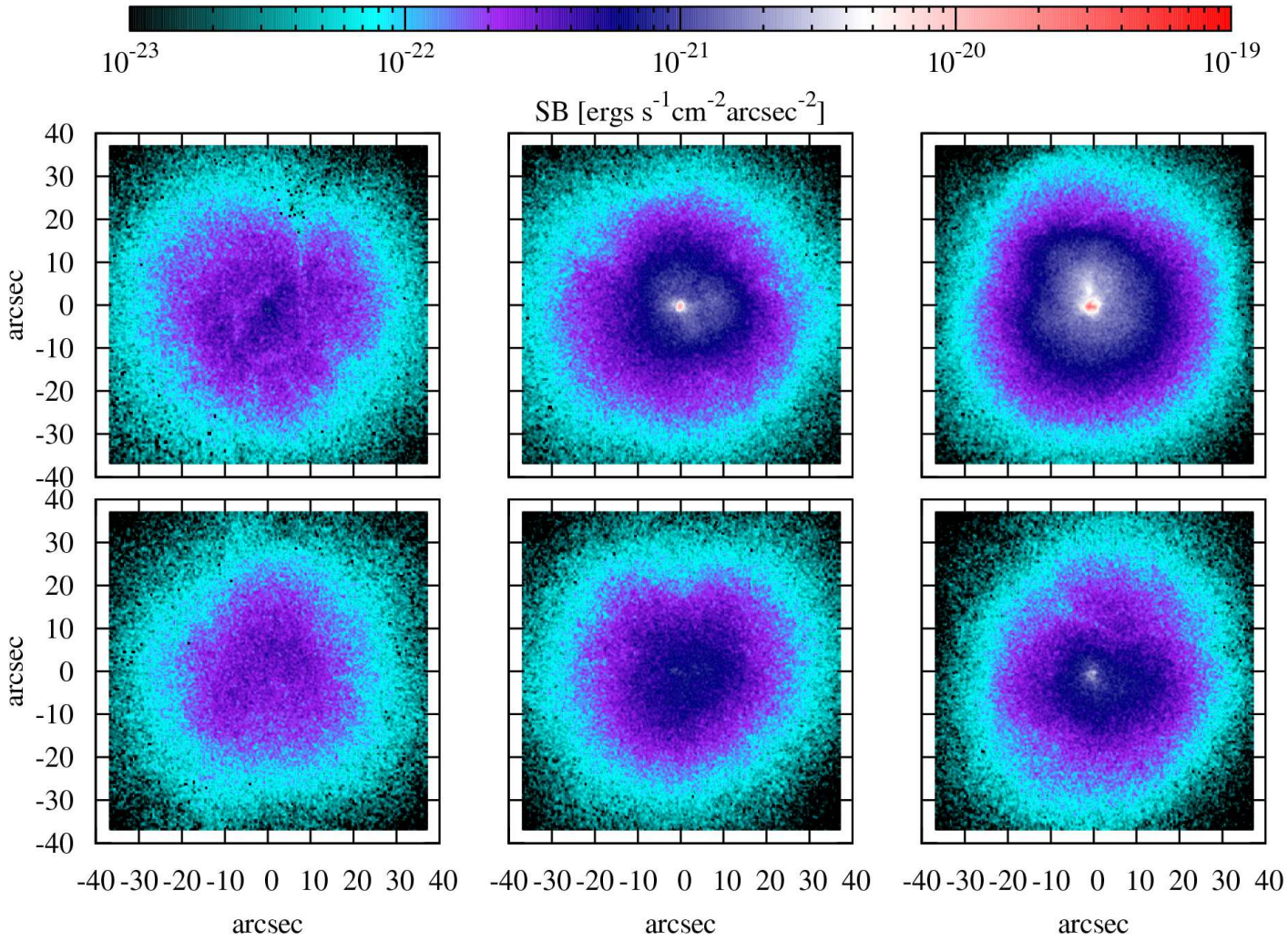

Figure 3.2: Surface brightness maps for the six sides of the radiative transfer simulation box for the most massive dark matter halo in L05. See text for details.

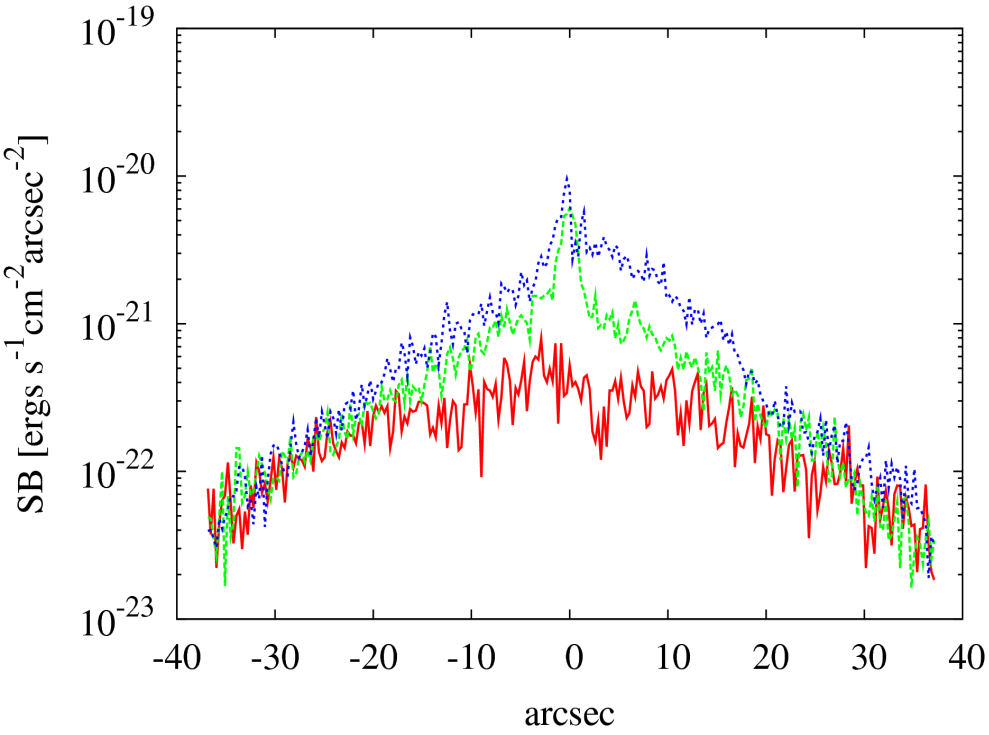

Figure 3.3: A cross section of the surface brightness profile for the maps in the top row of Figure 3.2. The lines are for surface brightness values along the y axis (i.e. x values are 0) of the maps. The solid red, dashed green and dotted blue lines refer to the top-left, top-middle and top-right maps, respectively.

propagating through the ionized gas will travel a large distance unscattered while being redshifted away from the Lyα rest wavelength, thus improving the chances of escape from the neutral IGM beyond the ionized region. Lyα photons encountering a high density filament instead will undergo several scatterings before being able to eventually escape the simulation box. Because of the different paths followed by the photons, we expect different surface brightness distributions and spectra, depending on the viewing direction. This can be clearly seen in Figure 3.2, where the surface brightness maps obtained from the six sides of the simulation box are shown for the same object.

The differences in the maps reflect the structure in the IGM. In general, the lesser the photons are scattered before reaching the edge of the box, the more point-source-like the surface brightness profile is. In this case, the central pixels also have a much higher value of surface brightness compared to cases in which the photons are scattered by high density neutral gas. Note that, as a result of the very inhomogeneous structure of the IGM, the maximum surface brightness value in the image differs by orders of magnitude for different lines-of-sight. This is more evident in Figure 3.3, which shows a cross-section of the surface brightness maps for the three directions shown in the top row of Figure 3.2. Plotted are the surface brightness values in the map for an x-axis value of 0 against the y-axis of the map. As noted earlier, the difference in the value of the central pixels in the maps is more than an order of magnitude for different viewing directions of the same object.

The detectability of the objects simulated in this work depends on the surface brightness thresholds $SB_{\rm th}$ set by different observational campaigns. Since the lowest SB value of a pixel in the maps is $SB_{\rm min} = 10^{-23}$ ergs s^{-1} cm^{-2} arcsec^{-2}, only for such a small (or lower) value of $SB_{\rm th}$ would it be possible to detect all the flux from this object. At this $SB_{\rm th}$ level the observed luminosities, calculated using Equations 2.9 and 2.10, for this object are in the range $L_{\rm obs} = 4 - 12 \times 10^{41}$ ergs s^{-1}. For reference, the input luminosity is $L^{\rm esc}_{{\rm Ly}\alpha} = 9.2 \times 10^{41}$ ergs s^{-1}. This would lead to an observed escape fraction $f_{{\rm esc,Ly}\alpha,{\rm IGM}} = L_{\rm obs}/L^{\rm esc}_{{\rm Ly}\alpha}$ in the range $0.43 - 1.3$.

Some of the lines-of-sights have an escape fraction greater than 1. This is because of the presence of a large void along the line-of-sight. A Lyα photon entering the edges of a high density region (like a filament close to the source) has a high probability of scattering towards the relatively lower density regions around it. Once it is scattered towards the lower density region (a void), the photon travels longer distances without scattering while redshifting out of resonance, improving the probability of escape towards the observer. This leads to the lines-of-sight through voids being preferred over the ones through a high density region. These preferred lines-of-sight get Lyα photon contributions from other lines-of-sight leading to more than average photon escape through these lines-of-sight giving an effective escape fraction > 1 This is different from the case of absorption of photons, where once the photons are removed from the line-of-sight, they do not contribute to any other lines-of-sight leading to an escape fraction ≤ 1 for all lines-of-sight. Where as in the case of scattering, the photons removed from one lines-of-sight appears in another which makes the proper treatment of 3D Lyα important for LAE modelling.

On another note, if e.g. $SB_{\rm th} = 3 \times 10^{-21}$ ergs s^{-1} cm^{-2} arcsec^{-2}, only two maps would result in a detection, with luminosities $L_{\rm obs}(> SB_{\rm th}) = 8 \times 10^{40}$ ergs s^{-1} and

7.9×10^{39} ergs s^{-1}, compared to $L_{\rm obs}(> SB_{\rm min}) = 1.2 \times 10^{42}$ ergs s^{-1} and 8×10^{41} ergs s^{-1}, respectively. Note that if observations had a threshold surface brightness at this level, only a few percent of the photons would be observed leading to a very low effective Lyα escape fraction $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = L_{\rm obs}(> SB_{\rm th})/L^{\rm esc}_{\rm Ly\alpha}$. For $SB_{\rm th} = 10^{-21}$ ergs s^{-1} cm^{-2} arcsec^{-2}, detections would be made in four maps with an observed luminosity in the range $L_{\rm obs}(> SB_{\rm th}) = 2 \times 10^{39} - 5 \times 10^{41}$ ergs s^{-1}, corresponding to $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = 0.002 - 0.6$.

3.1.2 Statistical Trends

Next we discuss how the distribution in surface brightness affects the statistical properties of a large sample of LAEs. The difference in surface brightness profiles in fact leads also to a spread in the observed Lyα luminosity of the same object observed from different directions as seen earlier. The objects in our sample have an intrinsic correlation between the stellar and the DM mass, with a log-log slope of 1.12 and a standard deviation σ=0.13 dex (see Fig. 3.4). The best fit line (dashed) has been obtained using only haloes from L05 and L10. In fact, due to resolution effects in the larger boxes, some of the haloes selected in L20 (blue circles) and L30 (pink triangles) have a stellar mass which lies below the expected trend. A correlation similar to the one shown in Figure 3.4 is also expected to exist between the input Lyα luminosity of the source, $L^{\rm esc}_{\rm Ly\alpha}$, and the DM mass. Any additional scatter in the observed Lyα luminosity $L_{\rm obs}$ is instead due to the effect of the IGM.

Figure 3.5 shows the observed Lyα luminosity $L_{\rm obs}$ of the full sample calculated for each of the six sides of the box plotted against the DM mass for each of the simulated LAE. As in Figure 3.4, there is a strong correlation with a best fit log-log slope of 1.12, which has been calculated using only haloes from L05 and L10. The scatter now though is 0.17 dex, larger than the intrinsic one, due to the structure in the IGM.

An alternative way to look at the scatter due to the IGM is to plot the escape fraction $f_{\rm esc,Ly\alpha,IGM}$ against the DM mass of the objects (Fig. 3.6). This removes the scatter present in the intrinsic luminosity $L^{\rm esc}_{\rm Ly\alpha}$. As we can see, there is no strong correlation between $f_{\rm esc,Ly\alpha,IGM}$ and the DM mass because the scatter due to the IGM structure dominates. The average escape fraction has a nearly constant value of ~ 0.73 with a σ scatter of 0.18. Some lines-of-sight have escape fraction greater than 1 due to the presence of a large void as explained earlier. But the statistical mean of the six observed IGM escape fractions for all the objects in the sample is 0.73. We would get an average escape fraction of 1 only if we observed all the directions for the same object and calculated the mean of all the lines-of-sight. The average of the six viewing directions is < 1, showing that scattering removes a large fraction of the flux for most lines-of-sight. Most of the flux from the object escapes through a small fraction of lines-of-sight through large voids. This again shows the importance of IGM structure close to the source for the observability of LAEs. The spread in the correlation decreases for higher DM masses. This could be due to two reasons - a more uniform environment for haloes of higher masses leading to lesser scatter; or the resolution effects caused by the larger physical size of the grid cells ($\propto r_{200}$) for higher mass haloes compared to low mass ones. But the general consistency of $f_{\rm esc,Ly\alpha,IGM}$ values

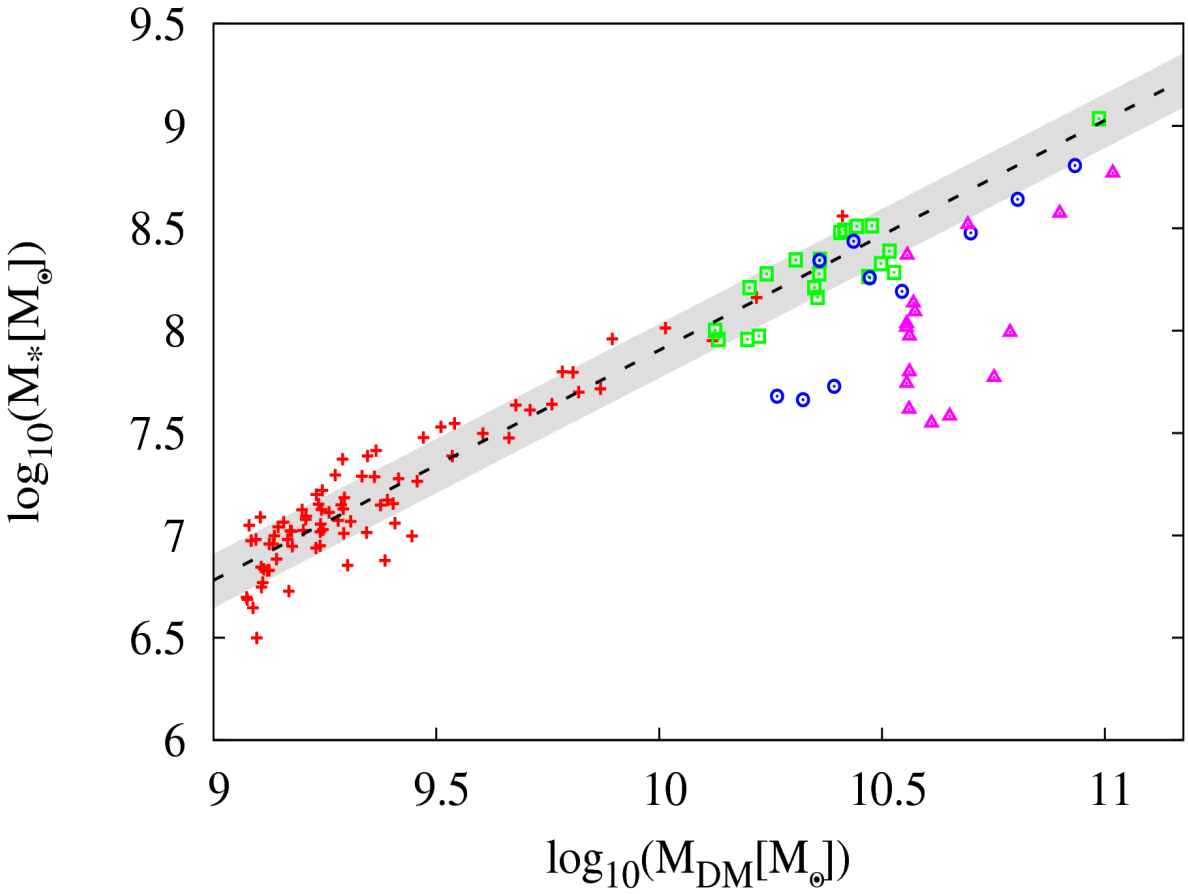

Figure 3.4: Stellar mass of the objects plotted against their DM mass. The objects from different simulations are marked as - L05 (red crosses), L10 (green squares), L20 (blue circles) and L30 (pink triangles). The best fit line (dashed) is obtained using only haloes from L05 and L10, and it has a log-log slope of 1.12 with a standard deviation σ=0.13 dex. The σ region is shaded in light grey.

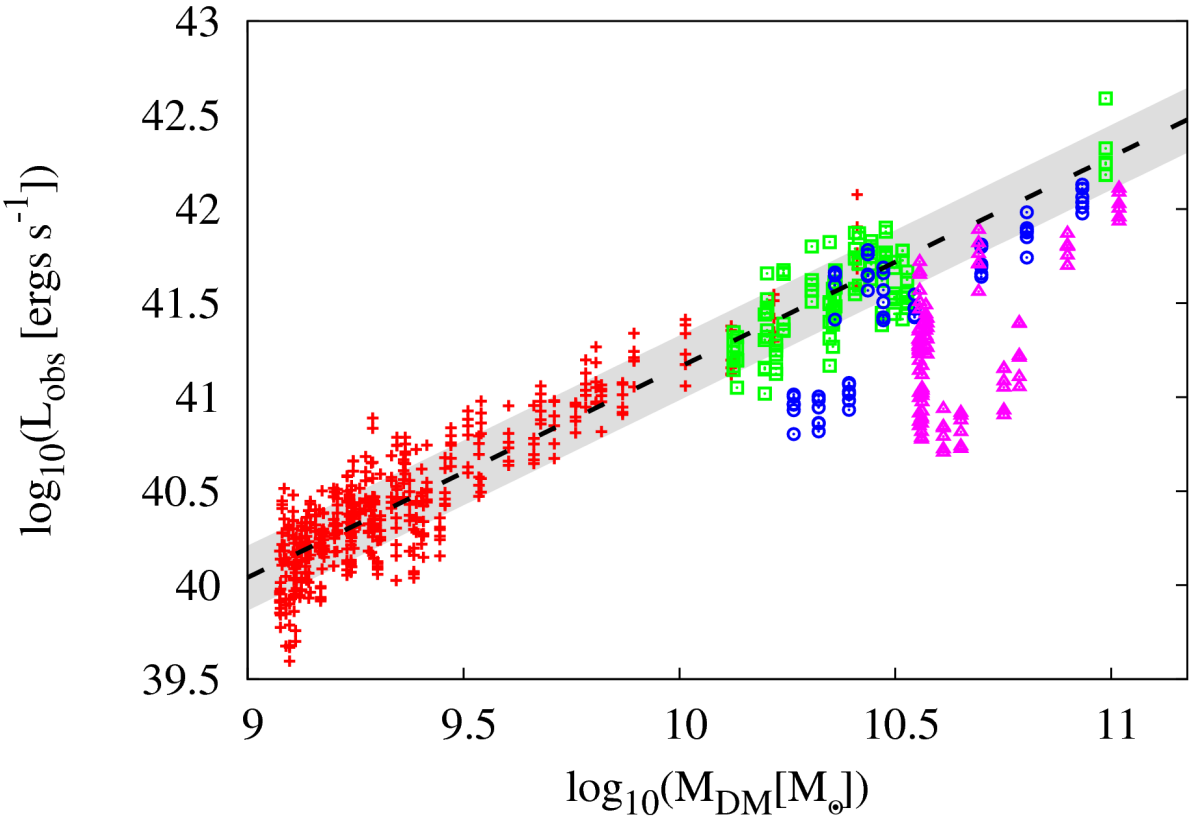

Figure 3.5: Observed Lyα luminosity of the objects plotted against their DM mass. The objects from different simulations are marked as - L05 (red crosses), L10 (green squares), L20 (blue circles) and L30 (pink triangles). The best fit line (dashed) is obtained using only haloes from L05 and L10, and it has a log-log slope of ~ 1.12 with a standard deviation σ=0.17 dex. The σ region is shaded in light grey.

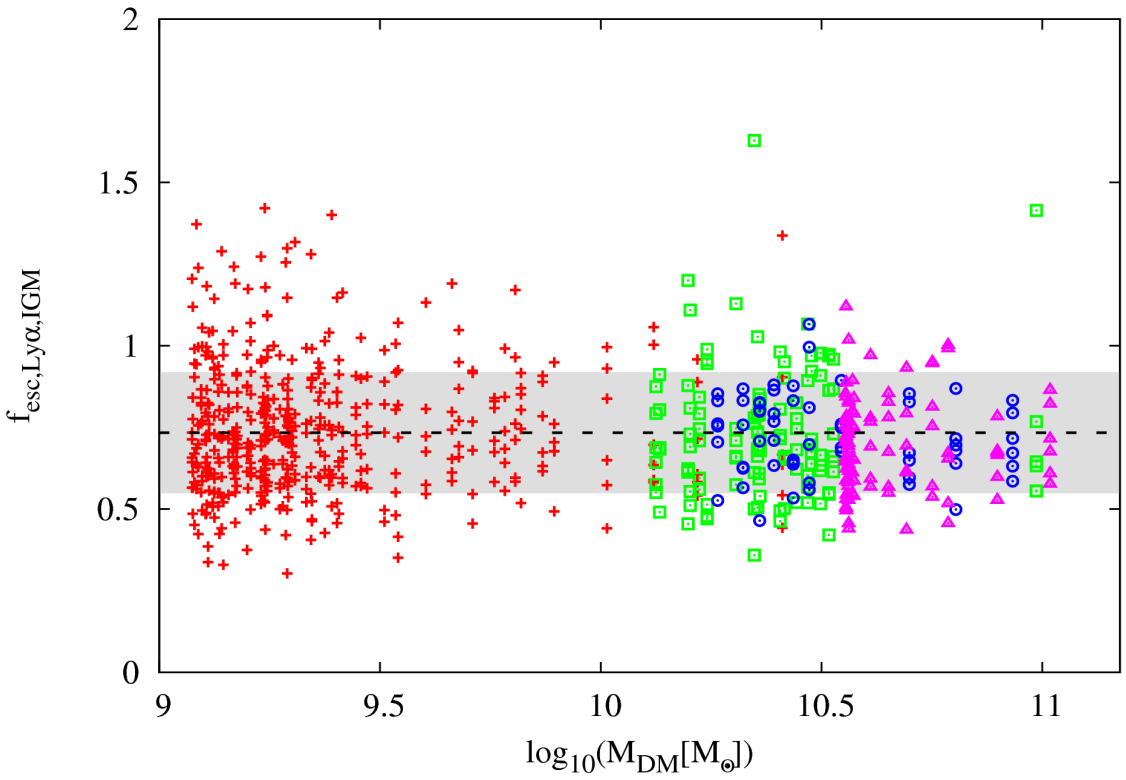

Figure 3.6: Lyα photon escape fractions due to the IGM, $f_{\rm esc,Ly\alpha,IGM}$, plotted against DM mass for objects from L05 (red crosses), L10 (green squares), L20 (blue circles) and L30 (pink triangles) simulations. The best fit line has a log-normal slope of -0.01, with mean $f_{\rm esc,Ly\alpha,IGM}$ of ~ 0.73 and a standard deviation σ=0.18 dex. The σ region is shaded in light grey.

in L20 and L30 with L05 seems to indicate that resolution issues might be less important. Higher resolution RT simulations of larger comoving volumes are needed to confirm this effect, which is beyond the scope of this study.

Next we investigate the detectability of this dataset by plotting in Figure 3.7 the observed luminosity of the objects calculated from the six sides of the simulation boxes for different surface brightness thresholds, $SB_{\rm th}$. Only one object is detected for $SB_{\rm th} \sim 5\times10^{-20}$ ergs s^{-1} cm^{-2} arcsec^{-2} (pink triangle). Note that this is not the most massive object neither in the full sample nor in the simulation box L05. Nevertheless it looks as the brightest due to the density and velocity fields in the surrounding IGM, which lead to a very concentrated surface brightness profile and ease the detectability of the source. From this example it is clear that the structure in the IGM plays an important role in the detectability of objects by shaping their surface brightness profile and only a full 3D RT approach can properly capture its effect.

We caution the reader that current observational surface brightness thresholds are sig-

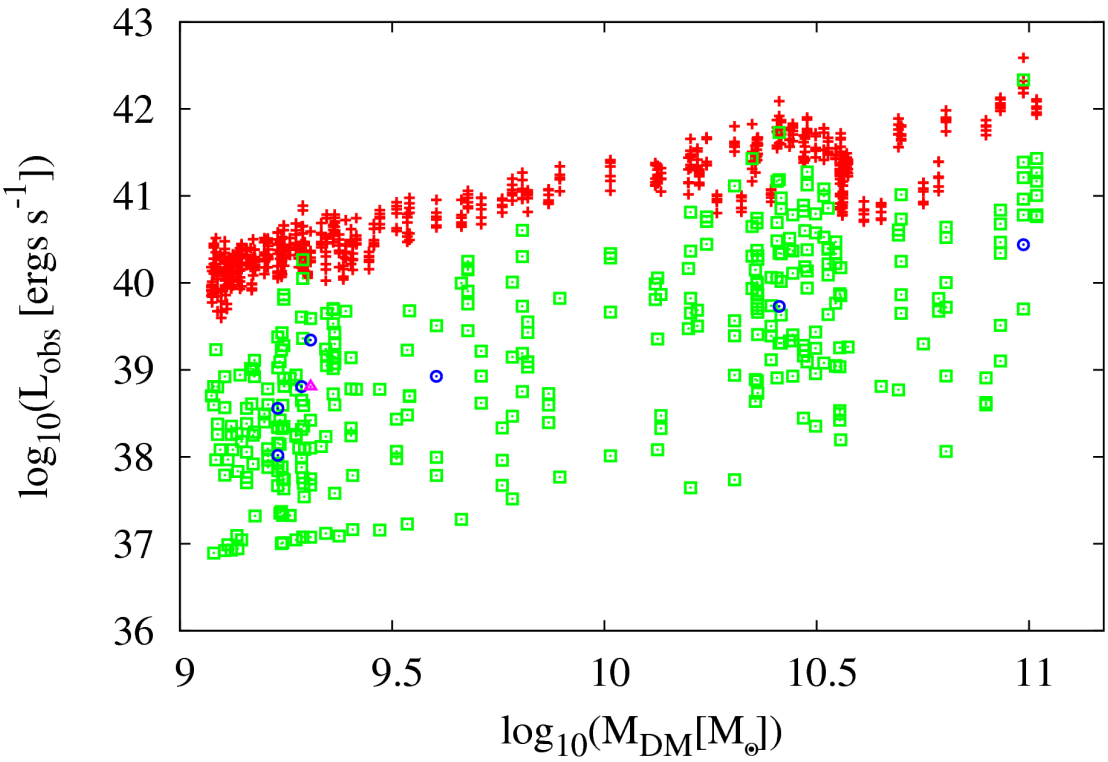

Figure 3.7: Observed Lyα luminosity of the objects from all simulations plotted against their DM mass. Symbols refer to different surface brightness cuts: 10^{-25} (red crosses), 10^{-21} (green squares), 10^{-20} (blue circles) and 5×10^{-20} (pink triangle) ergs s^{-1} cm^{-2} arcsec^{-2}.

nificantly higher than the 'typical' surface brightness levels of Lyα radiation scattered in the IGM as found in our simulations. This leads to an effective escape fraction due to the IGM (also see Laursen et al. 2011) which is very low (a few %). As we lower the surface brightness threshold, more objects are detected and a large fraction of the luminosity is accounted for. By $SB_{\rm th} = 10^{-20}$ ergs s^{-1} cm^{-2} arcsec^{-2} (blue circles), seven objects are detected covering a mass range of two orders of magnitude. Even though there is a correlation of luminosity with DM mass apparent in the data, a large scatter is present at small masses. Decreasing the threshold by another dex, leads to the detection of almost all objects with a scatter in luminosity spanning three dex in each mass range. A surface brightness level of $SB_{\rm th} = 10^{-25}$ ergs s^{-1} cm^{-2} arcsec^{-2} is needed to observe the total flux from all the sources (the same as Figure 3.5). The surface brightness values calculated here depend on the choice of a number of different parameters, which will be further discussed in the next Section.

The observed surface brightness is $SB_{\rm obs} \propto (1+z)^{-4}$ (see Eq. 2.6), making deep detections increasingly difficult at higher redshifts. Current observational thresholds for narrow band surveys are 2.64×10^{-18} ergs s^{-1} cm^{-2} arcsec^{-2} at $z = 5.7$ in SXDS (M. Ouchi; private communication; Zheng et al., 2010). Steidel et al. (2011) achieved deep surface brightness limits of 2.5×10^{-19} ergs s^{-1} cm^{-2} arcsec^{-2} for LAEs at $z \sim 3$ which was obtained through stacking of 92 images. They showed that at such low surface brightness levels, the escape fraction of Lyα is ~ 1. This is in agreement with Figure 3.7. Obtaining these surface brightness thresholds at higher redshifts is a very difficult task. With its predicted surface brightness levels, JWST ($\sim 10^{-19}$ ergs s^{-1} cm^{-2} arcsec^{-2})[1] would be an important step for the deep detections of LAEs.

3.2 Parameter Study

In this Section we study the dependence of our results on different parameters adopted in this study. To this aim we use the biggest object in L05 with a dark matter mass of 2.5×10^{10} M$_\odot$. The reference values are those described in Section 2.3.3.

A minor contribution to the observed flux comes from the photons emitted by the recombining gas in the IGM, which has been estimated to be $\sim 3\%$. This is not significant compared to the flux from the central source, but it should be kept in mind that this flux is not affected by the absorption in the ISM and thus has a higher effective escape fraction. Since it is not concentrated as a point source, it does not contribute to the detectability of the source. In our surface brightness method, it might reduce the strength of the point source by spreading the flux to the scattered component of the angular directional distribution. Due to this effect, we do not include the re-emission photons in our parameter study.

[1] From http://www.stsci.edu/jwst/science/sensitivity for 3σ detection in 10^5s with NIRSpec IFU mode at 0.1×3 arcsec^2 resolution.

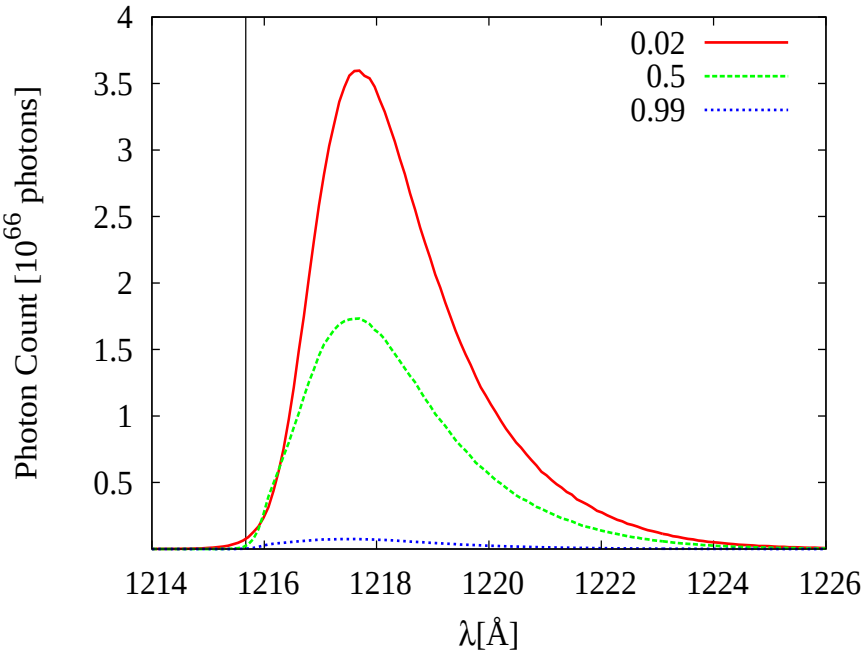

Figure 3.8: Spectrum of all the Lyα photons exiting the simulation box for runs with different $f_{\rm esc,ion,ISM}$. The lines refer to $f_{\rm esc,ion,ISM} = 0.02$ (red solid line, default), 0.5 (green dashed) and 0.99 (blue dotted). The vertical black line marks the Lyα rest-frame wavelength of 1215.67 Å.

3.2.1 Escape Fraction

As already discussed, we need to specify the value of both ionizing, $f_{\rm esc,ion,ISM}$, and Lyα, $f_{\rm esc,Ly\alpha,ISM}$, escape fractions from the ISM. The amount of ionizing photons reaching the IGM is controlled by $f_{\rm esc,ion,ISM}$ thus determining the extent of the ionized region affecting the propagation of Lyα photons. The value of the escape fraction also affects the production of Lyα photons by recombinations in the ISM. In our reference simulations, the box around the object has a size of $\sim 35\ r_{200}$ (comoving). Because this is not large enough to contain the HII region produced by $f_{\rm esc,ion,ISM} = 0.99$, for this parameter study we use an object with DM mass of 2.29×10^{10} M$_\odot$ from L10. We extract a cube with a side of $\sim 117\ r_{200}$ (comoving) which is gridded to a 256^3 grid.

To quantify the effect of changing the value of $f_{\rm esc,ion,ISM}$ on the spectrum, in Figure 3.8 we show the Lyα spectrum of the photons exiting the box for simulations with $f_{\rm esc,ion,ISM} = 0.02, 0.5, 0.99$, while $f_{\rm esc,Ly\alpha,ISM}$ is set to the default value of 0.3. Note that the photon packets are always blueshifted back to the source position from the edge of the box. Also note that the spectra is for all the photons exiting the box through out the simulation time of 2×10^8 years. The Lyα photon count increases with decreasing $f_{\rm esc,ion,ISM}$ (see Eq. 2.3), as recombination in the ISM is the dominant contribution. It also affects the shape of the spectrum by moving its peak closer to 1215.67 Å, as the photons escape with

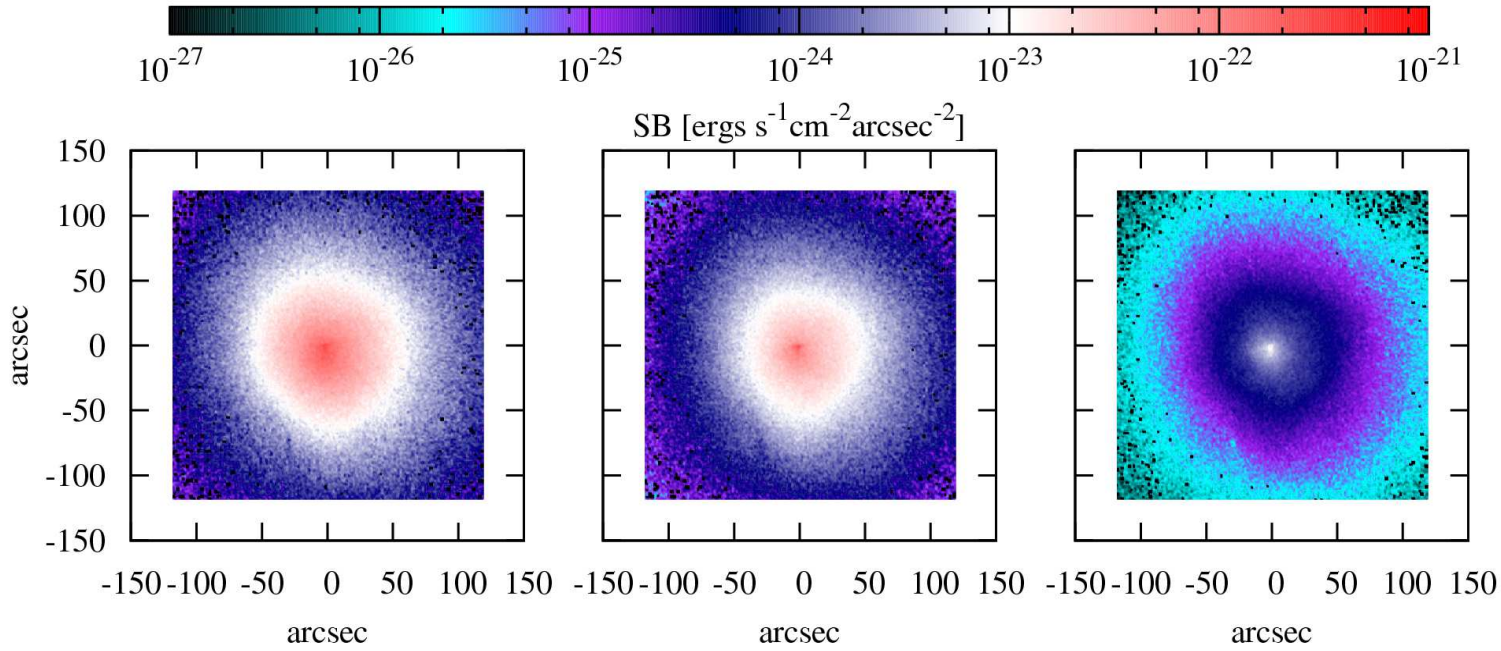

Figure 3.9: Surface brightness maps for one of the sides of the cube for runs with different $f_{\mathrm{esc,ion,ISM}}$. Plotted are the maps from the simulations with $f_{\mathrm{esc,ion,ISM}} = 0.02$ (left, default), 0.5 (middle) and 0.99 (right). See text for details.

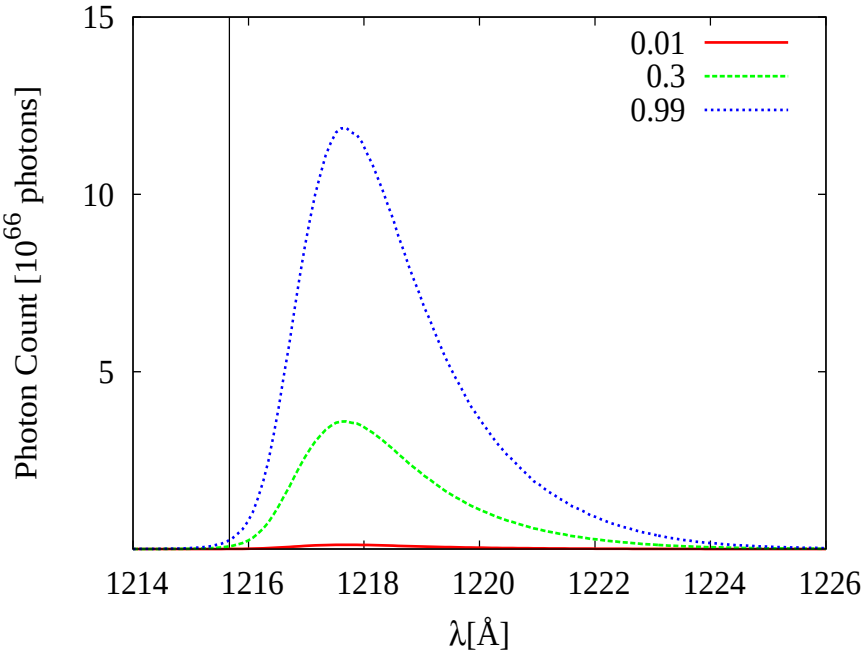

Figure 3.10: Spectrum of all the Lyα photons exiting the simulation box for runs with different $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}}$. The lines refer to $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}} = 0.01$ (red solid line), 0.3 (green dashed, default) and 0.99 (blue dotted). The vertical black line marks the Lyα rest-frame wavelength of 1215.67 Å.

less scatter. This can also be seen from the surface brightness maps in Figure 3.9. The higher the $f_{\mathrm{esc,ion,ISM}}$, the more concentrated is the surface brightness profile. But due to the lower production of Lyα photons in the ISM, the surface brightness values in each pixel decrease for higher $f_{\mathrm{esc,ion,ISM}}$. Thus, increasing $f_{\mathrm{esc,ion,ISM}}$ reduces the overall detectability of the source at a specific SB_{th}, but improves detectability of this galaxy at lower SB_{th} with respect to a low mass galaxy with lower $f_{\mathrm{esc,ion,ISM}}$ giving similar Lyα fluxes.

Direct estimation of $f_{\mathrm{esc,ion,ISM}}$ from observations is a very difficult task (e.g. Bland-Hawthorn & Putman, 2001; Steidel, Pettini, & Adelberger, 2001; Shapley et al., 2006; Iwata et al., 2009; Siana et al., 2010). A large effort has gone into modelling the correlation between halo properties and ionizing escape fraction from the galaxy (e.g. Dove, Shull, & Ferrara, 2000; Wood & Loeb, 2000; Ricotti & Shull, 2000; Ciardi, Bianchi, & Ferrara, 2002; Clarke & Oey, 2002; Fujita et al., 2003; Razoumov & Sommer-Larsen, 2006; Gnedin, Kravtsov, & Chen, 2008; Wise & Cen, 2009; Yajima et al., 2009; Yajima, Choi, & Nagamine, 2011; Paardekooper et al., 2011; Yajima et al., 2011). Estimated escape fraction values range from 10^{-5} to 1 depending, among others, on the mass, redshift, dust and gas distribution and SFR, with different studies giving different values. Since the correlation between halo properties and $f_{\mathrm{esc,ion,ISM}}$ is not well known, we choose our default value to be $f_{\mathrm{esc,ion,ISM}} = 0.02$, as in Gnedin, Kravtsov, & Chen (2008).

We have then tested the effect of the escape fraction of Lyα photons due to the ISM,

$f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}}$, which controls the fraction of Lyα photons reaching the IGM, as detailed in Equation 2.4. We use the same object and box of the previous tests, with $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}} = 0.01, 0.3$ and 0.99. Figure 3.10 shows the spectrum of the photons exiting the cube, which increases with increasing $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}}$ as more photons reach the IGM from the ISM. The effect of a different $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}}$ is to change the flux, while the shape of the spectrum is not altered. Similarly, the morphology of the surface brightness profile is not affected.

The escape fraction of Lyα photons from the ISM has been estimated in a large number of studies (e.g. Le Delliou et al., 2005, 2006; Davé, Finlator, & Oppenheimer, 2006; Nagamine et al., 2010; Ouchi et al., 2008; Dayal et al., 2009; Kobayashi, Totani, & Nagashima, 2007; Laursen, Sommer-Larsen, & Andersen, 2009; Dijkstra & Wyithe, 2010; Steidel et al., 2011; Schaerer et al., 2011; Forero-Romero et al., 2011; Yajima et al., 2011). Since the amount and type of dust in the ISM of a galaxy and its correlation with other halo properties is highly unknown, we choose a reference value for $f_{\mathrm{esc,Ly}\alpha,\mathrm{ISM}}$ of 0.3, similar to the one predicted by Dayal et al. (2009).

3.2.2 Effect of Input Lyα Spectrum

The radiative transfer of Lyα photons through the IGM is affected by their wavelength as they escape the ISM. In our default runs, we use a monochromatic line at 1215.67 Å, but we can expect the spectral shape of the photons emitted by the stellar sources to be distorted by their interaction with the ISM before entering the IGM. In fact, observations of LAEs at $z \sim 2-3$ show a non-monochromatic spectrum (Verhamme et al., 2007). Other authors (e.g. Dayal et al., 2009; Zheng et al., 2010; Dijkstra & Wyithe, 2010) use gaussian profiles with a width determined by a fraction of the circular velocity at the virial radius of the DM halo. Wider gaussian profiles lead to an easier escape of the Lyα photons from the surrounding medium. A lot of effort has gone into modelling more accurately the Lyα line profile for a range of parameters (Schaerer et al., 2011), but the exact shape of the spectrum and its connection to the galaxy properties are still unclear. Because of all these uncertainties, we choose to investigate the effect of the input Lyα spectrum simply by using a shifted monochromatic spectrum. All the tests were performed using the default configuration.

Figure 3.11 shows the spectrum of all the Lyα photons exiting the box for different values of the input Lyα monochromatic photons, i.e. with velocity shifts of 0 (reference case), 200, 400 and 800 km s^{-1}. Since the scattering cross-section is $\propto \Delta\nu^{-3}$, the larger the shift, the lesser the scatter suffered by the photons. The velocity field in the IGM also plays a role in determining the scattering probability of these photons. Since the velocity distribution of the IGM around the object has a complex structure, it is not possible to uniquely predict the velocity shift at which the photons stop being scattered. For our test object, as the photons are shifted by 200 km s^{-1}, the scattering is reduced by a factor of two, which in turn is reflected in a reduction of photons at 1215.67 Å. For a shift of 400 km s^{-1}, the scattering is reduced by an order of magnitude. Monochromatic spectra shifted by $\geq$ 800 km s^{-1} get hardly scattered. This also affects the surface brightness profiles, as shown in Figure 3.12, where the surface brightness in the four cases is plotted

for one of the side of the box (top-left map in Figure 3.2). As the velocity shift increases, the probability of scatter of photons decreases, leading to a more concentrated surface brightness profile. In addition, only a small fraction of the photons are scattered forming a low surface brightness halo. This can be seen more clearly in Figure 3.13 where the cross section of the maps is shown. The higher the velocity shift, the higher is the surface brightness value in the central pixel, improving the detectability of the source. Also, a larger fraction of the flux escapes from this pixel - 0.004% for no velocity shift case, 0.1% for 200 km s^{-1}, 1.5% for 400 km s^{-1} and 4% for 800 km s^{-1}. Therefore, the initial spectra of Lyα photons play an important role in determining the surface brightness profile of the object. Thus the velocity shift induced by the processing of the Lyα photons in the ISM has a very important effect on the observability of LAEs but quantifying that effect is beyond the scope of this study and will be investigated elsewhere.

From Lyα spectrum models of Schaerer et al. (2011), the peak of the spectra of photons escaping from the galaxy is at wavelengths where the velocity shift is equivalent to $\sim 1.5-2$ times the outflow velocities. The typical outflow velocities estimated from observations by Verhamme et al. (2007) is $\sim 150 - 200$ km s^{-1}, thus showing a Lyα spectrum peak at $200 - 400$ km s^{-1}. With a simplifying assumption that almost all the flux enters the IGM at these wavelength, we can estimate the expected changes in observed luminosity[2] at different $SB_{\rm th}$ due to this velocity shift. At $SB_{\rm th} = 10^{-21}$ ergs s^{-1} cm^{-2} arcsec^{-2}, for a velocity shift of 200 km s^{-1}, $f^{\rm eff}_{{\rm esc,Ly}\alpha,{\rm IGM}} = 0.01$ compared to 7×10^{-4} in our reference case with no velocity shift. At 400 km s^{-1}, the $f^{\rm eff}_{{\rm esc,Ly}\alpha,{\rm IGM}} = 0.1$. Thus detectability of a LAE at a specific $SB_{\rm th}$ can improve by orders of magnitude for realistic velocity shifts of $200 - 400$ km s^{-1} even with $\rm x_{ion} = 0$ (see also Dijkstra & Wyithe, 2010). Therefore, understanding the effect of the ISM of the galaxies along with their outflows is a crucial input for better modelling of LAEs. This topic will be addressed elsewhere in further detail.

3.2.3 Effect of IGM Ionization

Another important factor determining the observability of LAEs is the ionization level of the IGM outside the fully ionized region created by the source. The initial ionization fraction of our simulations is set to a uniform default value of $\rm x_{ion} = 0$ for convenience, but simulations of the reionization process indicate that the IGM might be substantially ionized at the redshifts studied in this thesis (e.g. Ciardi et al., 2012). The IGM ionization level is determined by a combination of a background radiation and local sources. Here we study how these two scenarios change the observability of our test source.

Uniform Ionization Distribution

To understand the effect of an ionizing background on Lyα RT, we repeat the reference simulation with an initial volume-averaged ionization fraction of $\rm x_{ion} = 0.5$ and 0.89. The first value corresponds to an epoch when only half of the IGM is ionized, while the second

[2]Note that a specific surface brightness threshold might select more than 1 pixel.

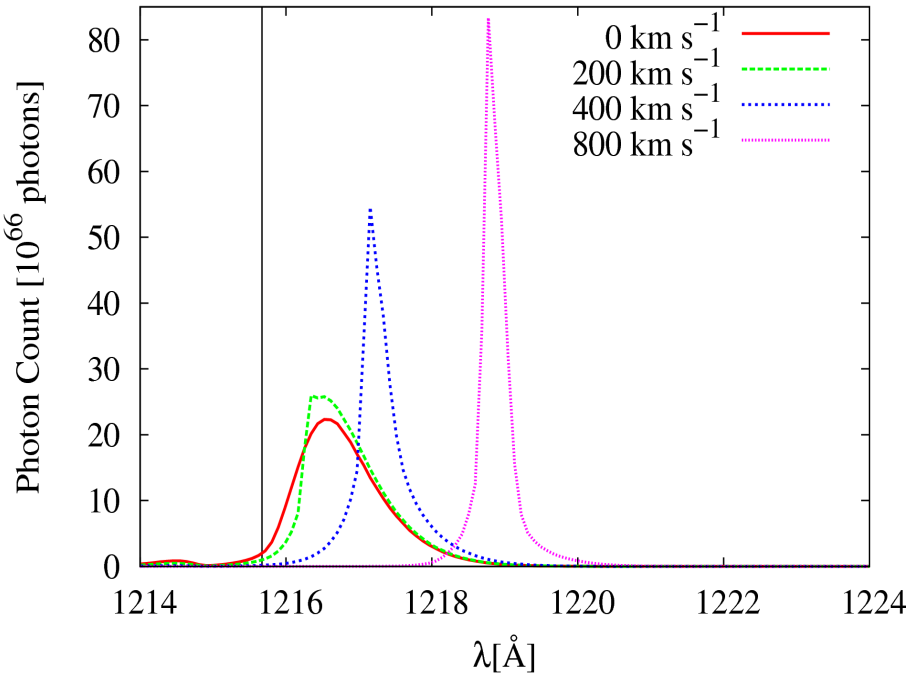

Figure 3.11: Spectrum of all the Lyα photons exiting the simulation box for different velocity shift of the input Lyα monochromatic spectrum. The lines refer to a velocity shift of 0 (red solid; reference case), 200 (green dashed), 400 (blue short dashed) and 800 (pink dotted) km s^{-1}.

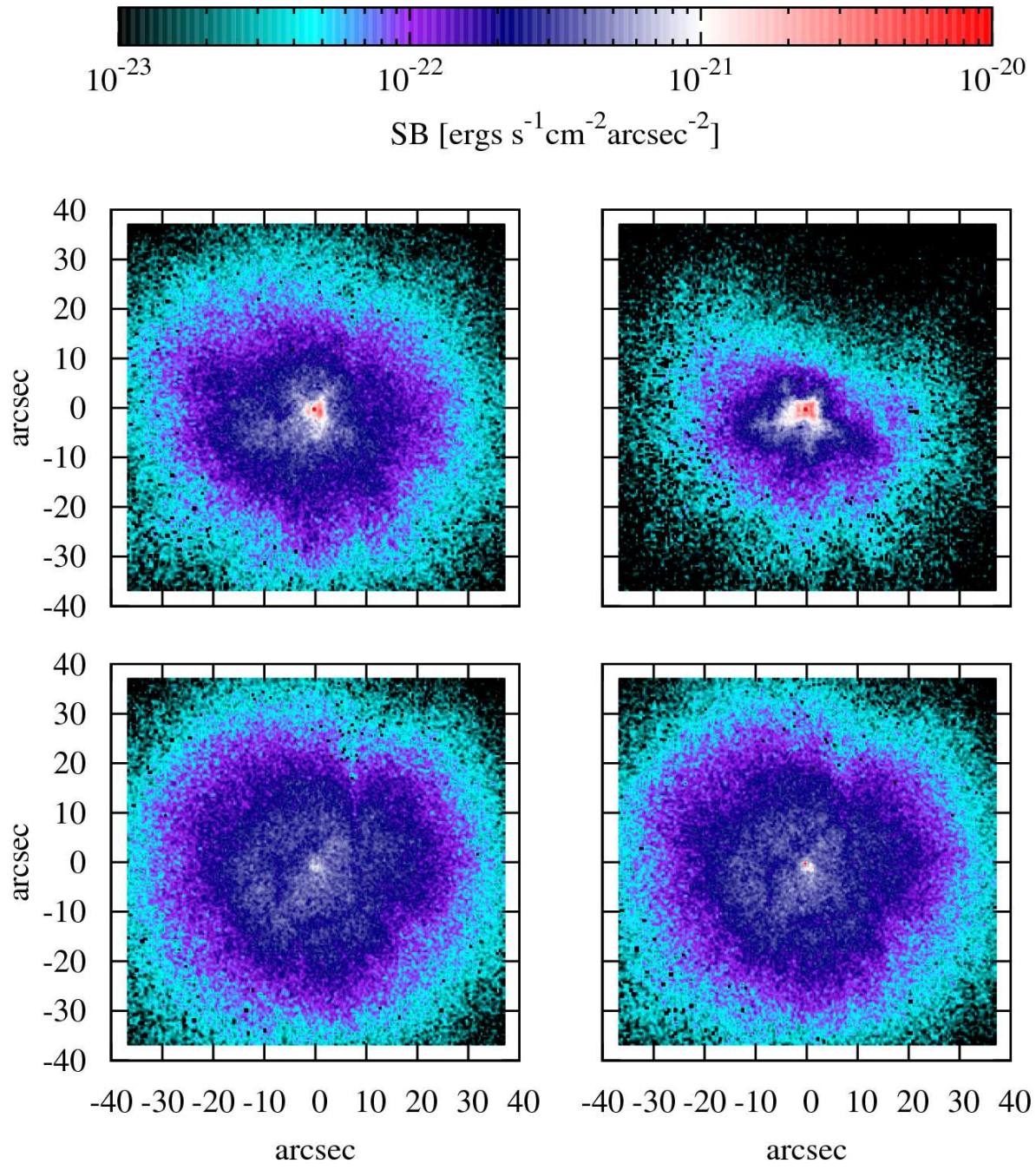

Figure 3.12: Surface brightness maps for one of the sides of the simulation box and for different velocity shift of the input Lyα monochromatic spectrum, i.e. 0 (bottom left panel), 200 (bottom right), 400 (top left) and 800 (top right) km s^{-1}.

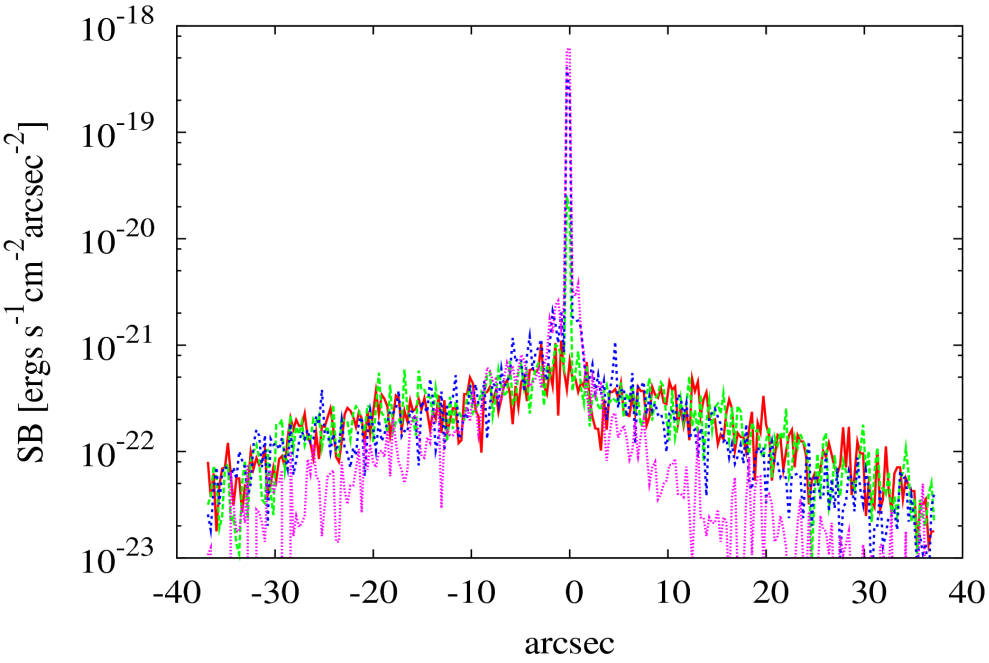

Figure 3.13: Cross section of the surface brightness profile for the maps in Figure 3.12. Plotted are the surface brightness values along the y axis for x axis value of 0. The lines refer to a velocity shift of the input Lyα monochromatic spectrum of 0 (red solid; reference case), 200 (green dashed), 400 (blue short dashed) and 800 (pink dotted) km s^{-1}.

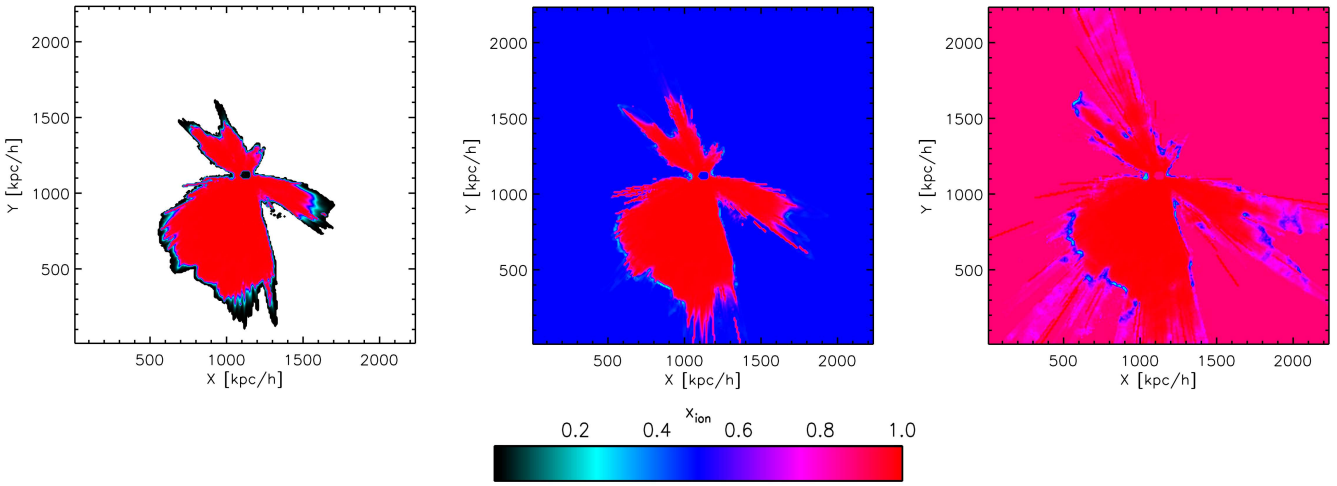

Figure 3.14: Slice through the final ionization structure for simulations with initial volume-averaged ionization fractions x_{ion}=0 (left panel; reference case), 0.5 (middle) and 0.89 (right).

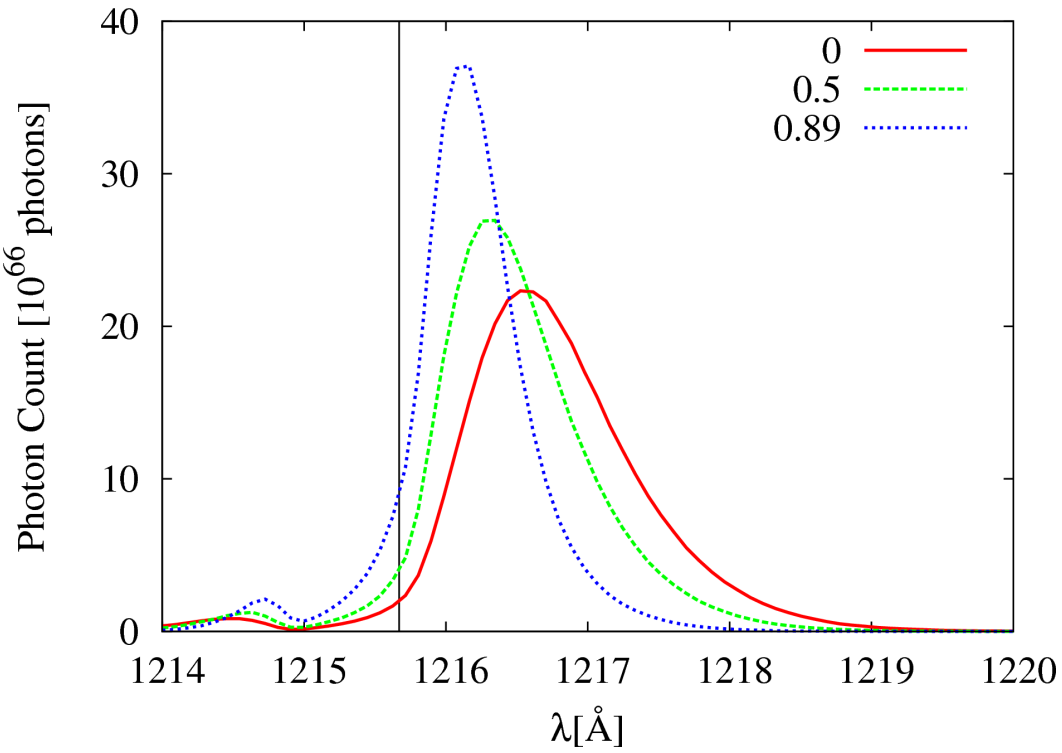

Figure 3.15: Spectrum of all the Lyα photons exiting the simulation box for runs with initial volume-averaged ionization fraction x_{ion}=0 (red solid line; reference case), 0.5 (green dashed) and 0.89 (blue dotted).

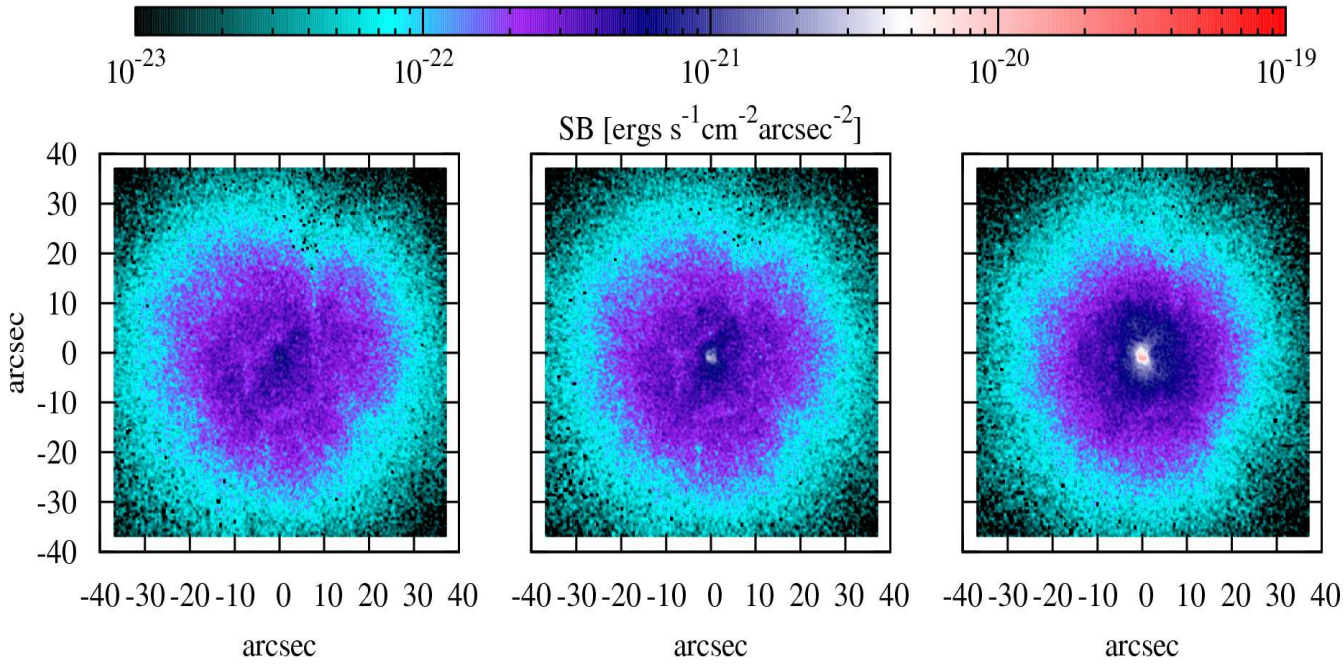

Figure 3.16: Surface brightness maps for one of the sides of the box for runs with initial volume-averaged ionization fraction x_{ion}= 0 (left panel; reference case), 0.5 (middle) and 0.89 (right).

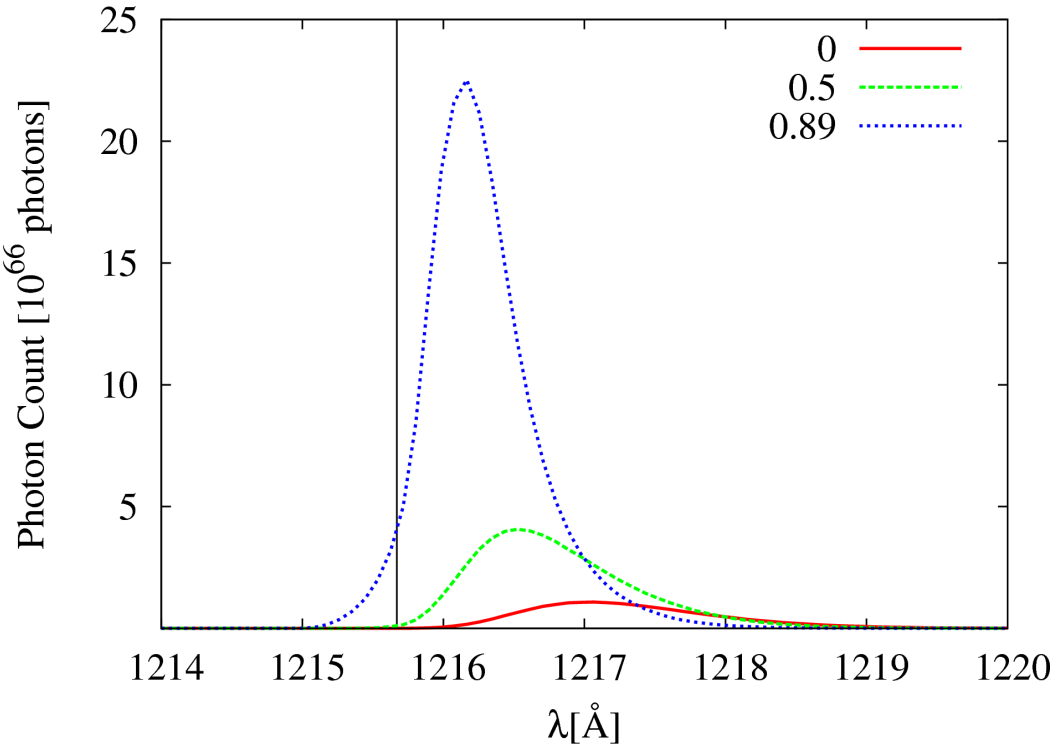

Figure 3.17: Same as Figure 3.15 but including scattering of Lyα photons by the column of neutral gas beyond the edge of the cube. The lines are for a volume-averaged ionization fraction x_{ion}=0 (red solid line; reference case), 0.5 (green dashed) and 0.89 (blue dotted).

is the value obtained from the reionization model $\mathcal{E}1.2 - \alpha 1.8$ described in Ciardi et al. (2012) (same as $35.12G128$ in Table 4.1) which is designed to be consistent with the existing observational constraints such as the Thomson scattering optical depth and the photo-ionization rate at $z < 6$.

Figure 3.14 shows a slice cut through the ionization structure at the end of the simulation for the three cases described above. As expected, the higher x_{ion} is, the larger is the fully ionized region produced by the source, although the differences are not dramatic because the high density filaments surrounding the source effectively confine the fully ionized region in all cases and determine its shape. Note that while the edges of the ionized region are controlled by the high density filaments in the IGM, the inclusion of ionizing flux from galaxies in those high density clumps could lead to a different ionization structure (as shown in Sec. 3.2.3).

As a larger ionized region results in less scatter for the Lyα photons, we expect an easier escape in this case. This is shown in Figure 3.15, where the spectrum of all the Lyα photons exiting the simulation box is plotted. Due to the lower number of scatterings undergone by the Lyα photons for higher values of x_{ion}, the peak shifts to bluer wavelengths and the photon count at the peak increases while reducing the width of the line profile. The surface brightness maps for one of the sides of the box for the different x_{ion} values is shown in Figure 3.16. As x_{ion} increases, the surface brightness profile becomes more concentrated while improving detectability. From the maps, we can infer that observations at $SB_{\rm th} = 3 \times 10^{-21}$ ergs s^{-1} cm^{-2} arcsec^{-2} would appear as a point source to the observer. For this viewing direction, the luminosity (effective escape fraction) of the point source detected for $x_{ion} = 0.5$ and 0.89 is 8.4×10^{38} ergs s^{-1} ($f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = 0.001$) and 3.7×10^{40} ergs s^{-1} ($f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = 0.04$), respectively, compared to no detection at $x_{ion} = 0$. More specifically, if we were to consider the maps from all six faces of the cube, at $x_{ion} = 0.89$ the object would be detected as a point source from all the six directions with $f^{\rm eff}_{\rm esc,Ly\alpha,IGM}$ in the range $0.004 - 0.6$, compared to only two detections as point source in six maps at $x_{ion} = 0$ with $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = 0.008$ and 0.09. Thus at high ionization fractions most of the sources can be detected as point sources at lower surface brightness thresholds compared to our reference case. Our results appear similar to Zheng et al. (2010) who find that even at a mean neutral fraction of 10^{-4}, only a small fraction of the photons (8-33 % of the intrinsic luminosity) escape as a point source.

Because this thesis is primarily focused on the effect of the IGM in the immediate surroundings of a source of Lyα photons, we do not follow the propagation of the photons to the observer. To do so, we would need much larger boxes, losing the resolution necessary to properly account for the scales we are mainly interested in here. Nevertheless, the IGM outside our simulation box is bound to scatter the photons in an amount dependent on their frequency when they escape the box. The optical depth τ due to scattering away from the line-of-sight to the observer (at $z = 0$) for a LAE at $z = z_s$ is given by (same as Equations 1.1 and 1.2):

$$\tau \simeq 6.02 \times 10^5 \; x_{\rm HI} \left(\frac{1 + z_s}{10} \right)^{3/2}, \tag{3.1}$$

for photons entering the IGM with $\lambda \lesssim \lambda_\alpha = 1215.67$ $\mathring{A}$ (e.g. Gunn & Peterson, 1965; Barkana & Loeb, 2001) and:

$$\tau = 2.9\, x_{\rm HI} \left(\frac{\Delta v}{600 \text{ km s}^{-1}}\right)^{-1} \left(\frac{1+z_s}{10}\right)^{3/2} \tag{3.2}$$

for photons entering the IGM with $\lambda = \lambda_\alpha\,(1 + \Delta v/c)$ (e.g. Miralda-Escude & Rees, 1998; Dijkstra & Wyithe, 2010). Here, $x_{\rm HI}$ is the neutral fraction of the IGM. For $x_{\rm HI} = 1$ at $z_s = 7.7$, the optical depth $\tau \sim 1$ for $\Delta v \sim 1400$ km s^{-1} which would be a comoving distance of 8.6 h^{-1}Mpc from the source which is much larger than our cube sizes. Even in the absence of peculiar velocities, beyond approximately this distance from the source, the Lyα photons would be redshifted away from resonance thus not affecting the Lyα surface brightness profles. But in some cases, especially for low mass objects for which smaller boxes are extracted, the photons emitted by the source do not get redshifted out of resonance when they reach the edge of the box and thus the neutral gas outside it might lead to further modification of the Lyα SB profile. As a simple estimate of the scattering further suffered by the Lyα photons due to the column of neutral gas outside the cube we can use Equation 3.2. The effect of this scattering on the spectra of photons escaping the box for our test object is shown in Figure 3.17. For $\rm x_{ion} = 0$, the τ integrated over the whole spectrum is 2.9, but as the ionization increases, the τ decreases to 1.8 at $\rm x_{ion} = 0.5$ and 0.55 at $\rm x_{ion} = 0.89$ leading to more photons being observed. As we can see, the scattering beyond the cube can be significant for low $\rm x_{ion}$ values leading to lower SB values per pixel extending over a larger angular size. Thus we would need to go to deeper surface brightness cuts to observe the total luminosity of the object. This would lower the $SB_{\rm th}$ values we quoted in our results section but the general trends we described (e.g. spread and increase in $f^{eff}_{\rm esc,Ly\alpha,IGM}$ for lower $SB_{\rm th}$) remain unchanged. For the case of high $\rm x_{ion}$, the scattering beyond the cube is very low retaining the point source like appearance of the SB profile seen in Figure 3.16. Since we don't do a proper RT transfer of photons beyond the edge of the cube, a direct comparison is beyond the scope of this study. As seen from our reionization simulations in Ciardi et al. (2012), in a realistic case the ionization fraction at $z = 7.7$ is not zero, but close to 0.89 and increases with decreasing redshift. At these high $\rm x_{ion}$ values, scatter beyond the cube is not significant.

Another factor which would reduce the significance of the scatter is the velocity shift discussed in Section 3.2.2. A shift of 200 km s^{-1} would slightly reduce the effective τ to 2.8 even for $\rm x_{ion} = 0$ and at 400 km s^{-1}, the effective τ is reduced to 2.3. Thus for most realistic configurations at $z = 7.7$, we expect the scatter inside the cube close to the source to be more important than the one due to the neutral gas outside. In any case, we do not include this extra scatter in the rest of our analysis.

Clustering of Nearby Sources

While a roughly uniform ionization fraction could be induced by an ionizing background, photons from neighboring haloes clustered around a source would induce further fluctuations in the gas ionization structure. To investigate the impact this has on our reference

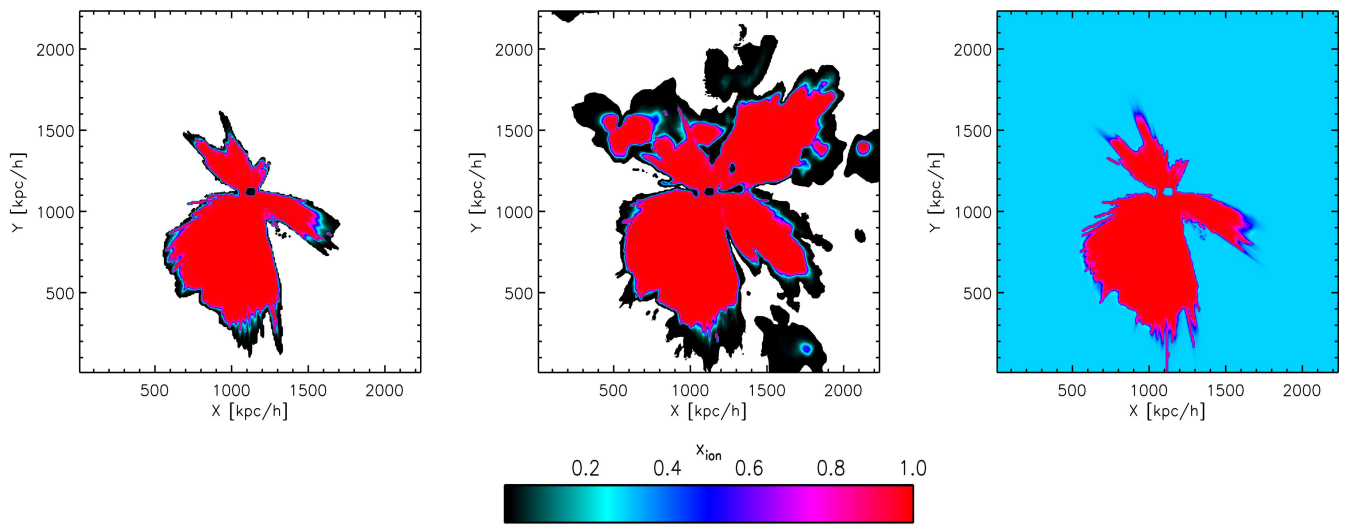

Figure 3.18: A slice though the ionization structure for RT simulations showing the effects of clustering of ionization sources. The left side shows the ionized region in the IGM produced by the central source alone, the middle plot shows the ionized region produced by all the sources in the simulation box (case α) and the right side plot shows the ionization structure produced by a central source on a uniform partially ionized IGM (case β). The volume averaged ionization fraction in the middle and right panel is the same. see text for further details.

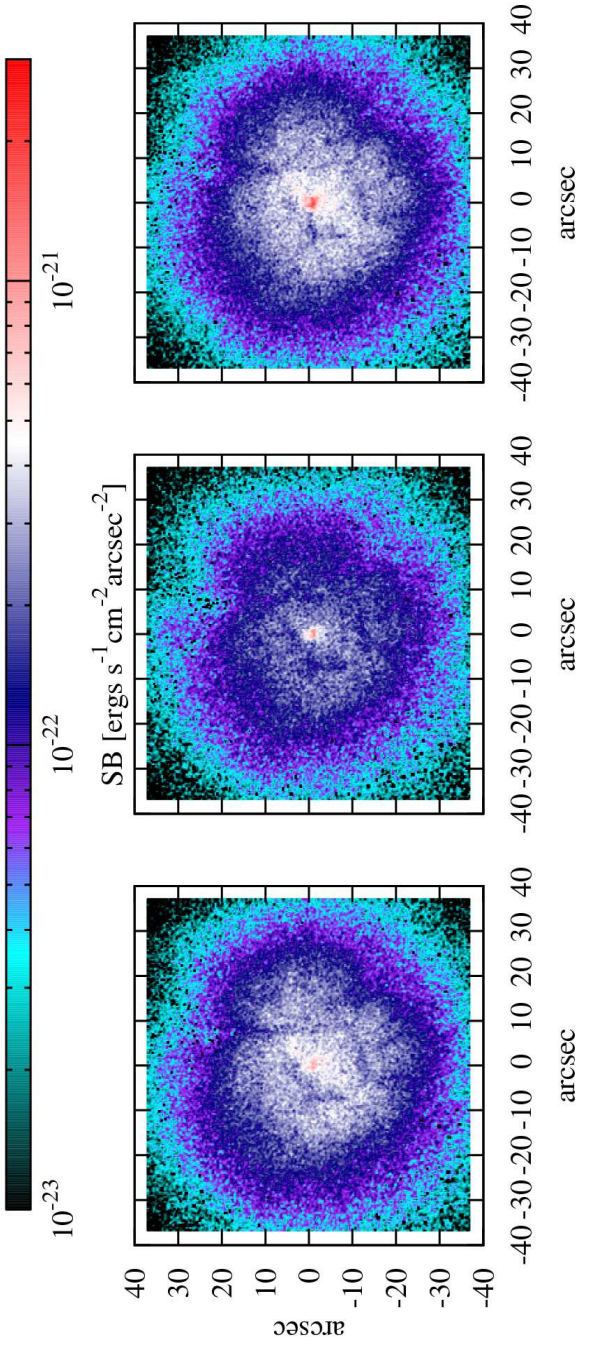

Figure 3.19: Surface brightness maps for one of the sides of the simulation box for the same simulations of Figure 3.18.

simulation, we propagate ionizing photons from all the stellar sources present in the cube (referred to as 'case α'), rather than only from the central source. Figure 3.18 shows a slice through the ionization structure obtained in the case α (middle panel) along with that from the reference simulation (left panel). As expected, the ionization bubble around the object is much larger and complex when the effect of all the sources is considered. In the case α, the final value for the volume averaged ionization fraction is $x_{ion} = 0.340$, compared to 0.047 of the reference case. As a comparison, test simulations in the previous sub-section with initial ionization levels of $x_{ion} = 0.5$ and 0.89 produce final levels of $x_{ion} = 0.547$ and 0.937, respectively. To understand how the ionization structure due to clustered sources affect the propagation of Lyα photons compared to a uniform ionization, we do another test run (referred to as 'case β') where the initial ionization level is set to a uniform value of 0.293 and then we follow the propagation of photons only from the central source. The final ionization fraction obtained is similar to the one obtained with the clustered sources (case α), i.e. ~ 0.340. Case β is shown in the right panel of Figure 3.18. We can see that the ionization structure shows significant differences for case α and case β. In case β the bubble where the gas is fully ionized is smaller than in case α, while the situation is reversed in the region beyond the bubble with some fully neutral gas still present in case α. These details of the ionization distribution have an effect on the Lyα RT. Note that in both the test cases (and the reference run), the Lyα photons are emitted only by the central source.

Figure 3.19 shows the surface brightness maps for the three cases discussed above. Both case α and case β show a more concentrated SB profile compared to the reference run as expected for a lower x_{ion}, but the details in the surface brightness maps differ due to the distribution of neutral gas. In particular, all the maps (from the six faces of the cube) of case β have a more concentrated structure compared to case α. Since such an appearance in the surface brightness improves detectability, case β is expected to have a better detectability for more viewing directions. The reason for the point source like appearance of case β is due to the higher ionization level, compared to case α, for the gas outside the fully ionized bubble. Since these ionization bubbles are not large enough to redshift the Lyα photons out of resonance, the ionization level outside these bubbles determine the surface brightness maps. At $SB_{\rm th} = 10^{-21}$ ergs s^{-1} cm^{-2} arcsec^{-2} case α has $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} \sim 0.001$, while case β has $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} \sim 0.004$, compared to the reference case of ~ 0.0007. It might seem that clustering of neighbouring sources does not change the observability of the LAE very much, but in reality the flux is redistributed in different directions due to the different ionization structure in the IGM. For the same $SB_{\rm th}$, another direction (top right panel in Fig. 3.2 instead of our reference direction top left panel) in case α has $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} = 1.5$ as a large amount of flux escape through line-of-sight. The escape fraction in this direction for case β is $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} \sim 0.7$ and for the reference case is $f^{\rm eff}_{\rm esc,Ly\alpha,IGM} \sim 0.6$. For reference, the input luminosity is $L^{\rm esc}_{\rm Ly\alpha} = 9.2 \times 10^{41}$ ergs s^{-1}. Thus one has to keep in mind that even though both neighbouring sources and uniform background flux improves detectability of LAEs by increasing the mean ionization fraction of IGM in the region, the detailed ionization structure around a source plays a crucial role in determining the surface brightness profiles in different directions.

3.3 Effects on Estimates of the Reionization History

Keeping in mind the discussions in the previous Sections, we now turn to study how structure in the IGM affects the methods which use the surface brightness profiles of LAEs to estimate the mean ionization fraction of the Universe. One of the most popular methods is to calculate the luminosity function of LAEs and determine how it changes with redshift (e.g. Malhotra & Rhoads, 2004; Dayal et al., 2009; Kashikawa et al., 2011). Using analytic models the above authors calculate the redshift evolution of the τ (employing equations similar to Eq. 1.2) assuming a specific reionization history. The intrinsic LF of LAEs is calculated from halo mass functions and assuming that the luminosity of an object is proportional to its mass; then it is attenuated by $e^{-\tau}$ to take into account IGM absorption/scattering and derive the expected LF. A comparison with the observed LF provides a constraint on the actual τ and thus on the gas neutral fraction in the IGM. But these methods rely on the assumption that observations of LAEs obtain the total flux from the source detected as a point source. Some recent observations though have detected faint extended SB profiles (Steidel et al., 2011), as seen in our simulations. As we have seen, the probability of detection of a LAE depends on the $SB_{\rm th}$, especially at low $x_{\rm ion}$. Here we investigate how $SB_{\rm th}$ affects the LAE LFs considering only changes in $x_{\rm ion}$. A direct comparison to observations is not made considering the uncertainties in the different parameters in our models, but the qualitative effect is applicable to the whole parameter range and thus is important. A more detailed comparison is part of a future study.

We simulate the 45 most massive objects in L05 using $x_{\rm ion}$=0, 0.5 and 0.89 as described in Section 3.2.3. For each object, we calculate the six SB profiles and their respective observed luminosities $L_{\rm obs}(> SB_{\rm th})$, which yields a total of 270 data points (LAEs) at each $SB_{\rm th} = 5 \times 10^{-20}$, 10^{-20}, 10^{-21} and 10^{-25} ergs s^{-1} cm^{-2} arcsec^{-2}. We consider the six sides of a simulation cube as part of six separate observational fields with comoving side of 5 Mpc h^{-1} each. Structure in the IGM leads to differences in the surface brightness profiles for the six sides (as seen in Fig. 3.2), which in turn leads to a scatter in the observed LFs calculated from the different fields. Therefore, LFs are calculated separately for each of the observational fields and then averaged in each luminosity bin to obtain the averaged luminosity functions. The averaged LFs calculated for the different $SB_{\rm th}$ at each $x_{\rm ion}$ are shown in Figure 3.20.

First we focus on $x_{\rm ion} = 0$, to understand the effects of surface brightness thresholds on detections (as seen in Figure 3.7) which in turn determines the luminosity functions. At $SB_{\rm th} = 5 \times 10^{-20}$ ergs s^{-1} cm^{-2} arcsec^{-2}, only one object is detected with an observed luminosity of $L_{\rm obs}(> SB_{\rm th}) = 3 \times 10^{38}$ ergs s^{-1}. Moving to deeper surface brightness thresholds of 10^{-20} ergs s^{-1} cm^{-2} arcsec^{-2} leads to more detections (5 LAEs). It also has higher observed luminosities than the previous case, thus moving the LF to higher values in both axes. Going one order of magnitude deeper in surface brightness leads to a huge increase in the number of detections ($\sim$ 130 LAEs). This gives rise to a shift in the LF of about two orders of magnitude in both luminosity and number density. Deeper surface brightness thresholds lead to more detections and a larger fraction of the true flux in that direction being detected. By $SB_{\rm th} = 10^{-25}$ ergs s^{-1} cm^{-2} arcsec^{-2}, all the flux in all the

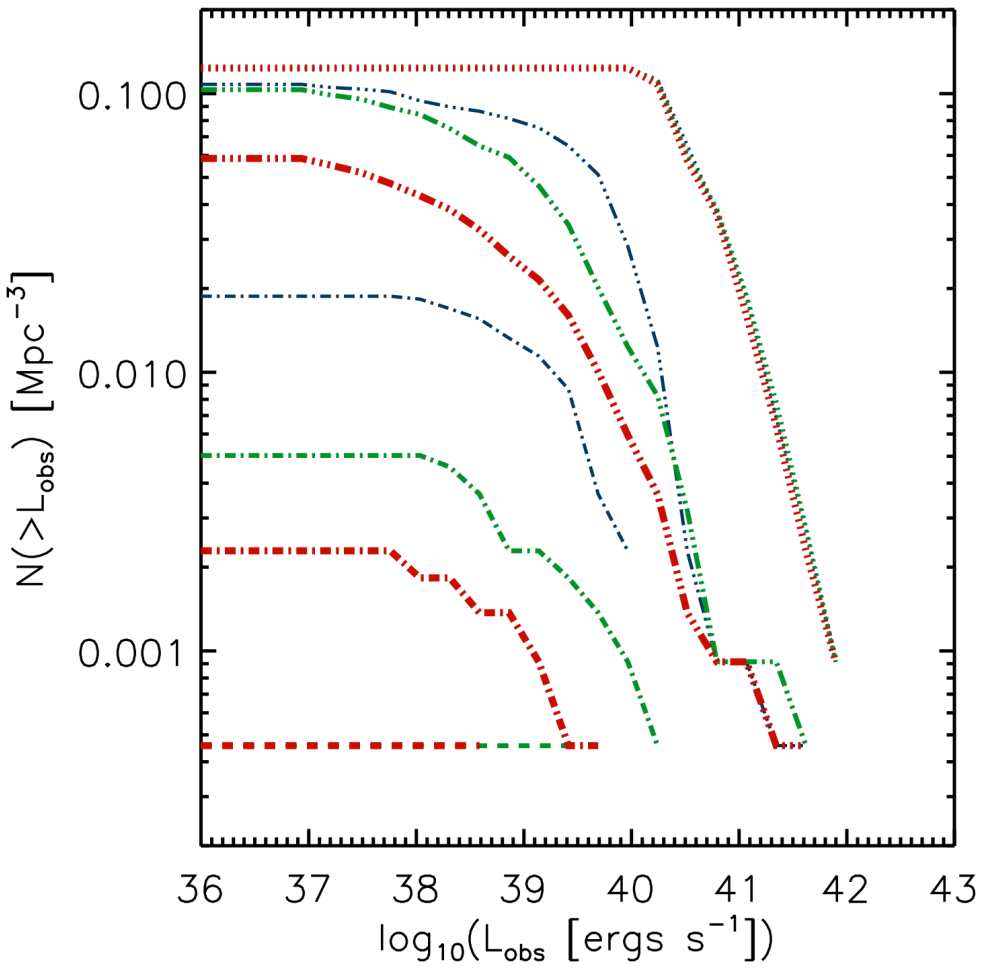

Figure 3.20: Luminosity functions calculated for the 45 most massive objects in L05 for different surface brightness cuts. The runs are for initial volume averaged ionization fraction x_{ion}=0 (red), x_{ion} = 0.5 (green) and x_{ion} = 0.89 (blue). The SB thresholds are $SB_{th} = 10^{-25}$ (dotted lines), 10^{-21} (triple dot dashed lines), 10^{-20} (dot dashed lines) and 5×10^{-20} (dashed lines) ergs s^{-1} cm^{-2} arcsec^{-2}.

LAEs is observed returning the intrinsic LF of the simulated LAE sample. It should be kept in mind that the value of $SB_{\rm th}$ needed to observe the full flux depends on the intrinsic luminosity of the sources and neutral IGM structure around it.

In the real Universe, both the mean ionization fraction and the intrinsic properties of galaxies change. This leads to a complex evolution of the LF of LAEs with redshift. In addition, the different $SB_{\rm th}$ from real observations add another level of complexity to the redshift evolution of observed LFs. To help constraining $\rm x_{ion}$ with observations of LFs, here we try to understand how the LF depends on the ionization fraction of the IGM for different choices of $SB_{\rm th}$. For $SB_{\rm th} = 10^{-25}$ ergs s^{-1} cm^{-2} arcsec^{-2}, the luminosity functions for different values of $\rm x_{ion}$ are equivalent, because all the flux from the sources is detected. For lower surface brightness cuts instead, the number of detections changes with $\rm x_{ion}$, leading to an increasing number density of LAEs at a given luminosity for higher $\rm x_{ion}$. The difference between LFs at different $\rm x_{ion}$ is higher for smaller values of $SB_{\rm th}$. It should be noted that shallow detections (e.g. 10^{-20} ergs s^{-1} cm^{-2} arcsec^{-2}) at higher $\rm x_{ion}$ (e.g. $\rm x_{ion} = 0.89$) have similar LFs as deeper but incomplete detections (e.g. 10^{-21} ergs s^{-1} cm^{-2} arcsec^{-2}) at lower $\rm x_{ion}$ (e.g. $\rm x_{ion} = 0$). Thus, when comparing luminosity functions at different redshifts, the surface brightness threshold of the observations should be taken into account to obtain a realistic estimate of the decrease in the number density of LAEs. Another point to note is that discrepancies between LFs for different values of $SB_{\rm th}$ are smaller at higher ionization fractions (e.g. $\rm x_{ion} = 0.89$). Thus only mild differences in LFs for orders of magnitude in $SB_{\rm th}$ at a specific redshift can be used as an evidence for high $\rm x_{ion}$. At high $\rm x_{ion}$, absorption outside the edge of the cube (as discussed in Section 3.2.3), which has not been included in this test, would lead to an even larger gap between LFs for different $SB_{\rm th}$. Nevertheless one has to keep in mind that velocity shifts discussed in Section 3.2.2 (e.g. due to outflows from the galaxy) could also lead to milder changes in luminosity functions with changing $SB_{\rm th}$.

Due to the difficulties in obtaining deep observations of LAEs (refer to Section 3.1), we look into the possibility of gathering additional information from stacked profiles. This was recently done by Steidel et al. (2011) at $z = 3$ to estimate the low surface brightness extended emission present in individual non-detections. In addition, also Zheng et al. (2010) looked at estimating total Lyα luminosities from stacked profiles for different mass and luminosity bins. In our study, we investigate the shape of the stacked surface brightness profiles without binning for halo properties, to understand how the profiles are changed by $\rm x_{ion}$. The data set used is the same one discussed above. In general, for $\rm x_{ion} = 0$ the individual profiles show a flatter surface brightness profile with substantial structure due to scattering through a mostly neutral but structured IGM. For higher initial ionization fractions, the strength of the point source increases as more photons escape with less scattering resulting in a steeper surface brightness profile. These characteristics can be exploited to study the mean ionization fraction of the Universe from stacked LAE SB profiles.

Figure 3.21 shows the stacked surface brightness maps from the 270 simulated maps of the 45 galaxies described previously. The top (bottom) panels refer to the mean (median) values in each pixel of the stack. The mean value is generally affected by the outliers when

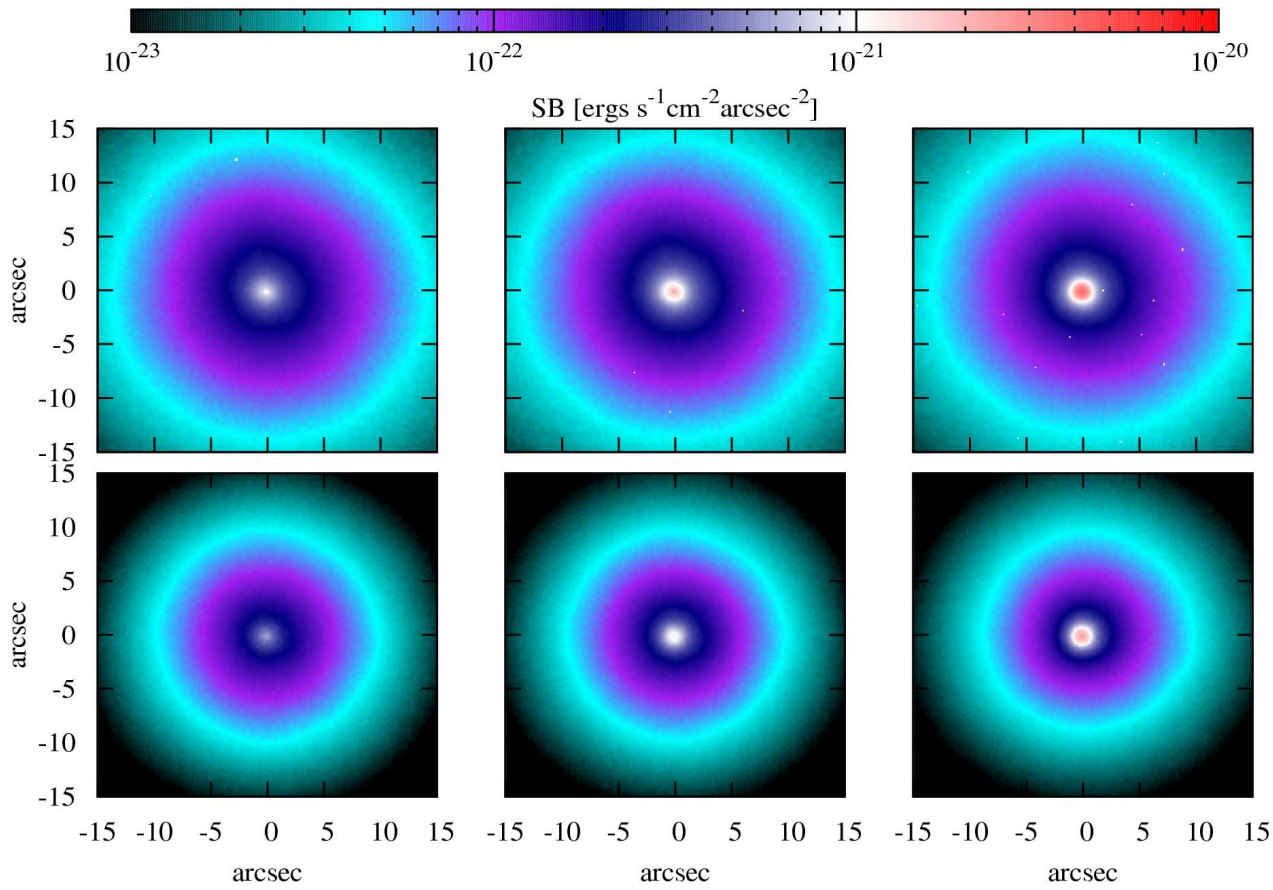

Figure 3.21: Stacked surface brightness profiles for initial volume averaged ionization fractions of x_{ion} =0 (left panels), 0.5 (middle) and 0.89 (right). The top (bottom) panels refer to the mean (median) values in each pixel of the stacked surface brightness profiles.

the distribution is not Gaussian. In our stacks, the higher mass haloes act as outliers being brighter and more point-source-like due to larger ionized bubbles around them, but are fewer in number compared to the majority of SB maps of low mass haloes. This leads to the mean profile being steeper and brighter than the median profile at each x_{ion}. But it is important to note that in both mean and median stacks, the trend of the stacked surface brightness profiles with x_{ion} is the same.

As we have seen in Section 3.2.3, the higher x_{ion} is, the steeper is the surface brightness profile of a LAE. This trend is more important for haloes of lower masses, as the ionization bubble around these objects are smaller due to lower ionizing photon flux from these sources, leading to a more diffuse SB profile at low x_{ion}. As x_{ion} increases, a larger fraction of the photons from these haloes escape unscattered, leading to a more concentrated SB profile. For high mass haloes, even at low $x_{ion} = 0$, the photons escape with lesser scatter due to larger ionized regions around them, leading to a steeper SB profile. Due to the shape of the halo mass function though, there is a much larger number of low mass haloes in a comoving volume of the Universe. Therefore, when stacking SB profiles of LAEs in a volume, the average of the stack is dominated by the low mass haloes. Thus the stacked SB profile at low $x_{ion} = 0$ is flat. The stacked profile has no structural details, because the profiles in individual SB maps have a random distribution, thus averaging out in the stack. At high ionization fractions of $x_{ion} = 0.5,\ 0.89$, the photons escape easier in all directions leading to a steeper SB profile in all directions. This also appears when stacking SB profiles, with the stacked SB profile becoming more concentrated with increasing x_{ion}. Thus this method of stacking SB profiles can be used to estimate the steepness of the profile and in turn the volume average mean ionization fraction of the Universe at that redshift. As discussed a few paragraphs above, the velocity shifts could also lead to the steepening of the individual SB profile even at $x_{ion} = 0$, thus reducing the difference in concentration between stacked profiles for different x_{ion}. But absorption due to IGM outside the cube at $x_{ion} = 0$ would lead to flattening of individual SB profiles. This would help in differentiating between x_{ion} in stacked SB profiles.

In short, a lot of information about the galaxies and their surrounding IGM can be obtained from the deep observations of LAE surface brightness profiles. But direct comparison of our simulations with observations is not possible due to the uncertainties in parameters involved in this model as shown in this thesis. A more detailed modelling and comparison with observations is the work of a future study.

3.4 Discussion

The aim of this chapter was to study how structure in the IGM affects the observability of LAEs. We simulate LAEs and examine the six different surface brightness maps for each object. We find that there is structure and variation in the different maps of the same object. This leads to a spread in the observed luminosity in each direction. Our simulations include diffuse radiation from the ionized IGM. But this contributes to only $\sim 3\%$ of the source luminosity and generally remains very scattered.

For the case of observing the total luminosity, the mean escape fraction of Lyα photons by the IGM is 0.73 with a scatter of $\sigma = 0.18$. Due to differences in the maps, we obtain a different luminosity for the source for different surface brightness thresholds. The shallower the surface brightness cut, the lesser number of objects detected and more important is the dependence on IGM structure. This effect improves the observability of low mass galaxies as they could have a larger ionized region around them than in the uniform IGM case. The different surface brightness thresholds also lead to differences in luminosity functions for the same sample of objects.

Two of the main parameters in our study are $f_{\mathrm{esc,ion,ISM}}$ and $f_{\mathrm{esc,Ly}\alpha\mathrm{,ISM}}$. The surface brightness values we obtain for each object depends heavily on the choice of these parameters. But in the current situation with a lot of uncertainties on the sources of Lyα, the ISM properties and their link to galaxy, we have to be satisfied with a parameter study.

Clustering of sources is another important effect which improves the detectability of objects especially at the low mass end in high neutral fractions. Neighboring sources enlarges the ionized region making the surface brightness profile of the sources steeper. Due to the time consuming nature of these simulations, we perform it only in one case. But even for clustering situations, the structure in the IGM close to the source is very important.

Both velocity of the photons exiting the ISM into the IGM and the ionization level of the IGM play an important role in the detectability of LAEs. At redshifts where the universe is highly neutral, the outflows from the galaxy would greatly improve the observability of these objects (Dijkstra & Wyithe, 2010). But due to the uncertainty in the link between the galaxy and outflows, this would add a challenge to estimating the ionization fraction of the IGM at high redshifts. Strong outflows could help the Lyα photons escape with very little scattering which suggests that the imprint of reionization on the statistics of LAEs is weaker than thought previously.

As our cube sizes are small compared to the distance from the source to the observer, we estimate the optical depth due to scattering in the IGM gas outside the cube using an analytic expression from Dijkstra & Wyithe (2010). We can see that the optical depth is large for the neutral case specially for Lyα photons emitted at rest wavelength. But a lower ionization fraction outside the cube reduces the optical depth.

The current observational limit of SB threshold is about 10^{-19} ergs/s/cm^2/arcsec^2 (Steidel et al., 2011). This was achieved through stacking of 100 profiles at $z \sim 3$. Achieving these limits at higher redshifts is harder, especially for individual observations. Thus stacking is a good method also shown by Zheng et al. (2010) who simulated stacked profiles at $z \sim 5.7$ and showed the dependence of surface brightness profiles on mass, etc. In our work, we stack surface brightness profiles for different $\mathrm{x_{ion}}$ and find that the lower the ionization, the flatter the surface brightness profiles. This can be used to estimate the ionization level of the IGM. Even though it has not been shown here, outflows would again lead to steeper profiles at low ionizations. Thus one needs to look at the relative steepening of mean/median surface brightness profiles at different redshifts to estimate $\mathrm{x_{ion}}$ under the assumption that the outflow properties do not change with redshift.

Chapter 4

Simulations of Reionization and Clumping Factors

In the previous chapter, we understood the importance of IGM ionization structure in determining the observability of an LAE. We also found that the initial ionization state of the gas due to background flux leads to easier escape of Lyα photons towards the observer. Also, clustering of ionizing sources lead to a more complex ionization structure around the source which affects the Lyα RT in imperceptible ways. Thus for proper modelling of LAEs we need the ionization structure around these galaxies as input to our RT simulations. We can use reionization simulations to compute the overall reionization history of the region which can be used to estimate the mean ionization fraction expected around the galaxy due to background flux and clustering of sources.

Simulating the reionization history is a complex task due to the involvement of a large number of crucial factors spanning orders of magnitude in spatial and mass scales (refer to e.g. Barkana & Loeb, 2001; Ciardi & Ferrara, 2005; Barkana & Loeb, 2007; Meiksin, 2009; Morales & Wyithe, 2010, and the references there in.). Due to the patchy nature of reionization, one needs to simulate a large comoving volume ($\geq 100\ h^{-1}$Mpc) to representatively sample the dark matter halo distribution as well as to contain the large ionized bubbles made by them with characteristic sizes of tens of Mpcs. Similarly, high mass resolution is needed to resolve the low mass galaxies ($M_{\rm DM} > 10^8$ M$_\odot$) which are thought to be dominating the ionizing photon contribution for reionization. Lyman-limit systems in the IGM regulate photon path lengths and delay the completion of EOR, thus being important objects to be resolved for accurate modelling of the late history of reionization. To add to this, the Jeans mass in the heated neutral IGM (a few 100 K) is a few times 10^6 M$_\odot$ which sets the mass scale of neutral gas clumps. Resolving all this needs simulation volumes of $\sim 100\ h^{-1}$Mpc with spatial scales of a few kpcs.

Simulating N-body dynamics of DM particles, hydrodynamics of gas (with H and He) and radiative transfer of ionizing photons from all the sources with different spectra in a simulation with the required mass resolution is a Herculean task with computational limits dictating the resolution one can achieve. To alleviate this problem it is possible to simulate large computational volumes at a limited resolution, and employ sub-resolution

prescriptions to include the physics which would be otherwise missed. A typical example is the adoption of a clumping factor, which allows to include in the calculation of the ionization state of the gas the effect of small scale high density regions. This has been done for a large number of reionization simulations (e.g. Iliev et al., 2007; McQuinn et al., 2007; Ciardi et al., 2012).

In grid based reionization simulations with gas consisting of only H, the ionization balance averaged over a cell can be written as (e.g. Kohler, Gnedin, & Hamilton, 2007):

$$\frac{d}{dt}\langle n_{\rm HI}\rangle = -3H\langle n_{\rm HI}\rangle - C_{\rm I,HI}\langle n_{\rm HI}\rangle\langle \dot{n}_\gamma\rangle + C_{\rm R,HII}\langle \alpha_{\rm R,HII}\rangle\langle n_{\rm e}\rangle\langle n_{\rm HII}\rangle \quad (4.1)$$

where $n_{\rm HI}$ and $n_{\rm HII}$ are the number density of neutral and ionized hydrogen, $n_{\rm e}$ is the number density of electrons, H is the Hubble parameter, $\dot{n}_\gamma$ is the ionizing photon rate and $\alpha_{\rm R,HII}$ is the recombination coefficient for HII. The angle brackets represent the mean value of the true underlying distribution the cell volume would have if the spatial resolution were enough to fully resolve IGM structures down to the smallest relevant scales. $C_{\rm I,HI} = \langle n_{\rm HI}\dot{n}_\gamma\rangle/\langle n_{\rm HI}\rangle\langle\dot{n}_\gamma\rangle$ is the clumping factor of HI and $C_{\rm R,HII} = \langle \alpha_{\rm R,HII} n_{\rm e} n_{\rm HII}\rangle/\langle\alpha_{\rm R,HII}\rangle\langle n_{\rm e}\rangle\langle n_{\rm HII}\rangle$ is the clumping factor of HII.

Low resolution gridding of the density field cannot resolve all the small scale gas clumps in a region which are at scales below the grid resolution. Depending on the distribution of gas (neutral/ionized) inside a cell, the ionization balance calculated using the density of the cell might be incorrect. Therefore, the clumping factors $C_{\rm I,HI}$ and $C_{\rm R,HII}$ can be used to estimate the true ionization and recombination rates in the cell due to unresolved small scale high density regions.

A large number of efforts have gone into calculating the clumping factor of different species and their evolution (e.g. Giroux & Shapiro, 1996; Gnedin & Ostriker, 1997; Miralda-Escudé, Haehnelt, & Rees, 2000; Iliev, Scannapieco, & Shapiro, 2005; McQuinn et al., 2007; Trac & Cen, 2007; Kohler, Gnedin, & Hamilton, 2007; Pawlik, Schaye, & van Scherpenzeel, 2009; Raičević & Theuns, 2010). While some works adopt an analytic approach (e.g. Giroux & Shapiro, 1996; McQuinn et al., 2007), others use DM only/gas simulations with/out radiative transfer (e.g Gnedin & Ostriker, 1997; Iliev, Scannapieco, & Shapiro, 2005; Trac & Cen, 2007; Kohler, Gnedin, & Hamilton, 2007; Pawlik, Schaye, & van Scherpenzeel, 2009; Raičević & Theuns, 2010) to compute clumping factors. Early calculations of Gnedin & Ostriker (1997) have high values of gas clumping factors (~ 40) at low redshifts ($z = 5$). But recent works quote lower values (~ 3 at $z = 6$), especially the ones with photoheating due to reionization (Pawlik, Schaye, & van Scherpenzeel, 2009).

Recent studies have shown variations in clumping factor values due to a large number of factors. Kohler, Gnedin, & Hamilton (2007) calculate the clumping factor of gas from 8 equal subregions (2 h^{-1}Mpc) in their simulation box of dimension 4 h^{-1}Mpc, and estimate correlations between clumping factor, gas density and neutral fraction. Pawlik, Schaye, & van Scherpenzeel (2009) show that photoheating reduces the clumping factor of gas, thus reducing the number of ionizing photons needed to keep the universe ionized. Raičević & Theuns (2010) show that the mean of the clumping factors calculated for the sub-regions

Table 4.1: The hydrodynamical simulations used in this work. The columns list, from left to right, the simulation identifier, the comoving box size L, the total number of particles (DM and gas), the DM particle mass $m_{\rm DM}$, the gas particle mass $m_{\rm gas}$, comoving softening length η and radiative transfer (RT) grid size $\mathrm{N_c^3}$.

Model	L $[h^{-1}\,\mathrm{Mpc}]$	Particles	$m_{\rm DM}$ $[h^{-1}M_\odot]$	$m_{\rm gas}$ $[h^{-1}M_\odot]$	η $[h^{-1}\,\mathrm{kpc}]$	$\mathrm{N_c^3}$
$35.12G128$	35.12	2×512^3	8.26×10^7	4.15×10^6	2.28	128^3
$8.78G128$	8.78	2×256^3	1.03×10^7	5.19×10^5	1.14	128^3
$4.39G128$	4.39	2×256^3	1.29×10^6	6.48×10^4	0.57	128^3
$2.20G128$	2.20	2×256^3	1.61×10^5	8.10×10^3	0.29	128^3
$2.20G64$	2.20	2×256^3	1.61×10^5	8.10×10^3	0.29	64^3
$2.20G32$	2.20	2×256^3	1.61×10^5	8.10×10^3	0.29	32^3
$2.20G128v2$	2.20	2×256^3	1.61×10^5	8.10×10^3	0.29	128^3
$2.20G128v3$	2.20	2×256^3	1.61×10^5	8.10×10^3	0.29	128^3

(local clumping factor) is lower than the clumping factor calculated using the simulation volume as a whole (global clumping factor). They also show that the value of the local clumping factors depend on the volume and the overdensity of the sub-regions studied. Sub-regions with a median overdensity values (0.3-3) have the highest local clumping factors compared to high/low overdensity sub-regions. Also, using the mean of the local clumping factors instead of the global clumping factor for the simulation volume in low resolution simulations (e.g. with cell sizes as big as 2 h^{-1}Mpc) brings the number of recombinations in the region closer to that of the true value from a high resolution simulation.

Previous works have mainly focused on computing the H gas clumping factor as most simulations only include H ionization (except for e.g. Trac & Cen, 2007; Pawlik & Schaye, 2011). Recent work by Ciardi et al. (2012) has shown that both H and He RT are essential to get the temperature and ionization structure accurately. Therefore, future large simulations need both H and He RT and would need clumping factors for the ionized species of both H and He. Thus, with our new suite of high resolution reionization simulations, we calculate the clumping factor of gas, HI, HII, HeI , HeII and HeIII and their redshift evolution. We also look in detail into the various factors affecting our clumping factor calculations such as resolution effects, comoving volume gas density and distribution. In Section 4.1 and 4.2, we describe the simulations used to study clumping factors. In Section 4.3 we discuss the definition of the clumping factor and the analysis performed on the different boxes. The conclusions are discussed in Section 4.4.

4.1 Simulations of Reionization

The simulations are based on those described in Ciardi et al. (2012). Here we briefly describe the main characteristics and refer the reader to the above paper. The reionization simulations are produced by post-processing hydrodynamic simulations with the 3D radiative transfer grid based Monte-Carlo code `CRASH` (Ciardi et al., 2001; Maselli, Ferrara, &

Ciardi, 2003; Maselli, Ciardi, & Kanekar, 2009; Partl et al., 2011; Pierleoni et al., 2012). The distribution and evolution of relevant physical quantities as gas number density, temperature and halo masses are provided by hydrodynamical simulations performed with the tree-smoothed particle hydrodynamics (Tree-SPH) code GADGET-3, which is an updated version of the publicly available code GADGET-2 (Springel, 2005). Star formation prescription converts all gas particles with overdensity $> 10^3$ and temperature $< 10^5$ K into collisionless stars. The simulations also include photoionization and heating of an assumed optically thin IGM due to instantaneous reionization at $z = 9$. This feedback is important to properly model the gas distribution towards the end of the simulation where pressure smoothing due to photoheating from EOR would have lead to lower gas clumping than in the case without radiation feedback (Pawlik, Schaye, & van Scherpenzeel, 2009).

In this work, we use hydrodynamical simulations performed in boxes of comoving size 35.12, 8.78, 4.39 and 2.20 h^{-1}Mpc. The parameters of the simulations are summarized in Table 4.1. The cosmology used is as follows : $\Omega_{0,m}$=0.26, $\Omega_{0,\Lambda} = 0.74$, $\Omega_{0,b} = 0.024\ h^2$, h=0.72, n_s=0.95 and σ_8=0.85. Haloes are recognized at each redshift using a friends-of-friends algorithm with a linking length of 0.2. Each snapshot provides gas densities, temperatures and halo masses which are then transferred to a uniform grid as input for `CRASH`.

The radiative transfer calculations are performed on these input fields as a post-process using `CRASH`, which follows self-consistently the evolution of the hydrogen (92% in number) and helium (8% in number) ionization state and the gas temperature. All the simulation boxes are gridded to a 128^3 grid. This leads to lower grid resolution for larger boxes. To study the effect of varying grid resolution for the same box size, we simulate the 2.20 h^{-1} Mpc box with lower grid resolutions of 64^3 (2.20G64) and 32^3 (2.20G32) to be compared to 2.20G128. Also to understand the effect of varying box size for the same grid resolution, we use 4.39G128 (same spatial resolution as 2.20G64) and 8.78G128 (same resolution as 2.20G32). To get the same resolution as 35.12G128, we would need to simulate 2.20 h^{-1} Mpc box with grid resolutions of 8^3 which is really low grid size for RT simulations. Simulations 2.20G128v2 and 2.20G128v3 have the same mean gas number density as 2.20G128 but with different realizations of initial conditions leading to different source and gas distributions, the effect of which on gas clumping are studied in Section 4.3.5.

Simulations start from snapshots at $z = 15$ till 2.6 with each RT run lasting for the Hubble time between snapshots. The emission properties of the sources are derived assuming that the total comoving hydrogen ionising emissivity at each redshift, $\epsilon_{\rm HI}$, is given by (Equation 3 of Ciardi et al. (2012)):

$$\epsilon_{\rm HI} = \begin{cases} \varepsilon \times 10^{50.89+\log(\chi(z))\frac{\alpha^{-1}(\alpha+3)}{2}} & z > 6, \\ \varepsilon \times 10^{50.50-0.06(z-6)\frac{\alpha^{-1}(\alpha+3)}{2}} & z \leq 6, \end{cases} \tag{4.2}$$

with

$$\chi(z) = \frac{\beta e^{\gamma(z-9)}}{(\beta - \gamma + \gamma e^{\beta(z-9)})}, \tag{4.3}$$

where $\beta = 14/15$ and $\gamma = 2/3$. The extreme-UV index of the source spectrum is set to a

power-law index $\alpha = 1.8$ which is typical of quasars with a soft spectra (Telfer et al., 2002). The amplitude of the emissivity $\varepsilon = 1.2$ which is obtained by solving Equation 1 of Ciardi et al. (2012). Equation 4.2 is consistent with constraints on the HI photoionization rate from Lyα forest at $z \leq 6$ (Bolton & Haehnelt, 2007) and the mean free path for Lyman limit photons (Songaila & Cowie, 2010). Equation 4.3 from Springel & Hernquist (2003) provides a parameterization for the rising emissivity at $z > 6$ required by the Lyα forest data at $z < 6$ (e.g. Pritchard, Loeb, & Wyithe, 2010). The emissivity is then distributed among the sources according to their gas mass. This leads to the total emissivity being the same in all the boxes in Table 4.1. This method of assigning the ionization rate avoids assuming an escape fraction of ionising photons and an efficiency of star formation, which are very uncertain parameters. This approach is designed to be consistent with the existing observational constraints on the UV background at $z < 6$.

4.2 Ionization Fractions

First we study the reionization history of the simulation boxes and their dependence on the source properties in the region and the gridding resolution used for RT. The mass range of sources (DM haloes) within a simulation box (selected to be at the mean density of the Universe) is determined by the size of the comoving volume (box size) and the DM halo mass resolution in the simulation. Larger comoving volumes contain rare, massive sources which ionize a larger region thus affecting the reionization history of a larger comoving volume. The lower limit of the mass range is determined by the DM particle mass in the simulation and the number of particles needed to resolve a DM halo (≥ 20 particles/DM halo). Due to the higher particle mass resolution in our smaller boxes, the range of resolved dark matter haloes extend towards lower masses than in the larger box. These low mass sources have a more uniform distribution (less clustering) than the high mass sources in a comoving volume thus affecting the topology of reionization. Therefore, the source properties (DM mass range and clustering) are different for different box size simulations.

The gridding resolution of the simulation determines the range of cell densities and the clumpiness of the gas in the gridded density field. This is important to determine the detailed ionization balance in the volume and sets the pace of the reionization history. Higher the gridding resolution, higher are the densities that are resolved which lead to a better calculation of the ionization balance in the simulation volume. Note that the gridding resolution also determines the total mass of ionizing source present in each cell. Higher resolution gridding has a more detailed distribution of sources in a region than a lower gridding resolution thus giving a more intricate reionization topology.

To visualize this effect, we show a slice through the H ionization structure of the largest $35.12G128$ and the smallest $2.20G128$ (equivalent to 8^3 cells of $35.12G128$) of our simulation boxes in Figure 4.1. To understand the differences in the evolution of the ionization structure in these two boxes, we plot slices at $z = 13.5, 12, 9.5$ and 6.6. The simulation $35.12G128$ with DM particle mass of $m_{\rm DM} = 8.26 \times 10^7\ h^{-1}M_\odot$ has haloes at $z = 13.5$ in the DM mass range $\sim 10^{9-12}\ h^{-1}M_\odot$ which produce HII regions of a few hundred kpcs

(proper) in diameter. The regions grow in size with redshift and eventually merge thus fully ionizing the computational volume by $z = 6.6$. Where as in the $2.20G128$ box at $z = 13.5$, the haloes are in the DM mass range $\sim 10^{7-9}\ h^{-1}M_{\odot}$, with typical HII regions of tens of kpcs. The reionization history is then determined by the growth and merging of a large number of smaller ionized regions compared to the case of $35.12G128$. Also, the distribution of bubbles in $2.20G128$ is more uniform than the $35.12G128$ due to the difference in clustering properties of sources in both the simulations. Keep in mind that the total emissivity is same in both the boxes.

Other than the source properties, another important factor which affects the reionization history is the gas number density of the cells. Since both simulations use a 128^3 grid for the RT calculation, the density field in $35.12G128$ is smoothed on a scale 8 times larger than in $2.20G128$. For example, at $z = 14.5$ the gas number densities in $35.12G128$ are in the range $10^{-3.7} - 10^{-2.1}$ cm^{-3} compared to wider range $10^{-4} - 10^{-1.5}$ cm^{-3} found in $2.20G128$. The range of gas number densities plays an important role in determining the ionization history of the comoving volume. Higher density cells need more photons to get ionized, which slows the reionization process. This also leads to some high density cells retaining a fraction of the neutral gas in the $2.20G128$ box towards lower redshifts (i.e. $z = 6.6$) in contrast to $35.12G128$ where the volume is fully ionized for all cell densities. The high density cells also lead to the formation of non-spherical ionized regions.

Therefore, we can see that both source properties and cell densities play an important role in determining the reionization history of the computational volume. The source properties has a large impact on the topology and the cell densities control the pace of reionization in the simulation.

4.2.1 Gas Ionization in Different Boxes

Next we look at the quantitative differences between the various simulations. Figure 4.2 shows the normalized probability density function (PDF) for density, temperature and ionization fields at $z = 14.5$ and 3.8. Comparing $2.20G128$, $2.20G64$ and $2.20G32$ gives us information about the effect of varying grid resolution for the same box size while comparing $4.39G128$ to $2.20G64$ and $8.78G128$ to $2.20G32$ bring out the effects of varying box size for the same grid resolution. The simulation $35.12G128$ is included in the plots for comparison.

First we focus on the density distribution in the different simulations. At $z = 14.5$ for $35.12G128$, the gas number density in cells range from $10^{-3.8}$ to 10^{-2} cm^{-3}. As grid resolution improves, the range in cell densities increases to 10^{-4} - $10^{-1.5}$ cm^{-3} for $2.20G128$. The $4.39G128$ and $8.78G128$ distributions lie in between these two extremes. The density distribution in $2.20G64$ is very similar to $4.39G128$, because of the same resolution. Mild differences are observed at very high and low densities. The same is true for $2.20G32$ and $8.78G128$. For simulations of different grid resolutions, a higher grid resolution resolves a wider range of cell densities. But, at the same grid resolution, the discrepancy at the high densities between simulations of different box sizes (e.g. $4.39G128$ and $2.20G64$) is caused by the presence of higher mass objects in larger boxes. At low densities, the cell densities

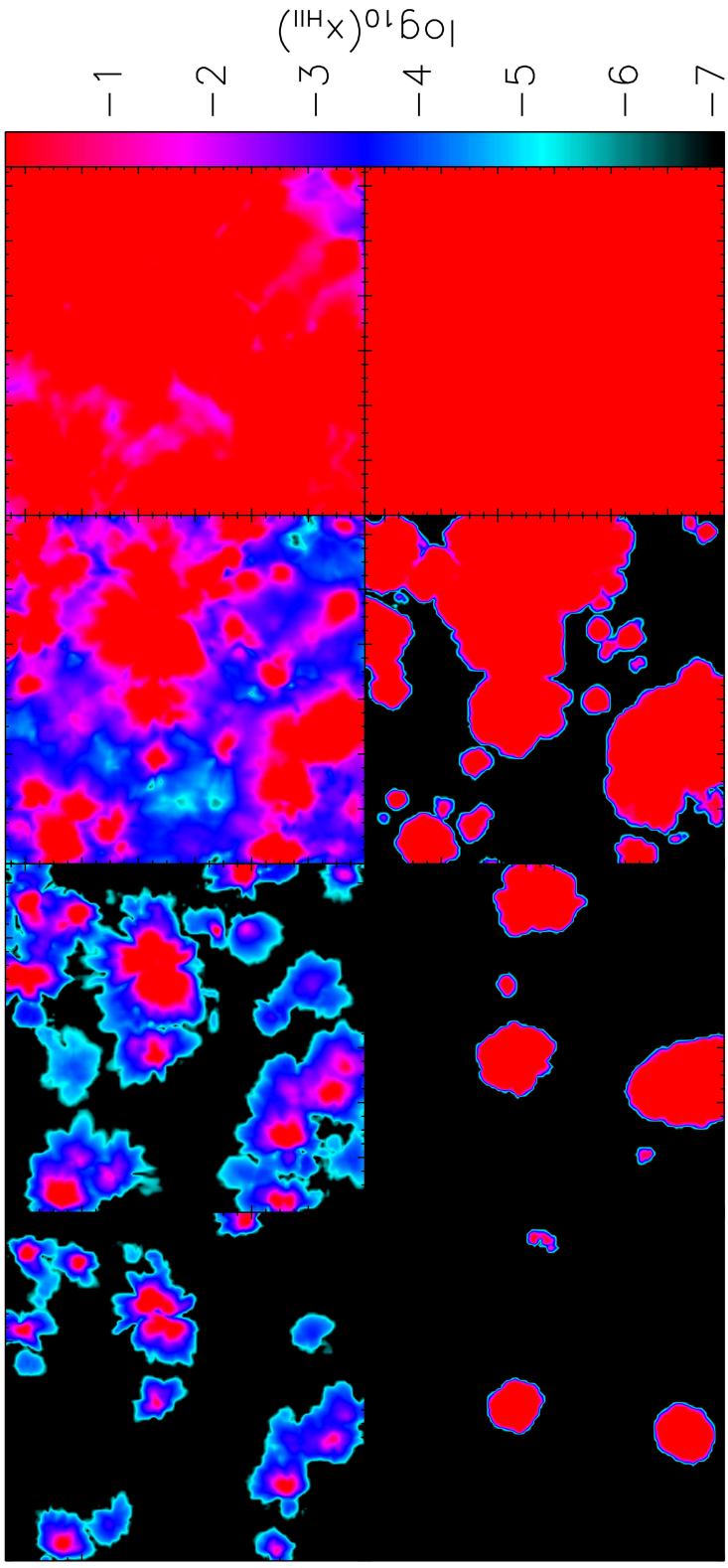

Figure 4.1: A slice cut through the simulation box showing the ionization structure for $2.20G128$ (top row) and $35.12G128$ (bottom row) at redshift (from left to right) $z = 13.5$, 12, 9.5 and 6.6. The color bar shows the fraction of H as HII in each cell, plotted in log scale.

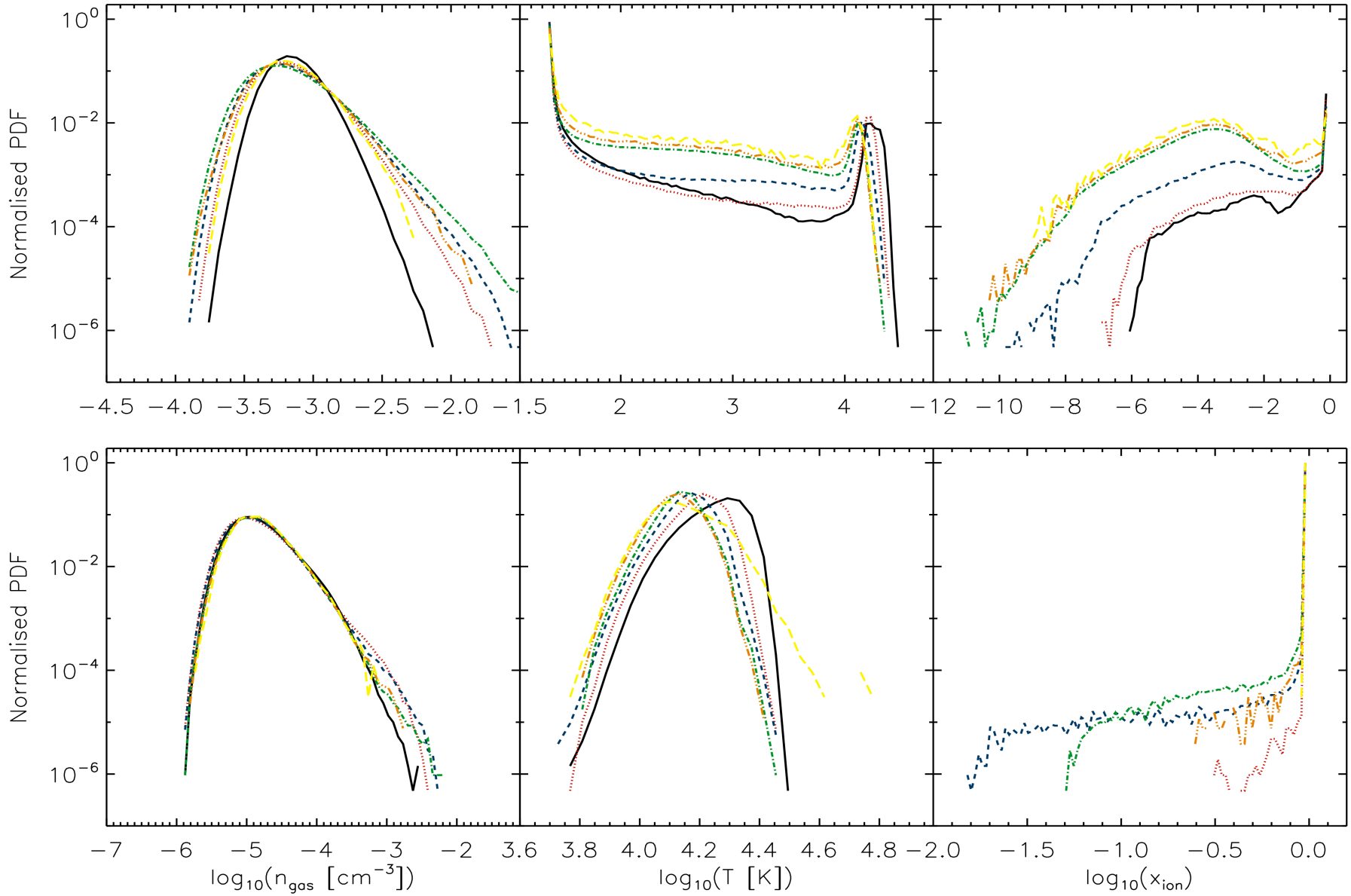

Figure 4.2: Normalized PDF of gas number density (left column), temperature (middle column) and hydrogen ionization fraction (right column) in simulations at redshifts z =14.5 (top row) and z =3.8 (bottom row). The lines are for - 35.12G128 (black solid line), 8.78G128 (red dotted line), 4.39G128 (blue dashed line), 2.20G128 (green dot dashed line), 2.20G64 (orange triple dot dashed line) and 2.20G32 (yellow long dashed line).

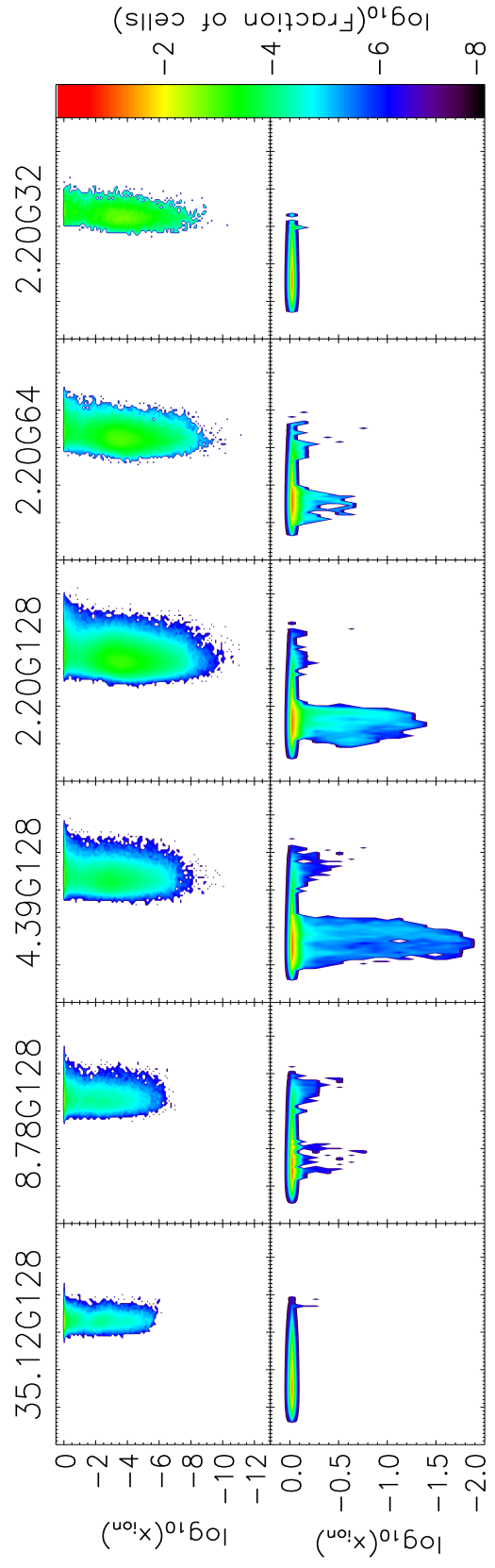

Figure 4.3: Gas ionization fraction plotted against gas number density with colors according to fraction of cells (in that 2D bin) at redshifts z=14.5 (top row) and z=3.8 (bottom row). The columns refer to, from left to right - 35.12G128, 8.78G128, 4.39G128, 2.20G128, 2.20G64 and 2.20G32.

depend on the detailed gas distribution in the voids. At high redshifts (e.g. $z = 14.5$), there is a lot of gas in the voids as the mean density of the Universe is higher compared to low redshifts. The variations in the gas are at approximately the scales of the cell size in $2.20G128$. Therefore, when we use lower grid resolution, the gas density field is smoothed to a smaller range in cell density values. At $z = 3.8$, the density distribution becomes similar for all boxes, with the exception of the higher densities where it is controlled by the virialised haloes. At low densities, the distribution is similar for all boxes as most of the gas at these redshifts is concentrated in high density filaments and virialised objects, which are at scales much smaller than the cell sizes even in $2.20G128$. Thus we can summarize saying that the cell density distribution depends on mainly grid resolution especially at high redshifts. At all redshifts, the high density cells are determined by the gas mass in the ionizing sources (i.e. galaxies) within those cells.

Next we study how the temperature distribution varies. At $z = 14.5$, most of the gas (i.e. 50-80%) is at 25 K because it still hasn't been reached by the ionizing radiation. Another peak is observed at $\sim 10^4$ K, corresponding to the highly ionized cells. At intermediate temperatures, the distribution has a flat profile with a higher fraction of the cells at intermediate temperatures in small boxes compared to large ones. This is due to the more uniform ionization topology of the small boxes (refer to Figure 4.1). The larger the box size, the higher is the maximum temperature achieved in a cell because of the higher ionization rates coming from the larger sources. As a confirmation, the behaviour of $2.20G64$ and $2.20G32$ follows that of $2.20G128$, rather than of $4.39G128$ and $8.78G128$. On the other hand, the smaller the box size is, the larger is the fraction of the cells with a temperature > 25 K. This is due to the larger range of ionization fractions present in the smaller boxes (see below). At $z = 3.8$, the temperature distributions within the various boxes are similar. The temperature distribution has a near Gaussian profile with the peak shifting towards higher temperatures for larger boxes (see earlier discussion). $2.20G32$ shows an unusual behaviour, most probably due to the poor gridding resolution, indicating that at least 64^3 cells are needed to properly resolve the temperature structure in the simulations.

Finally we look at the distribution of the ionization fraction for the gas in the cells. The ionization fraction of the gas $x_{\rm ion}$ is defined as the fraction of atoms present in an ionized state in that cell:

$$x_{\rm ion} = x_{\rm HII} \times f_{\rm H} + (x_{\rm HeII} + x_{\rm HeIII}) \times f_{\rm He}, \tag{4.4}$$

where $x_{\rm HII}$, $x_{\rm HeII}$ and $x_{\rm HeIII}$ are the volume averaged fractions of ionized H, single-ionized and double-ionized He respectively. $f_{\rm H} = 0.92$ is the fraction of H atoms and $f_{\rm He} = 0.08$ is the fraction of He atoms in the gas by number as mentioned earlier. At $z = 14.5$, about 15-35 % of the cells are highly ionized with the rest spread over lower ionization fractions. The larger the box size, the higher the number of cells that are highly ionized. The ionization structure in a box is mainly determined by the sources distribution as $2.20G64$ and $2.20G32$ follow that of $2.20G128$. At low redshift $z = 3.8$, almost all the cells are highly ionized. Also the ionization fraction depends on the density distribution as $2.20G64$ follows $4.39G128$ and

$2.20G32$ follows $8.78G128$. Thus one can say that source properties are more important at high redshifts in determining the temperature and ionization distribution, while at lower redshifts, density distribution is important to determine the ionization structure.

We plot in Figure 4.3 the ionization fraction of the cells against their gas densities to understand the correlation between these two quantities in the different boxes. Plots are shown for $z = 14.5$ and 3.8 as in the previous figure. At $z = 14.5$, all simulation boxes have a narrow range in gas densities but a large spread in ionization fractions. For $35.12G128$ only a few percent of the cells have an ionization fraction ~ 1. These are concentrated at number densities $\sim 10^{-3}$ cm^{-3}. The rest of the cells have lower ionization fractions, spanning a few orders of magnitude in ionization for the same gas number density. As we go to smaller boxes, as already observed, the range of number densities widens and lower ionization fractions are present. Similarly to the larger boxes though, the ionization fraction of cells with same number density can differ by orders of magnitude. Since the ionization fraction distribution in $2.20G64$ and $2.20G32$ is more similar to that of $2.20G128$ instead of $4.39G128$ and $8.78G128$, we can say again that at high redshifts the ionization structure is determined by source distribution more than cell densities.

At $z = 3.8$ in all the simulation boxes, there is hardly any neutral gas left. Almost all the hydrogen gas has been ionized and in case of He, almost all the atoms are either single/double ionized (refer to the next subsection for more details). Note that since the gas is 92% H, $x_{\rm ion}$ is dominated by $x_{\rm HII}$. In high spatial resolution simulations, a small fraction of cells still contain some neutral atoms. These are concentrated in the low density regions which haven't been reached by ionizing photons yet. Also, because $2.20G64$ and $2.20G32$ exhibit a distribution similar to that of $4.39G128$ and $8.78G128$ respectively, we conclude that the gas density is the most important factor in the determination of the ionization structure. Thus again we can say that at high redshifts, source properties play a larger role in determining ionization structure while at lower redshifts, gas distribution plays a larger role.

4.2.2 Different Species

Since we have both H and He in our RT simulations, we can look at how the ionization history differs for different species, i.e. HII, HeII and HeIII. Discrepancies in their evolution are expected due to the different ionization energy of the species. More specifically, the ionization energy of HI, HeI and HeII is 13.6 eV, 24.6 eV and 54.4 eV, respectively (Osterbrock, 1989). In Figure 4.4 slices through the mid plane of the ionization structure of HII, HeII and HeIII is shown for redshifts $z = 14.5,\ 10.5,\ 6.6$ and 3.8 for $2.20G128$. Due to paucity of higher energy photons in the spectra, HeIII is typically found only in the vicinity of the sources where helium is fully ionized. At larger distances instead HeII is more abundant. The ionization profiles of HeII are very similar to those of HII albeit smoother. This is consistent with previous studies investigating the ionization of both H and He. All three ionization species have similar evolution, with the ionized bubbles growing in size and number, and merging together as redshift decreases. At $z = 3.8$, while H is completely ionized with the exception of a few more neutral pockets, He still exhibits

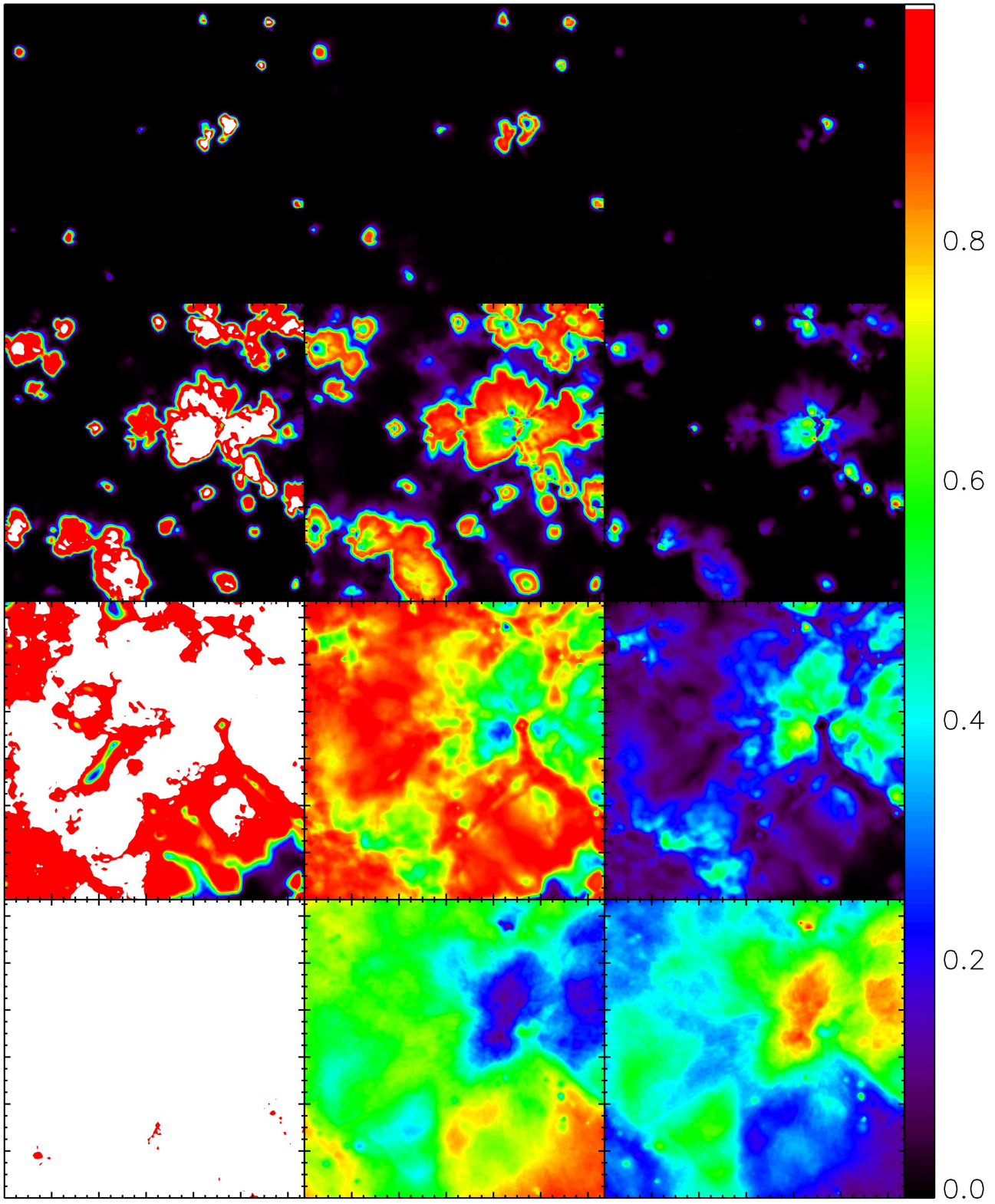

Figure 4.4: A slice through the mid-plane of the 2.20G128 simulation, showing $x_{\rm HII}$ (left column), $x_{\rm HeII}$ (middle column) and $x_{\rm HeIII}$ (right column). Plots are for redshifts (from top to bottom) $z = 14.5,\ 10.5,\ 6.6$ and 3.8.

both of its ionization states, with the high density regions dominated by HeIII and the low density regions by HeII.

On a more quantitative note, we can calculate the redshift evolution of the volume averaged ionization fraction of the different species, i.e. $x_{\rm HII/HeII/HeIII}$. These are shown in Figure 4.5. In this plot, we only plot upto $z = 3$ as we do not have outputs below that redshift for $35.12G128$ even though all the other simulations extend to $z = 2.6$. The evolution of $x_{\rm HII}$ is very similar for all the simulations and increases smoothly until full hydrogen reionization is reached. The larger the box size is, faster is the H reionization in the volume. Comparing $2.20G128$ to $2.20G64$ and $2.20G32$, we can see that as grid resolution decreases, the ionization of the volume goes faster. This is because in high resolution simulations, the high density regions are resolved which require more photons to be fully ionized. At the same grid resolution but different box sizes instead, i.e. $2.20G64 - 4.39G128$ and $2.20G32 - 8.78G128$ pairs, the evolution of $x_{\rm HII}$ is slower in the larger boxes. This is because of the few high density cells present in the larger boxes (as shown in the previous subsection) which need more photons to reionize.

From the three panels we can see that the reionization history of the three species show very different behaviour. First we focus on the curves for $2.20G128$ in all three panels. $x_{\rm HII}$ increases steadily to $z = 8$ and then slowly flattens out with the box reaching almost full ionization. $x_{\rm HeII}$ increases steadily, flattens between $8 < z < 6$ and then starts decreasing. $x_{\rm HeIII}$ increases steadily, flattens between $8 < z < 6$ and then starts increasing again as HeII is converted to HeIII. The shape (the flattening between $8 < z < 6$ and then rising) of the ionization history in the case of HeII/HeIII is determined by the shape of the emissivity (see Equation 4.2).

Comparing these curves to the ones from other boxes show that similar to $x_{\rm HII}$, $x_{\rm HeIII}$ increases faster for larger boxes at the same grid size (different gridding resolution). Focusing on the evolution of HeIII first, we can see that for the same box size at different gridding resolution, the $x_{\rm HeIII}$ evolution proceeds faster for low grid resolution. For different box sizes at the same gridding resolution, the behaviour is redshift dependent. At $6 \lesssim z \lesssim 10$, smaller boxes have higher $x_{\rm HeIII}$ compared to the large ones. But at high ($z \gtrsim 10$) and low ($z \lesssim 6$) redshifts, larger boxes have higher $x_{\rm HeIII}$. The exact reason for this trend is unclear. But it shows that both source properties (mass range and distribution) and grid resolution is important for HeIII.

Also for $x_{\rm HeII}$ evolution, both source properties and grid resolution seem to be important. $x_{\rm HeII}$ evolution shows strong redshift trends for varying box size and grid resolution. At $z \geq 9$, going to larger boxes at the same grid size (different gridding resolution), the curves stay close to each other. But for the same box size by different gridding resolution case, $x_{\rm HeII}$ is higher for low resolution. And for the same gridding resolution but different box sizes ($2.20G64 - 4.39G128$ and $2.20G32 - 8.78G128$ pairs), $x_{\rm HeII}$ is higher for smaller box sizes. Therefore, it seems that box size and gridding resolution changes counteract each other. At $z < 9$, the trends are different. Larger box size simulations (with the same grid size but different gridding resolution) have lower $x_{\rm HeII}$. This is because both lower grid resolution and larger box sizes lower the $x_{\rm HeII}$ values as HeII is converted to HeIII faster. Lowering grid resolution of $2.20G128$ to $2.20G64$ and $2.20G32$ lowers $x_{\rm HeII}$. At the same

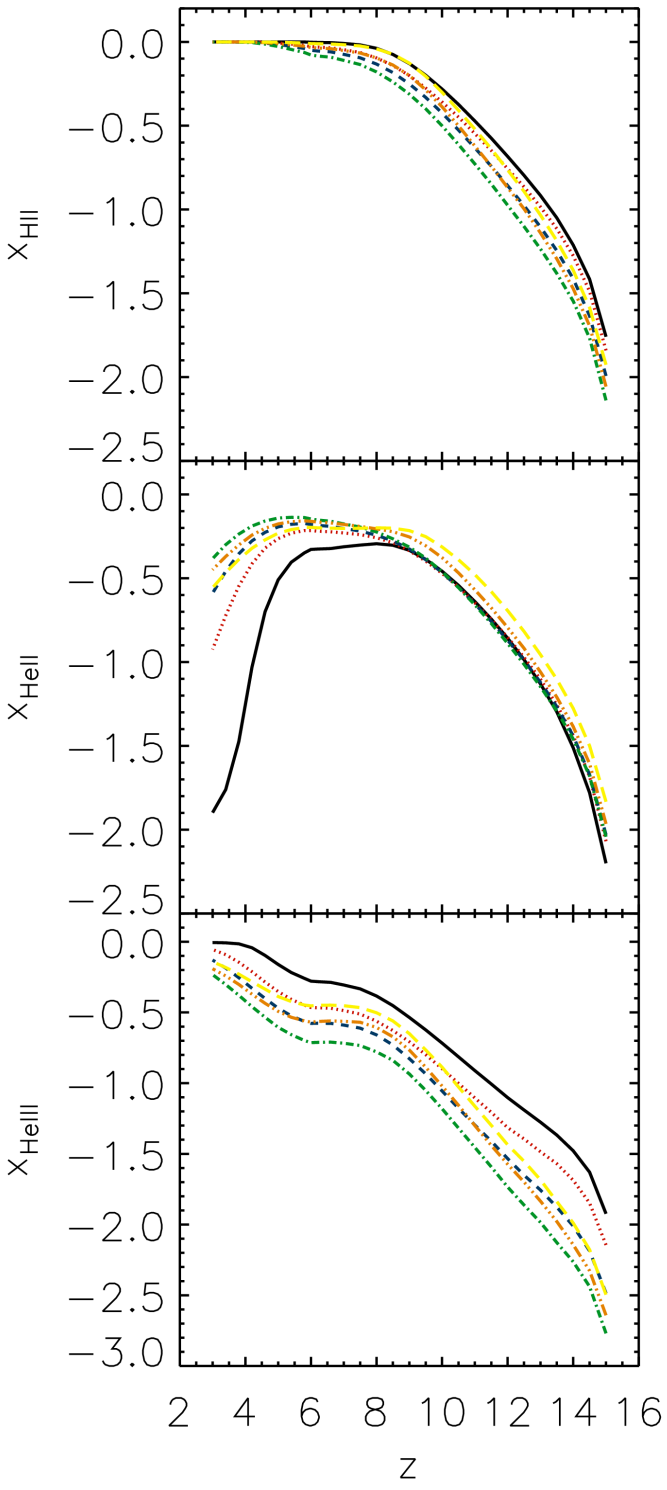

Figure 4.5: Redshift evolution of the volume averaged ionization fractions - $x_{\rm HII}$ (upper panel), $x_{\rm HeII}$ (middle) and $x_{\rm HeIII}$ (lower) for the $35.12G128$ (black solid line), $8.78G128$ (red dotted), $4.39G128$ (blue dashed), $2.20G128$ (green dot-dashed), $2.20G64$ (orange triple dot-dashed) and $2.20G32$ (yellow long-dashed) runs.

grid resolution for $2.20G64 - 4.39G128$ and $2.20G32 - 8.78G128$ pairs, $x_{\rm HeII}$ is lower for larger boxes as stronger sources convert more of HeII to HeIII.

Therefore, we can conclude that specially for the reionization history of He, both grid resolution and box sizes are important. But simulating large volumes at the grid resolution of $2.20G128$ is computationally expensive. To overcome this problem while simulating boxes larger than $35.12G128$ simulations, we need to use sub-grid clumping. For this we can use our suite of simulations to understand the important factors determining the estimation of clumping factors. This is done in the next section.

4.3 Clumping Factor

In RT simulations, the sub-grid clumping is included using the parameter 'clumping factor'. There are different ways to parameterize the clumping factor (e.g. Trac & Cen, 2007; Pawlik, Schaye, & van Scherpenzeel, 2009). In this work, we define the clumping factor as:

$$C_i = \langle n_i^2 \rangle / \langle n_i \rangle^2 \tag{4.5}$$

where n_i is the number density of the species i = HI, HII, HeI, HeII, HeIII and total gas. While typically the definition of clumping factor refers to the total gas (e.g. Pawlik, Schaye, & van Scherpenzeel, 2009), here we study its behaviour with respect to all the species using the $2.20G128$ simulation, i.e. the one with the best grid resolution. This definition of clumping factor estimates the spread in the number density values for each species in grid cells with respect to a mean value calculated for the whole simulation volume. This helps us understand the uniformity in the distribution of each species in the simulation box as the process of reionization proceeds. $C_{\rm HII}$ is equivalent to that of $C_{\rm R,HII}$ from Equation 4.1 under the assumption that the recombination rate is constant for the cell conditions involved in our study ($\alpha_{\rm R,HII}$ = constant) and that H atoms are the only source of free electrons leading to $n_{\rm HII} = n_{\rm e}$. Both these assumptions are approximately valid in our case with the deviations not significantly affecting our results.

4.3.1 Clumping Factor of Different Species

We start with studying how different species clump in our high resolution simulation. We calculate the clumping factor for total gas, HI, HII, HeI, HeII and HeIII. Since we are only interested in the clumping factor of the IGM, we need to remove the cells containing collapsed haloes. We define gas overdensity as:

$$\Delta = n_{\rm gas} / \langle n_{\rm gas} \rangle \tag{4.6}$$

where $n_{\rm gas}$ is the gas number density in a cell and $\langle n_{\rm gas} \rangle$ is the mean gas number density of the universe at that redshift. We assume that the IGM is composed only of cells with $\Delta \leq 100$ (Miralda-Escudé, Haehnelt, & Rees, 2000; Pawlik, Schaye, & van Scherpenzeel, 2009; Raičević & Theuns, 2010). The overdensity of a collapsed DM halo depends on the

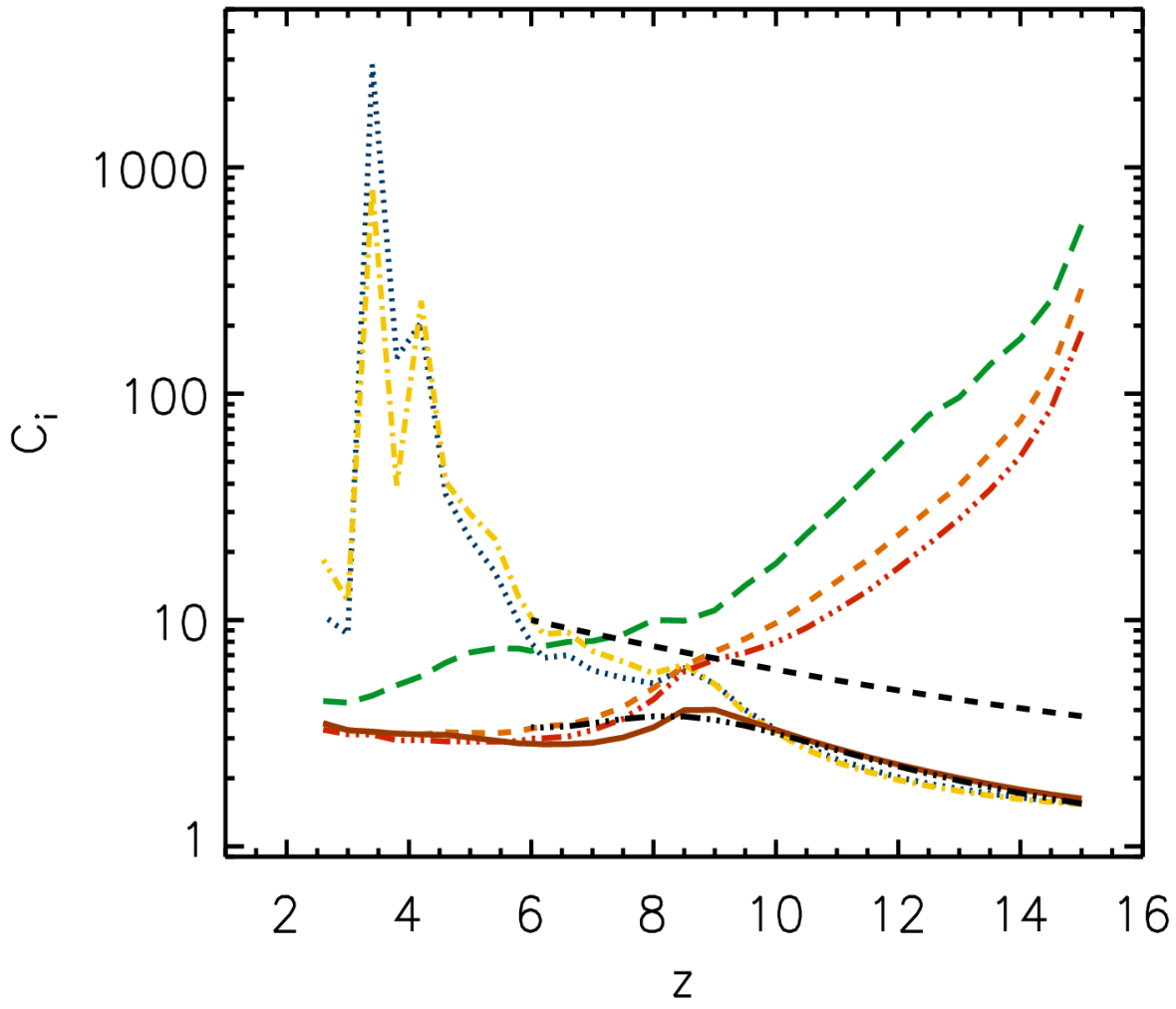

Figure 4.6: Clumping factor C_i of different species in the $2.20G128$ simulation. The lines are - gas (brown solid line), HI(blue dotted line), HII(orange short dashed line), HeI(yellow dot dashed line), HeII(red triple dot dashed line) and HeIII(green long dashed line). For comparison, plotted are the gas clumping factor from Iliev et al. (2007) (black dashed line) and Pawlik et al. (2009) (black triple dot dashed line).

definition used to compute its virial radius. For example, for spherical top hat collapse the DM overdensity at the virial radius is ~ 178 (e.g. Padmanabhan, 1993), while for an isothermal collapse it is ~ 60 (e.g. Lacey & Cole, 1994). For comparison and consistency with earlier works, we adopt the generally used overdensity threshold of $\Delta = 100$.

Figure 4.6 shows the clumping factor of different species plotted against z. First we look at the clumping factor of total gas. The clumping factor of gas increases with decreasing redshift from 1.5 at $z = 15$ to about 3 at $z \sim 8.5$. This trend is due to the self-gravity of the gas in the IGM. At $z = 9$ the simulation includes instantaneous photoionization and reheating of the IGM by a spatially uniform ionizing background (Haardt & Madau, 2001) assuming an optically thin IGM. This feedback leads to pressure smoothing of the gas in the IGM and a reduction of the clumping factor (e.g. Pawlik, Schaye, & van Scherpenzeel, 2009). Therefore, the gas clumping factor at $z < 8.5$ decreases to about 2 at $z \sim 6$. At $z \leq 6$ self-gravity becomes dominant again and the gas clumping factor starts increasing again, as the Universe evolves, to a value of 3 at $z = 2.5$.

For comparison, we plot the clumping factors calculated by Iliev et al. (2007) using high resolution dark matter only simulations of a 3.5 h^{-1} Mpc box with 1624^3 particles for WMAP3 cosmology. Assuming that the gas follows the dark matter distribution at high redshifts, the following fit was obtained from their simulations for the clumping factor of gas in the range $6 < z < 30$,

$$C_{\rm Iliev07} = 26.2917\ e^{-0.1822z+0.003505z^2}. \tag{4.7}$$

As we can see in Figure 4.6, the clumping factor derived from our hydrodynamic simulations is a factor of few lower than the one obtained by Iliev et al. (2007). A direct comparison of the two curves is not possible due to the differences in the details of the two simulations. In particular, photoheating of the gas is not accounted for in DM only simulations, nor was an overdensity threshold to remove virialised halo gas included, as was noted in Pawlik, Schaye, & van Scherpenzeel (2009). On the other hand, the evolution of the clumping factor in the r9L6N256 simulation of Pawlik, Schaye, & van Scherpenzeel (2009) (refer to the original paper for more details) is very similar to ours, with a decline of the clumping factor at $z \sim 8$, most probably because also their simulations include photoheating of the gas from $z = 9$. The small differences observed are due to differences in specific details between the two simulations.

Next we focus on the clumping factor of the two components of H, i.e. the neutral component HI and the ionized one HII. The HI clumping factor follows the gas up to $z \sim 11$. This is because the gas at high redshifts is mostly neutral with only very small pockets of ionized gas. At lower redshifts though $C_{\rm HI}$ keeps increasing as reionization proceeds with only a few pockets of neutral gas left, thus giving high clumping factor values. At $z < 5$ the curve becomes very noisy due to the small number of cells still neutral. A reversed behaviour is observed in the evolution of $C_{\rm HII}$, which is as high as ~ 250 at $z = 15$, when only a handful of small ionized bubbles are present around the sources. But towards lower redshifts, most of the gas becomes ionized leading to low clumping factors (< 10 at $z < 10$) and eventually being the same as total gas clumping factor at $z < 5$, when $x_{\rm HII} \gtrsim 0.95$.

Finally we look at the components of He, i.e. HeI, HeII and HeIII. The evolution of $C_{\rm HeI}$ is very similar to that of $C_{\rm HI}$, with values close to those of $C_{\rm gas}$ at $z > 11$ and progressively increasing with decreasing redshift. Similarly, $C_{\rm HeII}$ closely follows the evolution of $C_{\rm HII}$, but with slightly lower values due to the larger HeII regions with smoother edges compared to the corresponding HII regions (refer to Figure 4.4). $C_{\rm HeIII}$ exhibits the same qualitative evolution, but it has values higher than those of HeII by a factor of a few. This is due to the smaller dimensions of the HeIII regions. Thus we can say that at high redshifts ionized species are clumpier than gas where as at low z they follow the gas. HeIII has higher clumping factors through out the whole redshift range as it is mainly present in the high density regions close to the sources.

4.3.2 Dependence on Overdensity

After understanding how clumping properties of different species change with redshift, we look into how the calculations depend on the overdensity threshold used. The parameters adopted to run the hydrodynamic or N-body simulation used to calculate the gas and galaxy distribution (e.g. the box dimension, resolution and smoothing length) set the range of overdensities spanned by the simulation, together with the grid dimension adopted for the RT calculation. In addition to this range intrinsic to the simulation, one might want to include an artificial upper limit to exclude high density virialised gas from the calculation of the clumping factor (e.g. Miralda-Escudé, Haehnelt, & Rees 2000; Pawlik, Schaye, & van Scherpenzeel 2009; Raičević & Theuns 2010; but see also Iliev et al. 2007 for a different approach). Here we investigate the dependence of the clumping factor on the gas overdensity.

The range of gas overdensities in $2.20G128$ spans approximately two orders of magnitude and it changes with redshift, being $0.16 < \Delta < 52$ and $0.06 < \Delta < 688$ at $z = 15$ and $z = 2.6$, respectively. Most cells ($\gtrsim 80\%$) at all redshifts though are found at $0.3 < \Delta < 3$. As already mentioned, we set an upper limit of $\Delta = 100$ for our clumping factor calculations, while for convenience we choose a lower limit of $\Delta = 0.1$. We have verified that this choice of a lower overdensity cutoff does not affect our results.

For this analysis, the overdensity range of $0.1 < \Delta < 100$ has been split into six bins of width 0.5 dex. Note that almost all overdensity bins used in our calculation have > 1000 cells at all redshifts, except for the highest $30 < \Delta < 100$ bin which has 30 - 500 cells across the redshift range. The clumping factor for each species in different overdensity bins is defined as:

$$C_i(\mathrm{d}\Delta) = \langle \mathrm{n_i^2}(\mathrm{d}\Delta)\rangle / \langle \mathrm{n_i}(\mathrm{d}\Delta)\rangle^2 \tag{4.8}$$

where the species i = gas, HII, HeII and HeIII. Figure 4.7 shows the above clumping factors for the gas, HII, HeII and HeIII in the overdensity bins dΔ= 0.1-100, 0.1-0.3, 0.3-1, 1-3, 3-10, 10-30 and 30-100. We do not plot $C_{\rm HI}$ and $C_{\rm HeI}$ as at high redshifts they closely follow the gas clumping factor, while at low redshifts they are very noisy and involve a negligible number of cells. The top panel refers to the clumping factor of total gas. For $\mathrm{d}\Delta = 0.1 - 100$, we recover the total clumping factor of the IGM shown in Figure 4.6 as a

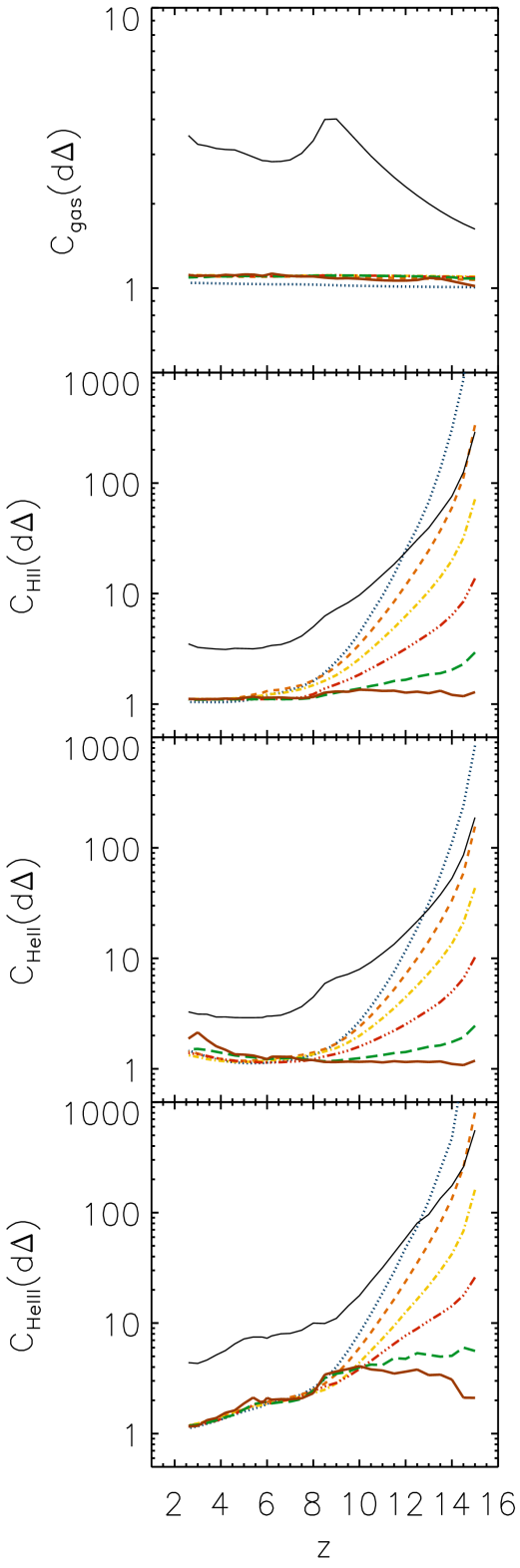

Figure 4.7: Clumping factors $C_i(\mathrm{d}\Delta)$ against redshift for different overdensity bins in the $2.20G128$ simulation. The different species plotted are (from top to bottom) - gas, HII, HeII and HeIII. The density bins are $\mathrm{d}\Delta = 0.1$-100 (black solid lines), 0.1-0.3 (blue dotted), 0.3-1 (orange short dashed), 1-3 (yellow dot dashed), 3-10 (red triple dot dashed), 10-30 (green long dashed) and 30-100 (brown solid).

brown solid line.

The clumping factor in the other overdensity bins is close to 1 at all redshifts, as the overdensity bins are narrow and the gas number density distribution in each of them is dominated by a narrow range of values due to the shape of the density PDF (see left column in Figure 4.2). This can be seen in Figure 4.8 where we plot the normalised PDF of the cell overdensities in each overdensity bin for $2.20G128$ at $z = 3.8$. We can see that the overdensity bins (except for 0.1-100) are very narrow and are parts of the total PDF. The narrow bins have a very biased distribution with one of the edges of the overdensity range dominating the distribution. Also, the mean overdensity value in these overdensity bins are almost equidistant which is important as it can be seen in the next sub-section.

For other species, instead, the clumping factor varies over the bins. Although some differences are evident, the behaviour of the clumping factor for different bins is similar for all the ionized species. The clumping factor is lower for higher overdensity bins, which typically correspond to the high density cells close to the sources and thus highly ionized, with similar values of H/He ionization fraction. For these high density cells, $x_{\rm HII/HeIII} \sim 1$ and $x_{\rm HeII}$ is very low as a consequence (as shown in the previous section). For lower overdensity bins, the clumping factors increase because of the larger range of $x_{\rm HII}$, $x_{\rm HeII}$ and $x_{\rm HeIII}$ encountered (see Figure 4.4). While large differences are present at high redshift, at low redshift the clumping factors in the different overdensity bins converge to values close to 1 due to the advanced ionization of the gas.

The clumping factors for $\mathrm{d}\Delta = 0.1 - 100$ have generally high values compared to the other overdensity bins especially at low redshifts. This is because in each of the individual $\mathrm{d}\Delta$s (other than $\mathrm{d}\Delta = 0.1 - 100$), especially at low redshifts, the range of ionization values (for the cells) is narrow leading to lower clumping factors. Thus we can conclude that both the values and the width of the overdensity bins affect the calculation of clumping factors.

4.3.3 Dependence on Definition of Clumping Factor

In our earlier plots, we looked at the spread in the ionization fractions for different cuts in the cell overdensities. Clumping factor for each species was calculated for within the selected species and selected overdensity cells. But following what done by other authors (e.g. Trac & Cen, 2007) for the clumping factor of HII, here we use an alternative definition to the one in Equation 4.8 and write:

$$C_i'(\mathrm{d}\Delta) = \langle \mathrm{n_i^2(d\Delta)} \rangle / \langle \mathrm{n_{gas}} \rangle^2, \tag{4.9}$$

where $\langle n_{\rm gas} \rangle$ is the volume averaged gas number density in the simulation volume, and substitutes $\langle n_i(\mathrm{d}\Delta) \rangle$ of Equation 4.8. This definition of clumping factor $C_i'(\mathrm{d}\Delta)$ is interesting as it aids in estimating the significance of recombinations $\propto n_i^2$ of each ionized species (HII/HeII/HeIII) in the comoving volume given a mean gas density in the region. At low ionization levels of a species, the clumping factor $C_i(\mathrm{d}\Delta)$ would be high due to poor statistic of a small number of ionized cells but $C_i'(\mathrm{d}\Delta)$ would be low showing that these few but highly ionized cells do not significantly contribution to the recombinations in the region.

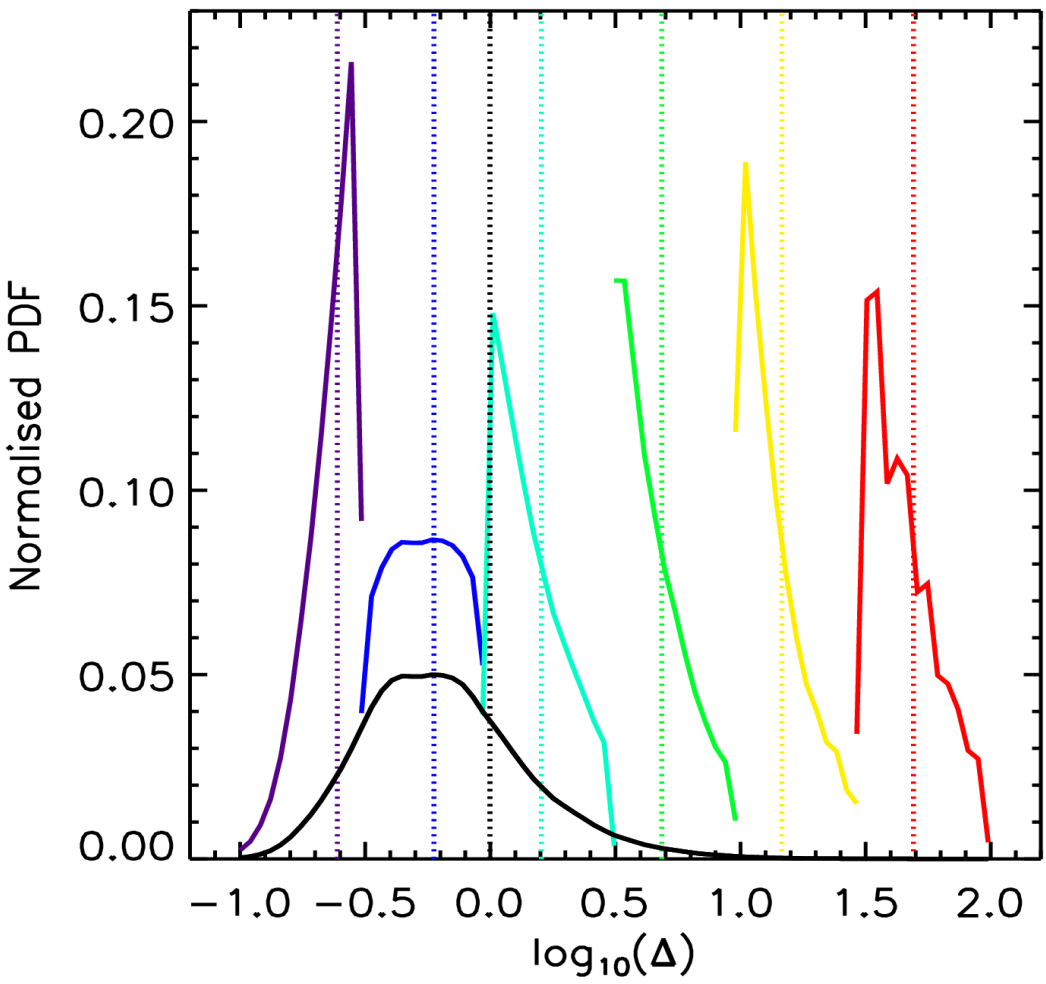

Figure 4.8: Normalised PDF of the cell overdensities Δ for different overdensity bins in the 2.20G128 simulation at $z = 3.8$. The different lines are for the overdensity bins - dΔ = 0.1-100 (black), 0.1-0.3 (violet), 0.3-1 (blue), 1-3 (light blue), 3-10 (green), 10-30 (yellow) and 30-100 (red). The solid lines show the normalised PDF of the overdensities in each bin. The dotted lines show the mean overdensity value in each bin.

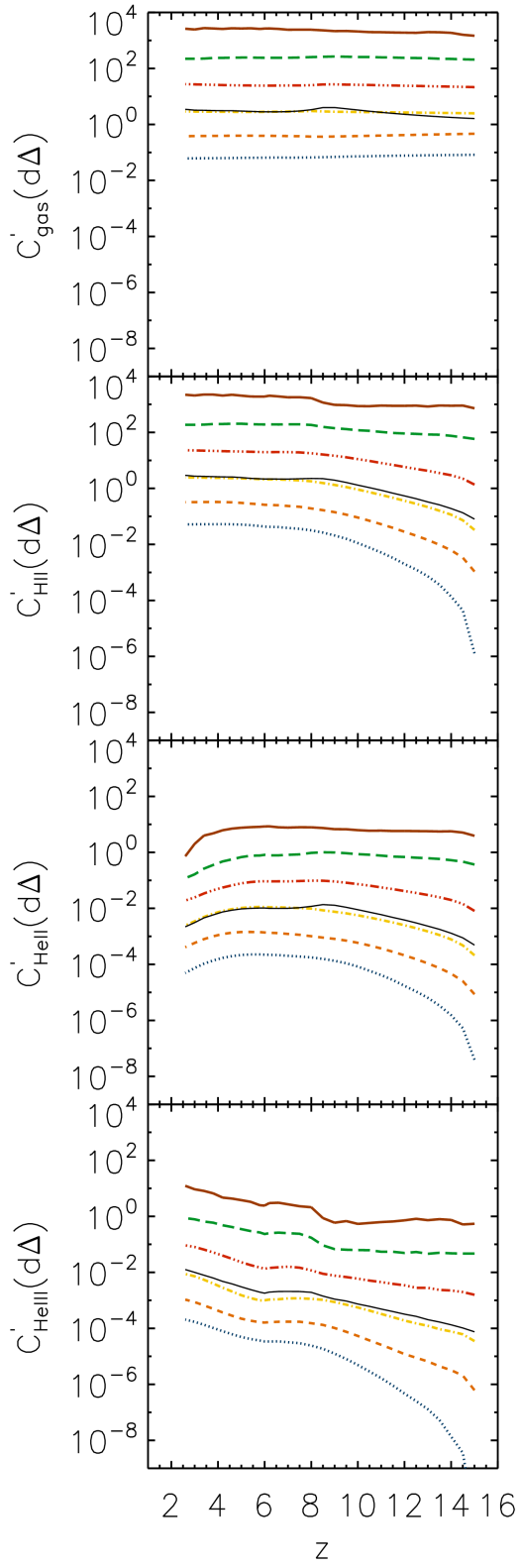

Figure 4.9: Plotted are the clumping factors $C_i'(\mathrm{d}\Delta)$ against redshift for different overdensity bins in the 2.20G128 simulation. The different species plotted are (from top to bottom) - gas, HII, HeII and HeIII. The lines are for $\mathrm{d}\Delta$ = 0.1-100 (black solid lines), 0.1-0.3 (blue dotted), 0.3-1 (orange short dashed), 1-3 (yellow dot dashed), 3-10 (red triple dot dashed), 10-30 (green long dashed) and 30-100 (brown solid).

Figure 4.9 shows $C_i^{'}(\mathrm{d}\Delta)$ for i = gas, HII, HeII and HeIII plotted against z for different $\mathrm{d}\Delta$. First we look at how $C_{\mathrm{gas}}^{'}(\mathrm{d}\Delta)$ evolves with redshift for different overdensity bins which is an interesting case to understand before moving to the ionized species. For $\mathrm{d}\Delta = 0.1 - 100$, the curve is almost identical to $C_{\mathrm{gas}}(\mathrm{d}\Delta)$ in Figure 4.7 as we can re-write:

$$C_{\mathrm{gas}}^{'}(\mathrm{d}\Delta) = \mathrm{C}_{\mathrm{gas}}(\mathrm{d}\Delta) \times \frac{\langle \mathrm{n}_{\mathrm{gas}}(\mathrm{d}\Delta)\rangle^2}{\langle \mathrm{n}_{\mathrm{gas}}\rangle^2}, \tag{4.10}$$

and including cells with $\Delta > 100$ and $\Delta < 0.1$ has only a negligible effect on the mean gas density ($< 2\%$).

$C_{\mathrm{gas}}^{'}(\mathrm{d}\Delta)$ calculated for the other overdensity bins, i.e. $\mathrm{d}\Delta$ = 0.1-0.3, 0.3-1, 1-3, 3-10, 10-30 and 30-100, follows a similar shape with a difference in normalization caused by $\langle n_{\mathrm{gas}}(\mathrm{d}\Delta)\rangle / \langle \mathrm{n}_{\mathrm{gas}}\rangle$. The values of $\langle n_{\mathrm{gas}}(\mathrm{d}\Delta)\rangle / \langle \mathrm{n}_{\mathrm{gas}}\rangle$ in each $\mathrm{d}\Delta$ is close to the centre[1] of the overdensity range (in log scale; see Figure 4.8) due to the shape of the gas distribution in each bin as explained in the previous section. This leads to almost fixed shift (in log scale) between the curves of $C_{\mathrm{gas}}^{'}(\mathrm{d}\Delta)$ for different $\mathrm{d}\Delta$. The $\mathrm{d}\Delta = 0.1 - 100$ curve goes close to $\mathrm{d}\Delta = 1 - 3$ curve as most of the cells have an overdensity in the range $0.3 < \Delta < 3$.

Next we look at the other species, i.e. HII, HeII and HeIII. The curves for $\mathrm{d}\Delta = 0.1 - 100$ for all the species are similar to those in Figure 4.5. $C_{\mathrm{HII}}^{'}(\mathrm{d}\Delta)$ rises gradually and then flattens at about $z = 8$ which shows the reionization history of H in the box. $C_{\mathrm{HeII}}^{'}(\mathrm{d}\Delta)$ rises but then starts decreasing after $z = 8$, where as $C_{\mathrm{HeIII}}^{'}(\mathrm{d}\Delta)$ shows a steady rise. The decrease in $C_{\mathrm{HeII}}^{'}(\mathrm{d}\Delta)$ is due to the conversion of HeII to HeIII as discussed previously.

The normalization of $C_{\mathrm{HII}}^{'}(\mathrm{d}\Delta)$, $C_{\mathrm{HeII}}^{'}(\mathrm{d}\Delta)$ and $C_{\mathrm{HeIII}}^{'}(\mathrm{d}\Delta)$ for $\mathrm{d}\Delta = 0.1 - 100$ reflects the fraction of H/He present in the total gas. For other $\mathrm{d}\Delta$s, the curves show a redshift evolution similar to the case for $\mathrm{d}\Delta = 0.1 - 100$, but the normalization around the $\mathrm{d}\Delta = 0.1 - 100$ curve (for each species) is due to $\langle n_{\mathrm{gas}}(\mathrm{d}\Delta)\rangle / \langle \mathrm{n}_{\mathrm{gas}}\rangle$ as explained above. Note that at high redshifts, the curve for $\mathrm{d}\Delta = 0.1 - 100$ shows slightly higher values than the case of $\mathrm{d}\Delta = 1 - 3$ showing that at high redshifts, the clumping factor of the total IGM overdensities is dominated by that of the high density regions as those are the cells that get ionized earlier. Towards lower redshifts, the clumping factor values fall on top of the $\mathrm{d}\Delta = 1 - 3$ curve for HII and HeII. The values for HeIII at low redshifts stay higher that the $\mathrm{d}\Delta = 1 - 3$ curve, showing that high density cells dominate the HeIII clumping factor calculations at all redshifts.

4.3.4 Resolution Tests

In this Section we investigate how the behaviour of the clumping factor discussed previously is affected by the box and grid size, by performing a number of resolution tests.

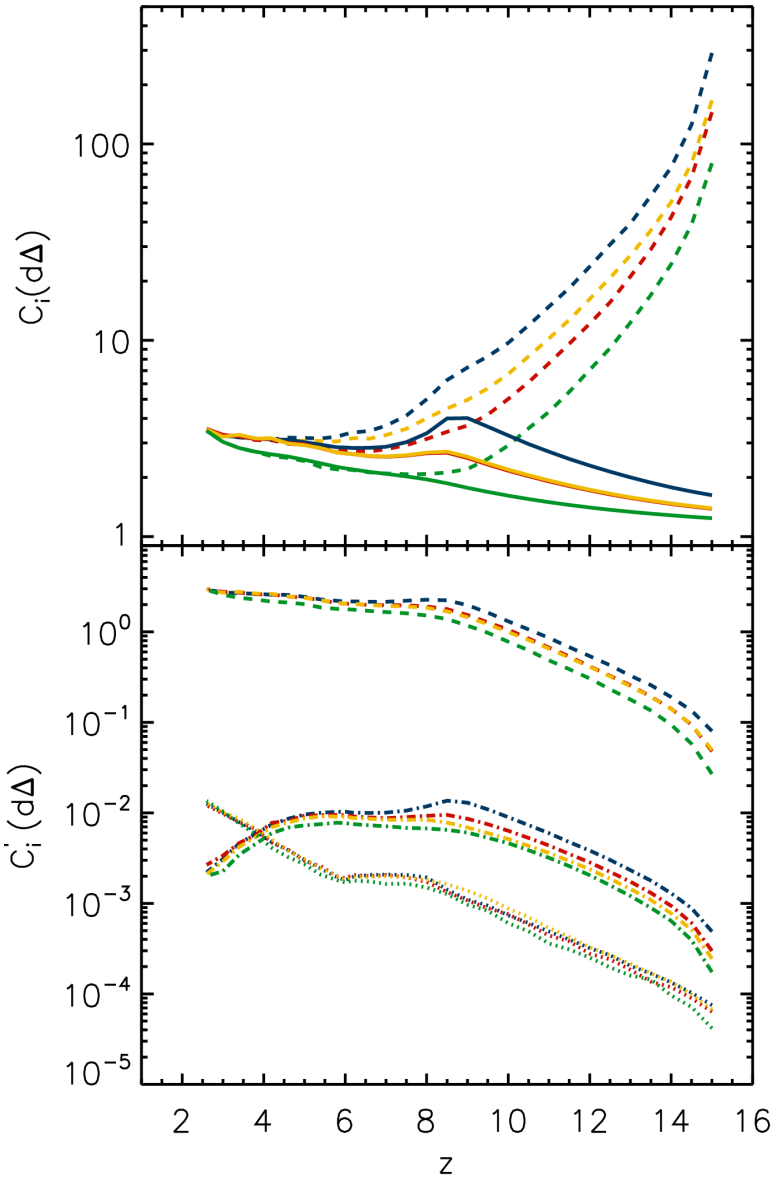

Figure 4.10: Evolution of the clumping factor $C_i(\mathrm{d}\Delta)$ (top panel) and $C_i'(\mathrm{d}\Delta)$ (bottom panel) for i = gas (solid line), HII (dashed line), HeII (dot dashed line) and HeIII (dotted line). In both panels, $\mathrm{d}\Delta = 0.1 - 100$. The lines refer to simulation $2.20G128$ (blue; reference), $2.20G64$ (red), $2.20G32$ (green) and $2.20G128S64$ (yellow). Refer to the text for more details.

Dependence on Grid Resolution

Hydrodynamic simulations have a spatial resolution equal to the softening length of the simulation. On the other hand, the resolution of the RT calculation is determined by the grid size, which, due to computational constraints, is limited to 128^3 cells for our reference simulations. This leads to a lower spatial resolution in the RT calculations. To understand the effect this has on the clumping factor calculations, we compare simulations $2.20G128$, $2.20G64$ and $2.20G32$ which have the same box size, but different gridding resolution. Difference in gridding leads to both smoothing of the initial density and temperature fields along with changes in ionization levels due to RT effects as discussed in Section 4.2. We also compare $2.20G64$ to results from the $2.20G128$ simulation smoothed to a 64^3 grid (referred to as $2.20G128S64$ for convenience). Note that $2.20G128S64$ is not a new RT simulation, but just a smoothing of the RT outputs.

Figure 4.10 shows the evolution of clumping factors $C_i(\mathrm{d}\Delta)$ and $C_i^{'}(\mathrm{d}\Delta)$ for $2.20G128$, $2.20G64$, $2.20G32$ and $2.20G128S64$. For $C_i(\mathrm{d}\Delta)$, we plot for i = gas and HII for $\mathrm{d}\Delta = 0.1 - 100$. HI/HeI follow similar behaviour to gas, while HeII/HeIII to HII and hence not plotted. Focusing on gas, we see that $C_{\mathrm{gas}}(\mathrm{d}\Delta)$ for $2.20G128$ has higher values ($\sim 50\%$) at high redshifts than $2.20G64$. Same is true for $2.20G64$ and $2.20G32$ showing that decreasing grid resolution leads to lower clumping factors. It is interesting to note that in $2.20G32$, the grid resolution is too low that the scale of gas smoothing is smaller than the cell size and thus the effect of radiative feedback of the UV background in the hydrodynamic simulations on gas clumping is not observed. Also note that $C_{\mathrm{gas}}(\mathrm{d}\Delta)$ for $2.20G64$ and $2.20G128S64$ have similar values (higher than that of $2.20G128$). This shows that gridding of the density field to lower resolution affects gas clumping factor calculations due to smoothing of the density field compared to the high resolution case. But for $C_{\mathrm{HII}}(\mathrm{d}\Delta)$, $2.20G128S64$ shows higher clumping factors than $2.20G64$ while still being lower than that of $2.20G128$ showing that the difference in gridding also affects ionizing photon radiative transfer leading to differences in ionization structure in $2.20G128$ and $2.20G64$. The latter has a smoother ionization structure with reionization proceeding at a faster and more uniform rate than the former. Thus we can conclude that lowering the grid resolution has a significant impact on the clumping factor calculations leading to lower clumping factor values.

The bottom panel of Figure 4.10 shows for $C_i^{'}(\mathrm{d}\Delta)$ for i = HII, HeII and HeIII. Since $C_{\mathrm{gas}}^{'}(\mathrm{d}\Delta) = \mathrm{C}_{\mathrm{gas}}(\mathrm{d}\Delta)$ (as discussed previously) and HI/HeI are complementary to HII/HeII/HeIII, we do not plot them here. Focusing on the curves for HII shows that even though there are differences in $C_{\mathrm{HII}}(\mathrm{d}\Delta)$ between the simulations $2.20G128$, $2.20G64$ and $2.20G128S64$, $C_{\mathrm{HII}}^{'}(\mathrm{d}\Delta)$ lines are consistent with each other. This is due to the lower mean gas density in low resolution outputs (both gridding and smoothing) canceling out the low clumping factors in those simulations. But $C_{\mathrm{HII}}^{'}(\mathrm{d}\Delta)$ for $2.20G32$ shows much lower values, suggesting that at least 64^3 cells are needed to resolve the evolution of $C_{\mathrm{HII}}^{'}(\mathrm{d}\Delta)$ for the $2.20\ h^{-1}\,\mathrm{Mpc}$ box.

$C_{\mathrm{HeII}}^{'}(\mathrm{d}\Delta)$ also shows differences for gridding resolution especially at high redshifts

[1] except for the edge bins - 0.1-0.3 and 30-100.

where higher resolution simulations have high clumping factors. For HeII, $2.20G64$ and $2.20G128S64$ have different clumping factors showing that cell density plays a very important role in determining the reionization history (as discussed in the previous section). HeIII being complimentary to HeII at low redshifts shows the reverse trend with higher $C'_{\rm HeIII}(\mathrm{d}\Delta)$ for lower grid resolution. This is mainly due to low number of very high density cells in low resolution simulations which dominate HeIII ionization structure. At high redshifts, due to poor statistics for high density cells, the trend is not very clear. Finally, we can note that the differences between the lines are larger in HeII and HeIII than HII showing the importance of grid resolution especially for He RT.

Dependence on Box Size

Next we look at the effect of the box size on the calculation of clumping factors. Changing the box size while keeping the same grid size changes both the spatial resolution and the source properties, as discussed in Section 4.1. These changes also affect the topology and the pace of reionization as seen in Section 4.2. In particular, more luminous sources together with lower density cells lead to faster reionization as seen in Figure 4.3. We investigate the effect of the box size using $2.20G128$, $4.39G128$ and $8.78G128$. To differentiate the effects due to a reduced spatial resolution and different source properties, we also use $2.20G64$, which has the same spatial resolution of $4.39G128$, but source distribution similar to $2.20G128$. Large comoving volume of $4.39G128$ would contain higher mass sources which are more clustered than the low mass sources in $2.20G128$ and $2.20G64$. Even though the detailed distribution of ionizing sources in each cell is different between $2.20G128$ and $2.20G64$, the ionizing photons produced in a cell of $2.20G64$ would be equivalent to that produced by the same subregion in $2.20G128$.

Figure 4.11 plots the evolution of $C_i(\mathrm{d}\Delta)$ and $C'_i(\mathrm{d}\Delta)$ with redshift. The top panel shows $C_i(\mathrm{d}\Delta)$ against z for i = gas and HII for $\mathrm{d}\Delta = 0.1 - 100$. We can see that at high redshifts, clumping factors of both gas and HII decrease as we go to larger box sizes from $2.20G128$ to $8.78G128$. But at $z < 6$, the trend flips with higher clumping factors for smaller boxes. We need to study the behaviour of $2.20G64$ to understand the reasons for the above trends. At high redshifts, we see that $C_{\rm gas}(\mathrm{d}\Delta)$ stays close to the $4.39G128$ curve but with slightly lower values. At low redshifts, $2.20G64$ joins with the $2.20G128$ curve. This shows that at high redshifts, grid resolution determines the gas clumping factor where as at low redshifts, source properties are more important. At high redshifts, source properties are mildly important which can be seen in the slightly higher gas clumping values of $4.39G128$ than $2.20G64$. $C_{\rm HII}(\mathrm{d}\Delta)$ curve for $2.20G64$ shows similar behaviour as that of gas but with a larger discrepancy compared to $4.39G128$ showing that source properties play a very important role in increasing ionization clumping factor even when the low resolution gridding pushes it the other way.

The bottom panel of Figure 4.11 shows $C'_i(\mathrm{d}\Delta)$ against z for i = HII, HeII and HeIII for $\mathrm{d}\Delta = 0.1 - 100$. As in the grid size test, we see that $C'_{\rm HII}(\mathrm{d}\Delta)$ does not differ much for $2.20G128$, $4.39G128$ and $2.20G64$. For $8.78G128$, the values are slightly lower at high redshifts which is probably due to the very low spatial resolution (equivalent to $2.20G32$).

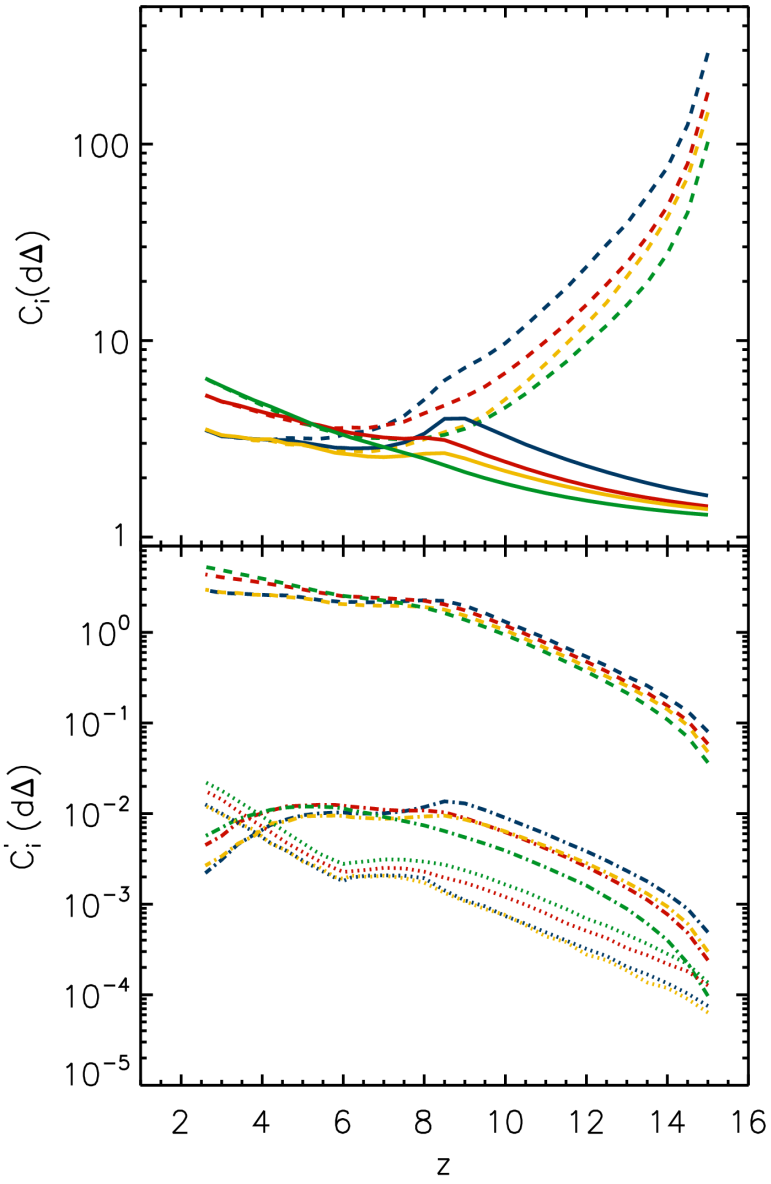

Figure 4.11: Redshift evolution of the clumping factor $C_i(\mathrm{d}\Delta)$ (top panel) and $C_i^{'}(\mathrm{d}\Delta)$ (bottom panel) for i = gas (solid line), HII (dashed line), HeII (dot dashed line) and HeIII (dotted line). In both panels, $\mathrm{d}\Delta = 0.1 - 100$. The simulations used in these plots are $2.20G128$ (blue, default), $4.39G128$ (red), $8.78G128$ (green) and $2.20G64$ (yellow).

But the discrepancy between the simulations are larger for HeII and HeIII. At high redshifts, $C'_{\rm HeII}(\mathrm{d}\Delta)$ shows a clear reduction with increasing box size. For HeIII, the trend is reversed at all redshifts with higher $C'_{\rm HeIII}(\mathrm{d}\Delta)$ for large boxes. The curve of $2.20G64$ lies above $4.39G128$ for HeII and below for HeIII showing that the source distribution plays an important role in converting a large fraction of HeII to HeIII at high redshifts. Thus we can conclude that box size plays a very important role in determining the clumping factors, especially of single/double ionized He.

4.3.5 Dependence on Mean Gas Density

Previous works have shown that the gas clumping factor correlates with the gas density (e.g. Kohler et al., 2005; McQuinn et al., 2007; Kohler, Gnedin, & Hamilton, 2007). Raičević & Theuns (2010) showed that sub-volumes within a simulation box have different gas clumping factors due to differences in the gas distribution. Interestingly, the average value of the clumping factors calculated within the different sub-volumes is much lower than the clumping factor calculated for the whole box. To investigate this further, we split $4.39G128$ into 8 sub-boxes each with size 2.20 h^{-1}Mpc of 64^3 cells. Each of these sub-boxes is equivalent to $2.20G64$ albeit with different gas distribution and source properties.

In Figure 4.12, similarly to Figure 4.2, we plot the normalized PDF of the gas number density, temperature and ionization within the sub-boxes at $z = 14.5$ and 3.8. As we can see, the different sub-boxes have a large variations in density distribution with differences being larger at low redshifts. Regions with high mean density have a generally wider density distribution at all redshifts. Temperature also shows a similar trend with high overdensity regions having higher fractions of high temperatures. Ionization fraction also shows a similar trend at high redshifts with high overdensity region showing higher ionization levels. At low z, a smaller fraction of the cells in a low overdensity sub-box is fully ionized compared to a high density one. These are due to the high mass objects present in that sub-volume. Next we look at the ionization history of the different sub-boxes. Figure 4.13 shows the ionization fraction of HII, HeII and HeIII against z. We can see that the higher the mean density of the cell is, the faster is the reionization, which is again due to the high mass objects present in that volume pumping in more ionizing photons leading to a faster reionization of the region. Thus we can suspect that this would affect clumping factor evolution.

Figure 4.14 shows the redshift evolution of $C_i(\mathrm{d}\Delta)$ and $C'_i(\mathrm{d}\Delta)$ of $\mathrm{d}\Delta = 0.1 - 100$ for the 8 sub-volumes of $4.39G128$. The top panel shows $C_{\rm gas}(\mathrm{d}\Delta)$ and $C_{\rm HII}(\mathrm{d}\Delta)$ both of which shows a huge range in clumping factors (a factor of 3-10 in range). The lines for the different sub-boxes are colored according to the overdensity of the region at $z = 15$, i.e. $\Delta_{\rm sub-box}(z = 15)$ (refer to the next section for more details). High overdensity regions have higher fraction of cells with high gas density, leading to larger gas clumping factors. The range in gas clumping factors is increasing with decreasing redshift. But in the case of HII, the trend goes the other way as high overdensity regions host higher mass sources leading to faster ionization of the region giving low HII clumping factor values. Also, the range is wider at high redshifts. But note that the clumping factors do not obey a strict

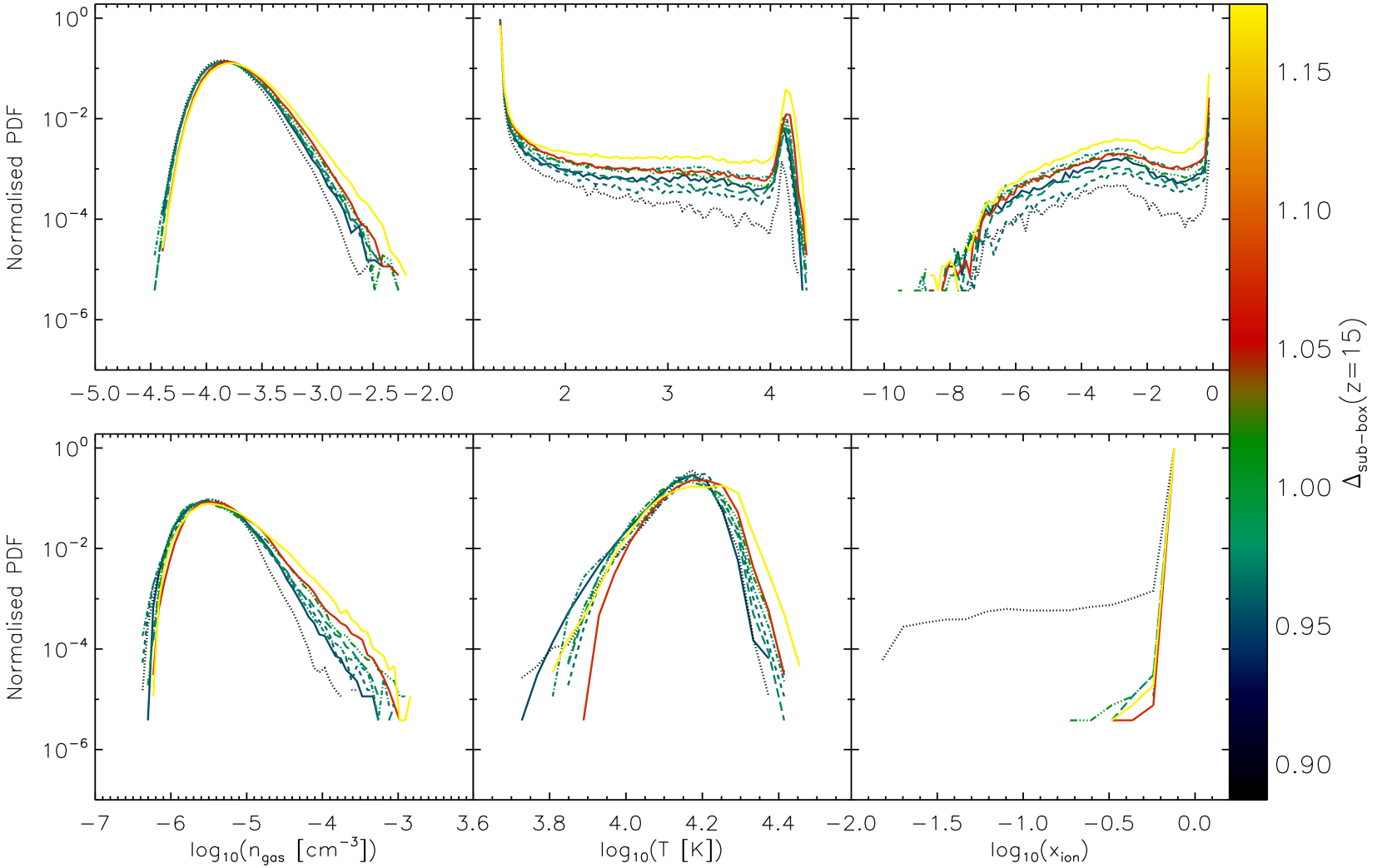

Figure 4.12: Normalized PDF of gas number density (left column), temperature (middle column) and ionization fraction (right column) in simulations at redshifts $z = 14.5$ (top row) and $z = 3.8$ (bottom row). The different lines are for the 8 sub-boxes of 64^3 each from the $4.39G128$ simulation. The lines are colored according to $\Delta_{\rm sub-box}(z = 15)$, i.e. the overdensity of the sub-box at $z = 15$.

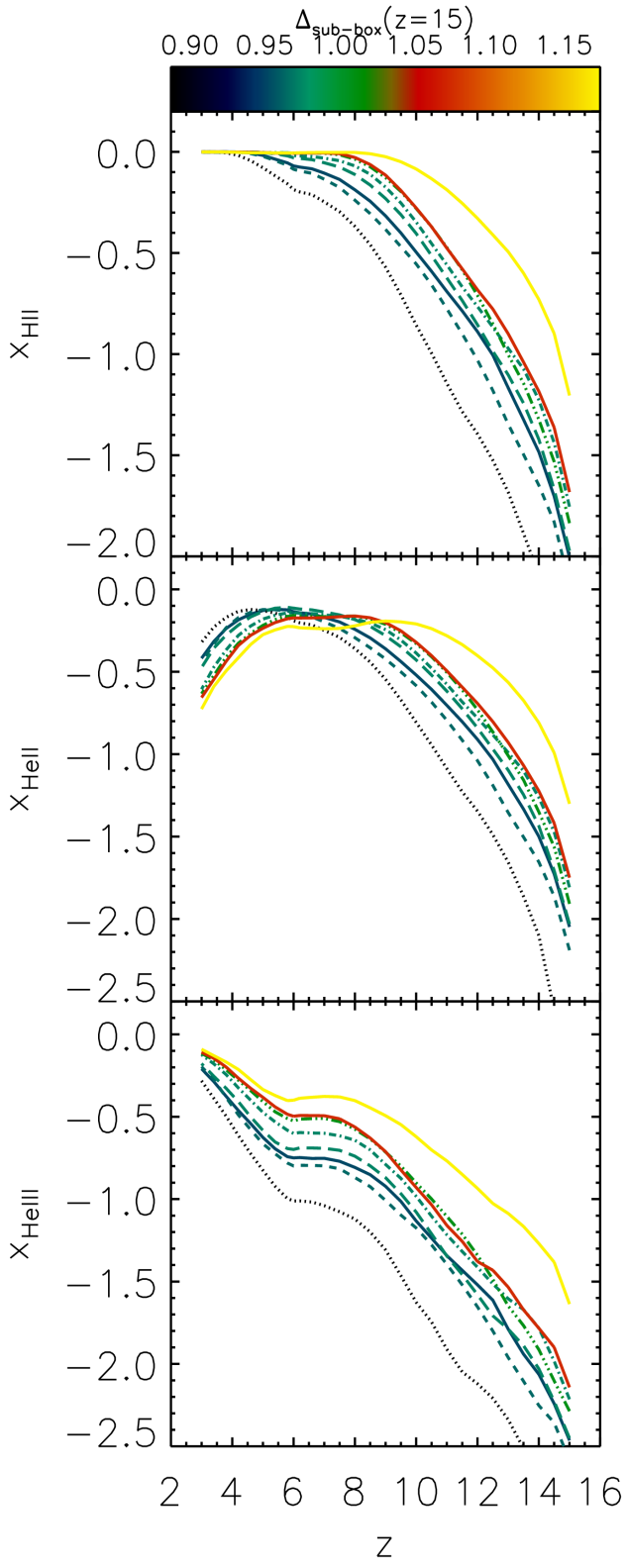

Figure 4.13: Redshift evolution of $x_{\rm HII}$ (upper panel), $x_{\rm HeII}$ (middle panel) and $x_{\rm HeIII}$ (lower panel) fraction. The different lines are for the 8 sub-boxes of 64^3 each from the $4.39G128$ simulation. The lines are colored according to $\Delta_{\rm sub-box}(z = 15)$, i.e. the overdensity of the sub-box at $z = 15$.

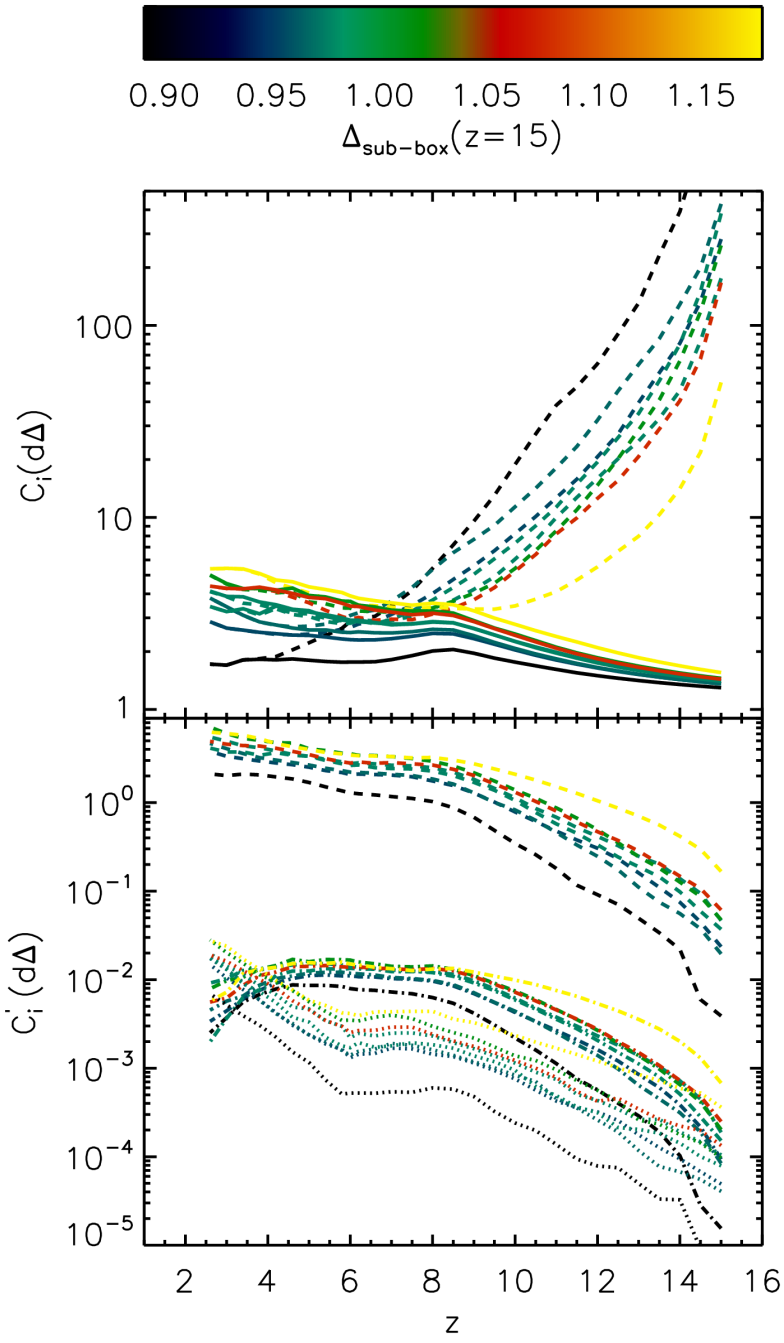

Figure 4.14: Evolution of the clumping factor $C_i(\mathrm{d}\Delta)$ (top panel) and $C_i'(\mathrm{d}\Delta)$ (bottom panel) for i = gas (solid line), HII (dashed line), HeII (dot dashed line) and HeIII (dotted line). In both panels, $\mathrm{d}\Delta = 0.1 - 100$. The different lines are for the 8 sub-boxes of 64^3 each from the $4.39G128$ simulation. The lines are colored according to $\Delta_{\mathrm{sub-box}}(z = 15)$ which is the overdensity of the sub-box at $z = 15$.

correlation with the mean gas overdensity of the region but do show a general trend with lower $C_{gas}(\mathrm{d}\Delta)$ for low overdensity regions. The bottom panel shows $C_i'(\mathrm{d}\Delta)$ for $i =$ HII, HeII and HeIII. Again, we see a large range in clumping factors from the sub-boxes for all the three species. Higher overdensity regions, show higher clumping factors and faster reionization in the region. The differences are larger at lower redshifts and can span orders of magnitude for all three species. This will be studied in detail in the next section. At low redshifts, the differences span a factor of few but HeIII shows more variance compared to HII and HeII due to the dependence on high density regions for HeIII ionization. Thus we see that the mean density of the region play a very important role in determining the reionization history and the clumping factor of the region. Compared to grid size and box size, the mean density of the region seems to be the dominant factor in clumping factor calculations

Dependence on Initial Conditions of Hydrodynamic Simulations

Since the gas density plays an important role in determining the clumping factor, we now explore the effect of initial conditions of hydrodynamic simulations for regions with the same mean density. To investigate this further, we use two more simulations $2.20G128v2$ and $2.20G128v3$ with the same mean gas number density as $2.20G128$ but with different realizations of initial conditions, thus leading to different source and gas distributions.

Figure 4.15 shows the normalized PDF of density, temperature and ionization distribution of $2.20G128$, $2.20G128v2$ and $2.20G128v3$ at $z = 14.5$ and 3.8. As it can be seen, both density and temperature distribution in the different simulations are the same. Only the ionization distribution shows mild differences at low redshifts. Moving to the ionization history, Figure 4.16 shows the evolution of x_{HII}, x_{HeII} and x_{HeIII} with redshift. Also the ionization history of the three simulation boxes is very similar with differences much smaller that the ones seen in the previous section. This manifests also in the calculations of $C_i(\mathrm{d}\Delta)$ for $i =$ gas, HII and $C_i'(\mathrm{d}\Delta)$ for $i =$ HII, HeII, HeIII where $\mathrm{d}\Delta = 0.1 - 100$ (Figure 4.17). Again we see that the differences between these simulations are mild and are present mostly at low redshifts. Compared to HII especially at low redshifts, the differences seem slightly larger for HeII and HeIII, showing the importance for detailed gas and source distribution for He reionization. Thus we can say that once the mean density of a region is fixed, the initial conditions only mildly affect the clumping factor calculations.

4.3.6 Clumping Factor-Overdensity Correlation

Since the main factor which leads to a spread in clumping factor at each redshift z is the mean overdensity of the cell, in this section we try to parameterise the correlation between these quantities. To do this we divide $8.78G128$ into 64 sub-boxes of 32^3 grid cells each which would be equivalent to 64 different $2.20G32$ simulations. Even though this grid size of 32^3 is not sufficient for proper RT (as seen in Section 4.3.4), this is enough to obtain an approximate trend. Using sub-boxes of $4.39G128$ split to sub-volumes of size 2.20 h^{-1}Mpc as in Section 4.3.5 gives better resolution by poorer statistics due to the lesser number of

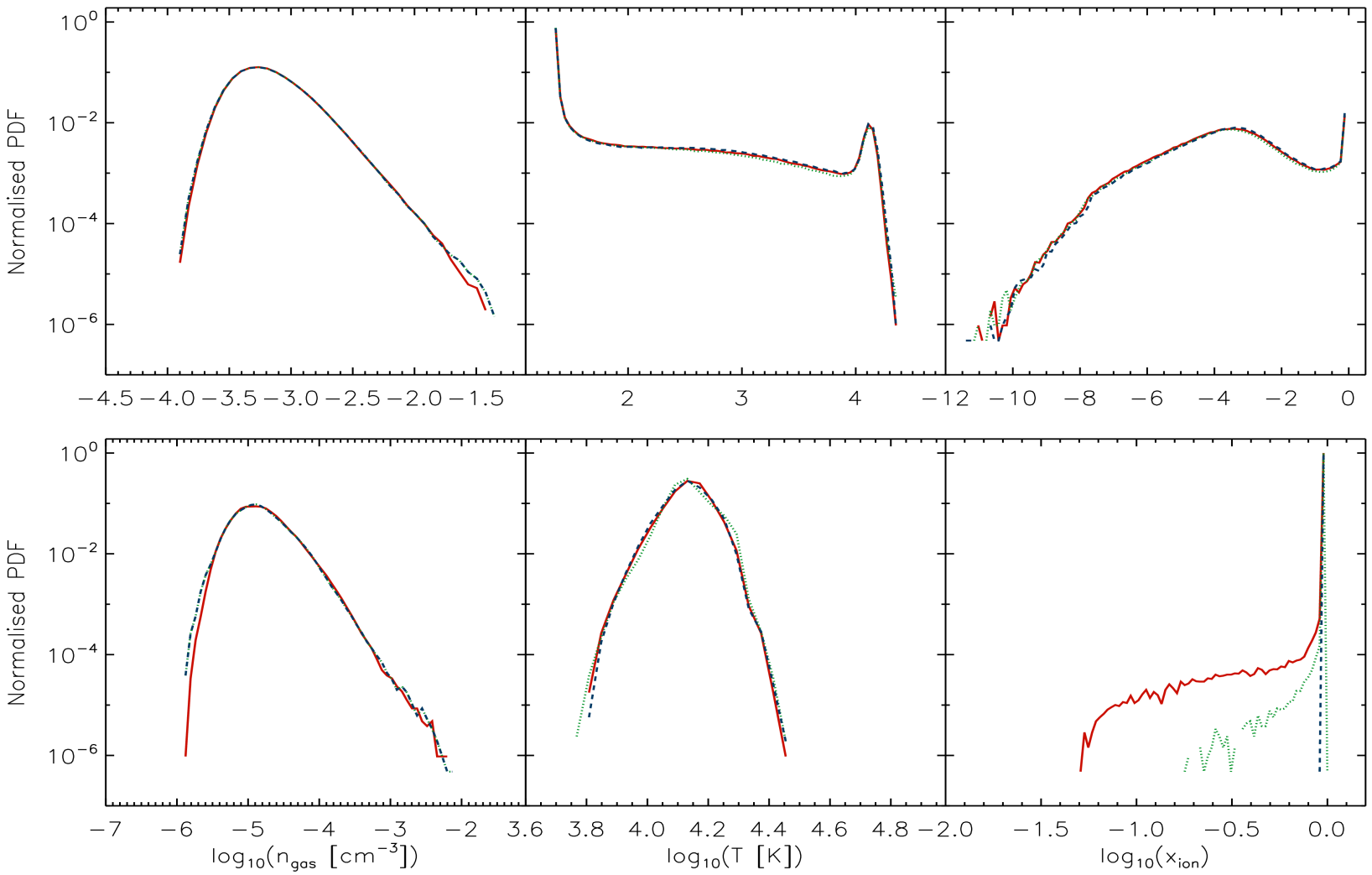

Figure 4.15: Normalized PDF of gas number density (left column), temperature (middle column) and ionization fraction (right column) in simulations at redshifts $z = 14.5$ (top row) and $z = 3.8$ (bottom row). The different lines are for - $2.20G128$ (red solid line), $2.20G128v2$ (green dotted line) and $2.20G128v3$ (blue dashed line)

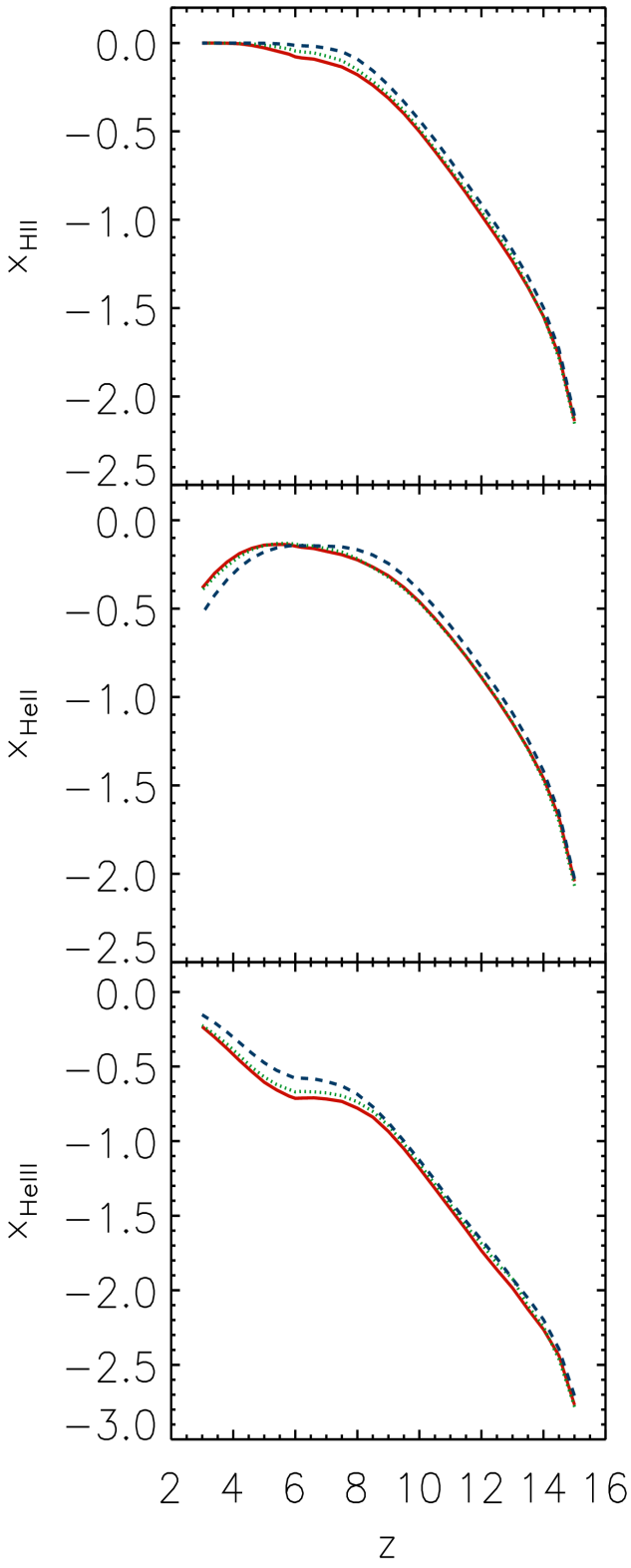

Figure 4.16: Evolution of $x_{\rm HII}$ (upper panel), $x_{\rm HeII}$ (middle panel) and $x_{\rm HeIII}$ (lower panel) fraction with redshift. The different lines are for a box 2.20 h^{-1}Mpc with 3 different initial conditions - $2.20G128$ (red solid line), $2.20G128v2$ (green dotted line) and $2.20G128v3$ (blue dashed line).

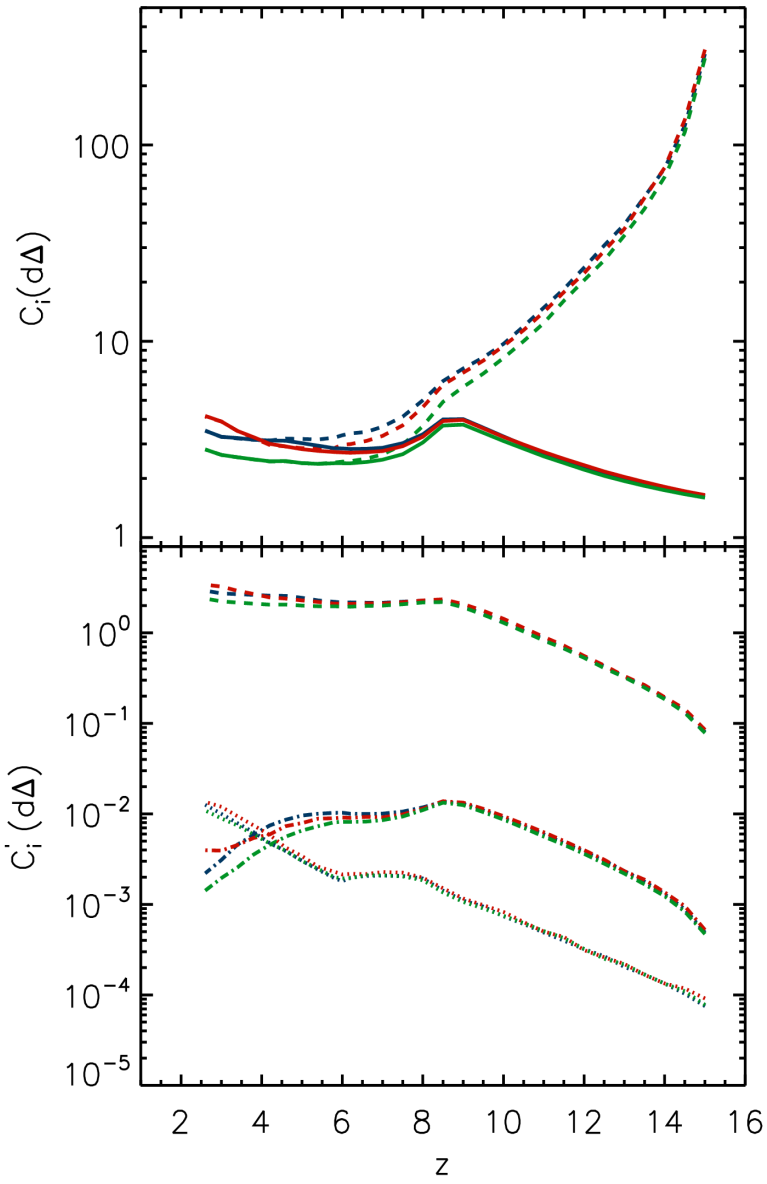

Figure 4.17: Redshift evolution of the clumping factor $C_i(\mathrm{d}\Delta)$ (top panel) and $C_i'(\mathrm{d}\Delta)$ (bottom panel) for i = gas (solid line), HII (dashed line), HeII (dot dashed line) and HeIII (dotted line). In both panels, $\mathrm{d}\Delta = 0.1 - 100$. The lines in these plots are for a box 2.20 h^{-1}Mpc with 3 different initial conditions - $2.20G128$ (red solid line), $2.20G128v2$ (green dotted line) and $2.20G128v3$ (blue dashed line).

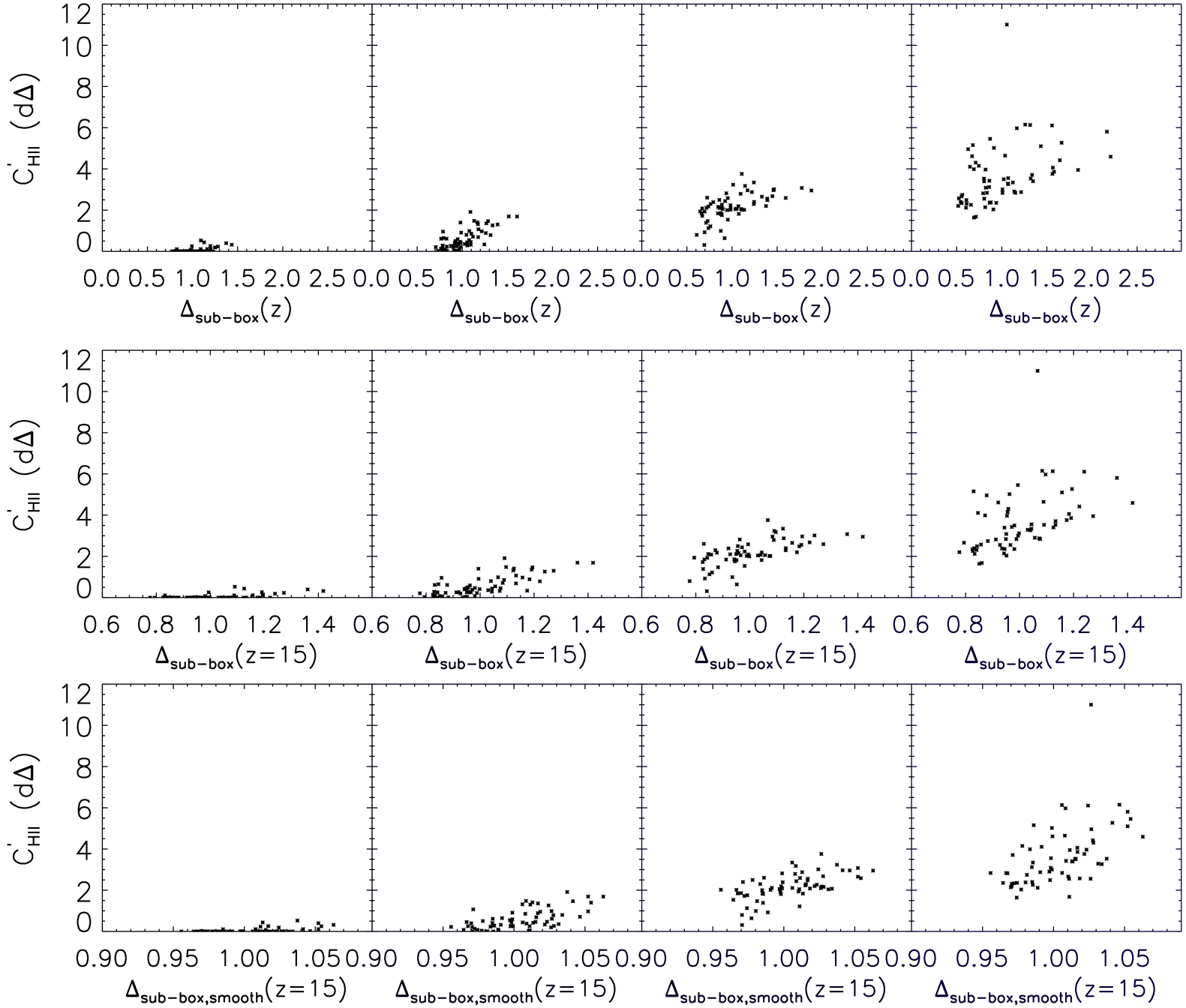

Figure 4.18: $C'_{\rm HII}(\mathrm{d}\Delta)$ as a function of the overdensity of the sub-box for each of the 64 sub-boxes of the 32^3 grid cells from the $8.78G128$ simulation. From left to right the columns refer to $z = 15$, 10, 6.6 and 2.6. The overdensities on the x-axis are - $\Delta_{\rm sub-box}(z)$ (top row), $\Delta_{\rm sub-box}(z = 15)$ (middle row) and $\Delta_{\rm sub-box,smooth}(z = 15)$ (bottom row).

sub-boxes. From Figure 4.10, we saw that the clumping factor values from $2.20G32$ are lower by a factor of a few compared to values from $2.20G128$ but the general trends seen in this analysis would be the same. Thus we can use these for the current analysis.

Then we estimate the overdensity of each of the sub-boxes ($\Delta_{\rm sub-box}(z)$) with respect to the mean density of the universe at any given redshift. As the Universe expands with redshift, the mean density of the Universe decreases. Also due to the flow of gas in the IGM from the voids to the overdense regions (seen in Figure 1.3), the mean density of the sub-box also changes mildly with redshift, leading to a redshift evolution of $\Delta_{\rm sub-box}(z)$. Since gas, HeII and HeIII show trends similar to HII, we will discuss only the case of HII. The top row panels of Figure 4.18 show $C'_{\rm HII}({\rm d}\Delta)$ for each of the sub-boxes calculated for ${\rm d}\Delta = 0.1-100$ plotted against $\Delta_{\rm sub-box}(z)$ at $z = 15,\ 10,\ 6.6$ and 2.6 and plotted in linear scale to show the trend and the scatter clearly. At high redshifts, the clumping factors have values $\lesssim 0.1$ and thus are not clearly visible in the linear-linear plot. But as the redshift decreases, the clumping factor increases till the region in the sub-box reionizes and the clumping factor value converges to an equilibrium value. For high overdensity sub-boxes, this is reached faster halting the evolution of clumping factors at a higher redshift.

As explained earlier, also present in $\Delta_{\rm sub-box}(z)$ is the effect due to change of overdensity of a sub-box with redshift. Therefore, in the middle row panels, we look at $C'_{\rm HII}({\rm d}\Delta)$ (as in the previous paragraph) plotted against $\Delta_{\rm sub-box}(z = 15)$, i.e. the overdensity of the sub-box as it was at $z = 15$. Even though the trend looks similar to that in the top row, note that some of the data points have a different location in this row compared to the top one especially the sub-box with a very high clumping factor. This is due to the change in overdensity of sub-boxes with redshift due to interaction with neighboring sub-boxes as gas from some sub-boxes flow toward higher density regions in other sub-boxes.

To understand this further, we plot $C'_{\rm HII}({\rm d}\Delta)$ against an overdensity parameter which has been smoothed $\Delta_{\rm sub-box,smooth}$ to take into account the overdensity of the neighboring sub-boxes. This is important as a low overdensity region close to a high overdensity region will get ionized faster than the average rate. $\Delta_{\rm sub-box,smooth}$ is obtained by smoothing the overdensity of a cube of 27 sub-boxes centred on the selected sub-box. This gives us the averaged overdensity smoothed over a cube including the 26 neighbouring sub-boxes for each sub-box. $\Delta_{\rm sub-box,smooth}$ is calculated for $z = 15$ in our plots. The bottom panel of Figure 4.18 shows the correlation between $C'_{\rm HII}({\rm d}\Delta)$ and $\Delta_{\rm sub-box,smooth}(z = 15)$. The trend with redshift is similar to the top panels. But at low redshifts, we see a dependency of $C'_{\rm HII}({\rm d}\Delta)$ of smooth overdensity albeit with a large spread. Even though the sub-box with the very high value of clumping factor does not fall on the trend, it does lie in a high density neighborhood where the clumping factors are generally seen to be higher.

Thus these plots show that neighboring sub-boxes have an effect (even though small) on the clumping factor of a certain region, due to their ionization fronts.

4.4 Discussion and Summary

In this chapter, we studied the effect of box size and gridding resolution on reionization simulations. We used boxes ranging from $2.20 - 35.12\ h^{-1}$Mpc and RT grid sizes from $32^3 - 128^3$. We found box sizes play an important role in determining source properties which seem important for proper estimations of temperature and ionization distribution, especially at low redshifts. Grid size determines the density range of the cells and controls the ionization rate of the box. Large box sizes and high spatial resolution is crucial for the proper modelling of He ionization.

Simulating large boxes with high resolution is a very difficult task. To aid that, we calculate the clumping factors of the different species from our high resolution simulation $2.20G128$. We found that at high redshifts, ionized species have higher clumping factors which converges with gas clumping factors at low redshifts. High overdensity cells close to the sources have low clumping factors as they are easily ionized. Clumping factor evolution also depends on the definition used. We also study the different factors which affect clumping factor calculations. Both box size and grid size does affect clumping factor calculation with larger influence on He ionizations. Overdensity of the region is the most important factor which determines the ionization history of the region and clumping factors. At a specific overdensity, differences in initial conditions only have a mild effect. Finally we find that there is a correlation between the overdensity of a region and the clumping factors and the trend evolves with redshift. Also neighbouring region around the one under consideration can have an effect in the clumping factors calculated from the region. Using all these information we can simulate the ionization history of larger regions which can then be used as input to LAE modelling.

Chapter 5

Conclusions

The Epoch of Reionization (EOR) is an interesting and important event in the history of the Universe, the time being when H in the gas changed from a mainly neutral to a highly ionized state. The details of the process are still unclear and are the topic of a large number of theoretical and observational studies. Lyman Alpha Emitters (LAEs), due to their strong Lyα emission line which scatters for even tiny amounts of neutral hydrogen, are used as one of the tools to probe the EOR. Modelling LAEs is a complex task with a large number of parameters and scales being important. Different studies either focus on a detailed understanding of individual aspects of LAEs or use numerical/semi-analytic approaches to model the LAEs and compare them to observations. In this thesis, we focus on the Inter-Galactic Medium (IGM) close to the ionizing/Lyα source and its effect on the observability of LAEs through Lyα radiative transfer effects. The IGM close to the source is important: the density, temperature and ionization structure controls the Lyα Radiative Transfer (RT), determining the intensity and Surface Brightness (SB) distribution along different lines-of-sight.

To achieve this, we simulated a sample of more than 100 LAEs, using galaxies and surrounding IGM extracted from simulation boxes of 5-30 h^{-1}Mpc at $z = 7.7$. Coupled RT of ionizing and Lyα photons was performed using CRASHα, to determine the ionization structure carved in the IGM by the ionizing photons exiting from the galaxies and the spectrum of the Lyα photons scattering through the remaining neutral hydrogen along different lines-of-sight. The outputs were also used to produce Lyα surface brightness maps for the lines-of-sight perpendicular to the six sides of each simulation cube. Analyses were done to study individual objects as well as statistical trends. A parameter study was also undertaken to understand the dependence of the results on the different, currently uncertain factors involved in the simulations. The following conclusions were drawn from our analysis.

- Inhomogeneities in the IGM affects Lyα RT, leading to structure in the simulated SB maps of LAEs. There are huge variations in the total flux in SB maps along different lines-of-sight for the same object. This leads to $\sim 30\%$ more scatter in the observed luminosity-mass relation than the intrinsic one for our sample of simulated LAEs.

This also leads to a spread in the values for the escape fraction of Lyα photons from the IGM: $f_{\rm esc,Ly\alpha,IGM} = 0.73 \pm 0.18$ (1σ). Note that some lines-of-sight, especially through voids, can have $f_{\rm esc,Ly\alpha,IGM} > 1$ due to Lyα photon contributions from other lines-of-sight by scattering and higher probability of Lyα photon escape towards the observer through voids.

- Observational campaigns have surface brightness thresholds, $SB_{\rm th}$, which limit the fraction of the flux observed from the SB distribution. At low $SB_{\rm th}$ (e.g. 10^{-20} ergs s^{-1} cm^{-2} arcsec^{-2}), the observability of the LAE strongly depends on the IGM ionization structure and velocity field around the objects. Therefore, in observational campaigns at very low $SB_{\rm th}$, a single detection within the observed comoving volume need not be that of the most massive LAE in the region but could be a lower mass object seen through a biased line-of-sight. At higher $SB_{\rm th}$, more detections are possible, but the observed luminosity values can vary by orders of magnitude for a single object depending on the line-of-sight. To obtain the total flux from the object along each line-of-sight, one needs deep observations at $SB_{\rm th} \sim 10^{-25}$ ergs s^{-1} cm^{-2} arcsec^{-2}. Thus we find that the impact of surface brightness thresholds on estimates of observed luminosity is very important and needs to be taken into account in the calculation of luminosity functions of LAEs from observational campaigns.

- One of the main factors which affect the Lyα RT through the IGM is the wavelength of the Lyα photons when they escape from the ISM into the IGM. Outflows/winds in the ISM would redshift the photons reducing their probability of scattering in the IGM and aiding the escape of Lyα photons towards the observer. We find that the higher the redshifting of the Lyα photons before entering the IGM is, the more concentrated is the SB profile, making the detection easier at low $SB_{\rm th}$ and improving the observability of LAEs.

- The ionization structure of the IGM also plays a very important role in Lyα RT thus determining the SB profiles. We estimate the effects due to an ionizing background using RT simulations with an initial non-zero uniform ionization level. At low levels of mean ionization in the IGM ($x_{\rm ion} \sim 0$), the Lyα photons undergo scattered diffusion through the IGM and the SB profiles thus produced are more extended and faint, making it harder to observe the total flux from the object. At high levels of mean IGM ionization ($x_{\rm ion} > 0.5$), the photons scatter less, leading to a more point source-like SB profile, making it easier to detect more flux for low $SB_{\rm th}$. Clustering of sources also affects the ionization structure around the object by making the region more ionized, but with a complex structure for the ionized bubble due to the distribution of neighbouring ionizing sources and IGM gas distribution. The qualitative effect on the SB profiles due to clustering is similar to that of a uniform ionization case with the same mean ionization level. But the detailed distribution of flux along different lines-of-sight vary due to the differences in the detailed IGM ionization structure.

Thus, this shows that proper treatment of the ionization structure in the region is important for better modelling LAEs.

- Due to the difficulty in achieving very high $SB_{\rm th}$ in observational campaigns, stacking of SB maps can be used to extract more information from the current LAE samples mapped upto relatively lower $SB_{\rm th}$. We find that the mean/median stacked SB profile of LAEs becomes steeper at higher mean IGM ionization levels thus giving an additional technique to use LAEs to study EOR.

We can thus conclude from the study that the IGM close to the source plays a very important role in determining the SB profiles of LAEs.

As proper modelling of the ionization structure in the IGM is crucial for a better description of LAEs, the second part of this thesis explores the aspect of simulating the ionization distribution and history of a region. Simulating the ionization history of a representative volume of the Universe needs large box sizes, to encompass the patchy nature of EOR and the massive ionized bubbles created towards the end of reionization. But high spatial/density resolution is also necessary to resolve the high density Lyman Limit Systems which control the rate of evolution of reionization history towards low redshifts. Since simulating large comoving volumes with very high spatial resolution is a difficult task, the general approach is to simulate large volumes at a lower resolution but using 'clumping factors' to achieve the correct ionization level in a cell as would be obtained from a higher resolution RT simulation. Recent work has shown that He along with H in the gas plays an important role in determining the temperature and ionization structure of the IGM. In this work, we study the different factors which affect the reionization history of the IGM and the estimation of clumping factors using these simulations. We analyse a suite of simulations 2.20-35.12 h^{-1}Mpc at $32^3 - 128^3$ RT grid sizes to understand the important factors which govern reionization simulations at different redshifts. We also calculate clumping factors from the highest resolution simulation of the IGM we have and study the different factors which affect clumping factor calculations. The conclusions made are as follows.

- We used different box sizes and grid resolution to learn that both source properties (mass distribution and clustering behaviour) and gas density resolution are important for ionization history. In particular, for high redshifts, source properties are more relevant, while at low redshifts, the IGM density distribution plays a more important role.

- We study the clumping behaviour $C_i = \langle n_i^2 \rangle / \langle n_i \rangle^2$ of different species, where n_i is the grid cell number density for each species i = gas, HI, HII, HeI, HeII and HeIII. This definition of clumping factor estimates the spread in the number density values for each species in grid cells with respect to a mean value calculated for the whole simulation volume. This helps us understand the uniformity in the distribution of each species in the simulation box as the process of reionization proceeds. At high redshifts, the behaviour of total gas, HI and HeI are similar with low clumping values

($\sim 1.5-3$), while HII, HeII and HeIII show similar high values ($\sim 10-500$) for the corresponding clumping factors. HeIII has more clumpiness than HeII (about a factor of 3) due to higher ionization energies and the dynamics between the two ionization states of He. At low redshifts, HII and HeII follow the gas while HeIII has slightly higher values. HI and HeI are negligible in the gas and thus show very high values of clumping factors.

- The clumping factors also depend on the value and the range of the overdensity of the cells taken into account in the calculations especially at high redshifts. We split the range of overdensity values (0.1-100) of cells in the simulations into narrow bins of 0.5 dex width and calculate the clumping factors within each bin. We find that the clumping factor is large for bins with low overdensity values and vice versa as the process of reionization proceeds from high density regions to low density ones.

- We also investigated the behaviour of another commonly used definition of clumping factor $C_i^{'} = \langle n_i^2 \rangle / \langle n_{\rm gas} \rangle^2$. This definition gives an easier way to estimate the number of recombinations in the gas for each ionized species $\propto n_i^2$ with respect to the mean gas density in the simulation volume. We find that at high redshifts, even though the values of C_i are high for HII/HeII/HeIII, $C_i^{'}$ is low ($\ll 1$) as only small pockets of gas are ionized and the fraction of ionization is the box is very low. As redshift decreases, more of the Universe is ionized and $C_{\rm HII}^{'}$ increases to ~ 3. For He ionizations, $C_{\rm HeII}^{'}$ and $C_{\rm HeIII}^{'}$ have values about two orders of magnitude lower (as He is only 8% of gas) and show complex evolution with redshift due to the higher ionization energy for HeIII compared to HeII and the conversion of HeII to HeIII.

- Different parameters affect the clumping factor values calculated from an individual simulation box: RT grid size, simulation box size, mean gas density of the region, detailed source distribution within the region. The main factor which leads to variations in the clumping factor values is the mean gas density of the region where it is estimated. For example, high overdensity regions give high clumping factor values $C_{\rm HII}^{'}$ for HII. At a fixed mean gas density of the region, the detailed source distribution only mildly affects the clumping factor values.

Thus we can conclude that a clumping factor is a complex parameter depending on numerous factors such as box size, grid size, mean gas density, species, redshift, etc and these must be taken into account while calculating and using it.

Future work includes simulating the reionization history of a large comoving volume using these clumping factors and extracting the IGM information from that simulation to better model LAEs at $z = 6-10$ to study EOR.

Appendix A

Criterion for the Removal of the ISM

In this work, we focus on resolving the IGM close to the object. The simulations we use do not have the resolution to resolve the structure in the ISM, which has been shown to be crucial to get the correct values of escape fraction of ionizing photons from the galaxy (e.g. Paardekooper et al., 2011). Thus we remove the high density cells from the gridded density field which represent the ISM. We pick the largest object in L05 to perform our tests. The IGM around the object was extracted for a cube of size $\sim 12\ r_{200}$ where r_{200} is the radius in comoving units at which the mean DM density inside the sphere is 200 times the critical density of the universe at that redshift. The rest of the parameters for the runs were set as in Section 2.3.3. Figure A.1 shows a slice through the ionization structure of the simulations where the ISM cells above different density thresholds - 0.1 cm^{-3}, 0.05 cm^{-3}, 0.03 cm^{-3} and 0.01 cm^{-3} have been removed. To perform the calculation, all cells within a radius of $\sim 0.7 \times r_{200}$ from the source with densities above the threshold value were set to 10^{-35} cm^{-3}. The choice of the radial distance is to make sure that only ISM cells of the reference galaxy are removed without affecting the other high density regions present in the cube. For a threshold value of 0.1 cm^{-3}, only a ring of gas is ionized as all the ionizing photons get trapped within the high density gas which hasn't been removed. Only if we lower the threshold, the photons escape into the IGM. As expected, the lower the threshold, the larger the ionized region is. We can see that for density thresholds below 0.05 cm^{-3} the shape of the ionized regions starts to converge. We have checked that the exact choice of the threshold does not affect our main results as long as it is in the range $0.05 - 0.01$ cm^{-3}. For this reason we choose the value 0.03 cm^{-3}.

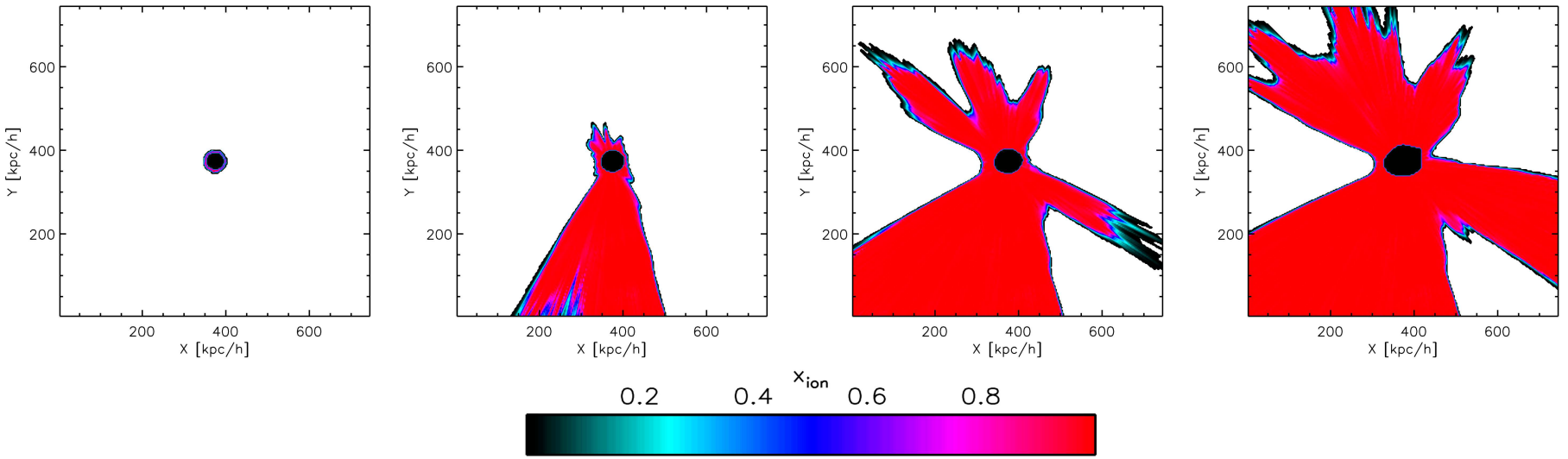

Figure A.1: A slice through the ionization structure in the cube for simulations with different density thresholds for ISM. The degree of ionization of the gas is shown by the color bar at the bottom. The different simulations are for ISM density thresholds of 0.1 cm^{-3} (left panel), 0.05 cm^{-3} (centre left panel), 0.03 cm^{-3} (centre right panel) and 0.01 cm^{-3} (right panel).

Appendix B

Criterion for the Cube Size

In Appendix A, we chose the density threshold to remove the high density cells which represent the ISM in our simulation cubes. For the value we chose, i.e. 0.03 cm^{-3}, we can see in the centre right panel of Figure A.1 that the ionization region is not fully contained within the simulation box, which is needed for the correct treatment of Lyα RT. Thus we test for the cube size required for our simulations. The input fields are always gridded for 256^3 cells which would lead to varying density and spatial resolution for these simulations. Thus we need to chose a cube size which is large enough to contain the ionized bubble but small enough to give a high density resolution for our RT simulations. Figure B.1 shows the ionization structure through the simulation cubes of size $\sim 12\ r_{200}$, $\sim 23\ r_{200}$, $\sim 35\ r_{200}$, $\sim 47\ r_{200}$, where r_{200} is the radius in comoving units at which the mean DM density inside the sphere is 200 times the critical density of the universe at that redshift. The simulation for $\sim 12\ r_{200}$ is the same as the centre right panel of Figure A.1, where we can see that the ionization structure is not confined within the cube. As we increase the size of the cube, this becomes large enough to contain the ionized bubble. A cube size of $\sim 23\ r_{200}$ is large enough to contain the bubble but not with much distance between the edge of the bubble and the edge of the box. Simulations with cube size $\sim 35\ r_{200}$ and $\sim 47\ r_{200}$ are both large enough to hold the ionized region, with $\sim 35\ r_{200}$ providing a higher spatial resolution. Thus we chose a cube size of $\sim 35\ r_{200}$ for our RT simulations. The main results do not seem to be affected too much for box sizes between $25 - 50\ r_{200}$.

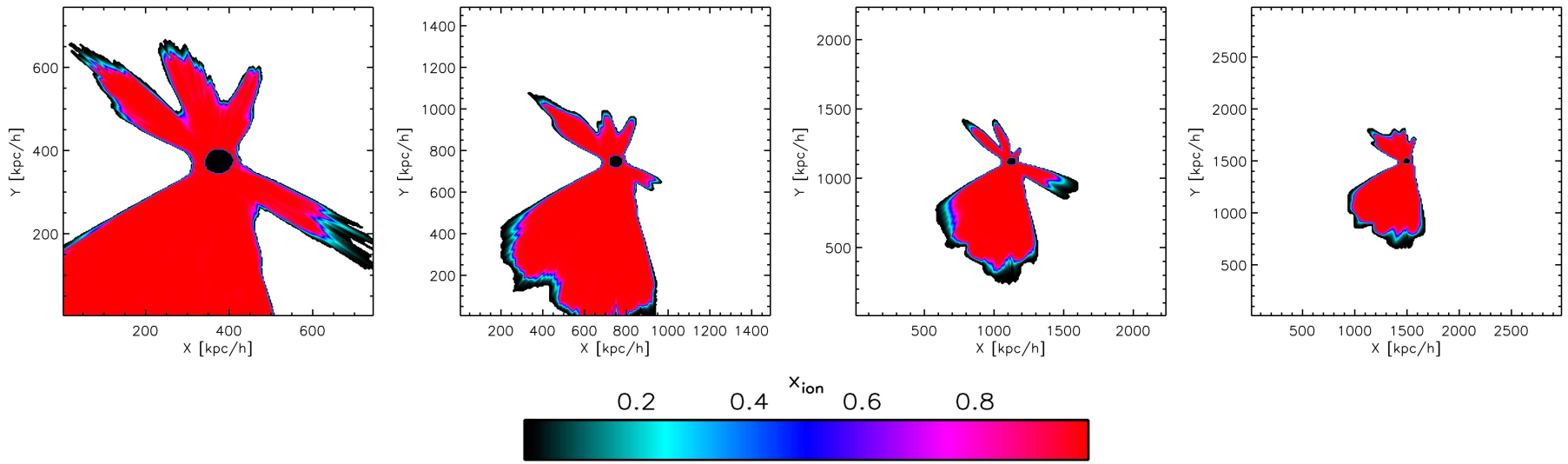

Figure B.1: A slice through the ionization structure in the cube for simulations with different cube sizes of the IGM around the source. The degree of ionization of the gas is shown by the color bar at the bottom. The panels are for simulations of different cube sizes - $\sim 12\ r_{200}$ (left panel), $\sim 23\ r_{200}$ (centre left panel), $\sim 35\ r_{200}$ (centre right panel), $\sim 47\ r_{200}$ (right panel) where r_{200} is the radius in comoving units at which the mean DM density inside the sphere is 200 times the critical density of the universe at that redshift.

Bibliography

Abel T., Bryan G. L., Norman M. L., 2002, Sci, 295, 93

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Physical Review, 73, 803

Alpher R. A., Herman R. C., 1948, PhRv, 74, 1737

Altay G., Croft R. A. C., Pelupessy I., 2008, MNRAS, 386, 1931

Alvarez M. A., Busha M., Abel T., Wechsler R. H., 2009, ApJ, 703, L167

Ambarzumian V. A., 1932, MNRAS, 93, 50

Astier P., et al., 2006, A&A, 447, 31

Atek H., Kunth D., Hayes M., Östlin G., Mas-Hesse J. M., 2008, A&A, 488, 491

Aubert D., Teyssier R., 2008, MNRAS, 387, 295

Barkana R., Loeb A., 2001, PhR, 349, 125

Barkana R., Loeb A., 2004, ApJ, 601, 64

Barkana R., Loeb A., 2007, RPPh, 70, 627

Barnes L. A., Haehnelt M. G., Tescari E., Viel M., 2011, MNRAS, 416, 1723

Becker R. H., et al., 2001, AJ, 122, 2850

Blanc G. A., et al., 2011, ApJ, 736, 31

Bland-Hawthorn J., Putman M. E., 2001, ASPC, 240, 369

Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325

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., et al., 2011, in prep.

Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2008, ApJ, 686, 230

Bromm V., Kudritzki R. P., Loeb A., 2001, ApJ, 552, 464

Bromm V., Coppi P. S., Larson R. B., 2002, ApJ, 564, 23

Bromm V., Loeb A., 2003, Nature, 425, 812

Bromm V., Yoshida N., 2011, ARA&A, 49, 373

Cantalupo S., Porciani C., Lilly S. J., Miniati F., 2005, ApJ, 628, 61

Cassata P., et al., 2011, A&A, 525, A143

Ciardi B., Ferrara A., Marri S., Raimondo G., 2001, MNRAS, 324, 381

Ciardi B., Bianchi S., Ferrara A., 2002, MNRAS, 331, 463

Ciardi B., Stoehr F., White S. D. M., 2003, MNRAS, 343, 1101

Ciardi B., Ferrara A., 2005, SSRv, 116, 625

Ciardi B., et al., 2012, in prep.

Charlot S., Fall S. M., 1993, ApJ, 415, 580

Choudhury T. R., Ferrara A., 2006, MNRAS, 371, L55

Choudhury T. R., Haehnelt M. G., Regan J., 2009, MNRAS, 394, 960

Clarke C., Oey M. S., 2002, MNRAS, 337, 1299

Clark P. C., Glover S. C. O., Klessen R. S., Bromm V., 2011, ApJ, 727, 110

Clément B., et al., 2011, arXiv, arXiv:1105.4235

Cowie L. L., Hu E. M., 1998, AJ, 115, 1319

Cowie L. L., Barger A. J., Hu E. M., 2010, ApJ, 711, 928

Cuby J.-G., Hibon P., Lidman C., Le Fèvre O., Gilmozzi R., Moorwood A., van der Werf P., 2007, A&A, 461, 911

Cucchiara A., et al., 2011, ApJ, 736, 7

Davé R., Finlator K., Oppenheimer B. D., 2006, MNRAS, 370, 273

Davis M., Wilkinson D. T., 1974, ApJ, 192, 251

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dawson S., Malhotra S., Rhoads J. E., Stern D., Dey A., Spinrad H., Jannuzi B. T., 2004, AAS, 36, 746

Dawson S., Rhoads J. E., Malhotra S., Stern D., Wang J., Dey A., Spinrad H., Jannuzi B. T., 2007, ApJ, 671, 1227

Dayal P., Ferrara A., Gallerani S., 2008, MNRAS, 389, 1683

Dayal P., Ferrara A., Saro A., Salvaterra R., Borgani S., Tornatore L., 2009, MNRAS, 400, 2000

Deharveng J.-M., et al., 2008, ApJ, 680, 1072

Dijkstra M., Haiman Z., Spaans M., 2006, ApJ, 649, 14

Dijkstra M., Lidz A., Wyithe J. S. B., 2007, MNRAS, 377, 1175

Dijkstra M., Wyithe J. S. B., Haiman Z., 2007, MNRAS, 379, 253

Dijkstra M., Wyithe J. S. B., 2007, MNRAS, 379, 1589

Dijkstra M., Wyithe J. S. B., 2010, MNRAS, 408, 35

Dijkstra M., Mesinger A., Wyithe J. S. B., 2011, MNRAS, 414, 2139

Djorgovski S., Thompson D. J., 1992, IAUS, 149, 337

Dodelson S., 2003, moco.book,

Dopita M. A., Krauss L. M., Sutherland R. S., Kobayashi C., Lineweaver C. H., 2011, Ap&SS, 335, 345

Dove J. B., Shull J. M., 1994, ApJ, 430, 222

Dove J. B., Shull J. M., Ferrara A., 2000, ApJ, 531, 846

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Fan X., Narayanan V. K., Strauss M. A., White R. L., Becker R. H., Pentericci L., Rix H.-W., 2002, AJ, 123, 1247

Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415

Fan X., et al., 2006, AJ, 132, 117

Fardal M. A., Katz N., Gardner J. P., Hernquist L., Weinberg D. H., Davé R., 2001, ApJ, 562, 605

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

Fernandez E. R., Shull J. M., 2011, ApJ, 731, 20

Finkelstein S. L., Rhoads J. E., Malhotra S., Pirzkal N., Wang J., 2007, ApJ, 660, 1023

Finkelstein S. L., et al., 2011, arXiv, arXiv:1110.3785

Finlator K., Özel F., Davé R., Oppenheimer B. D., 2009, MNRAS, 400, 1049

Fontana A., et al., 2010, ApJ, 725, L205

Forero-Romero J. E., Yepes G., Gottlöber S., Knollmann S. R., Cuesta A. J., Prada F., 2011, MNRAS, 415, 3666

Freedman W. L., et al., 2001, ApJ, 553, 47

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385

Fujita A., Martin C. L., Mac Low M.-M., Abel T., 2003, ApJ, 599, 50

Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2011, arXiv:1111.2334

Furlanetto S. R., Loeb A., 2004, ApJ, 611, 642

Furlanetto S. R., Oh S. P., Briggs F. H., 2006, PhR, 433, 181

Gamow G., 1948, Natur, 162, 680

Gawiser E., et al., 2006, ApJ, 642, L13

Gawiser E., et al., 2007, ApJ, 671, 278

Geil P. M., Wyithe J. S. B., 2008, MNRAS, 386, 1683

Giallongo E., Cristiani S., D'Odorico S., Fontana A., 2002, ApJ, 568, L9

Giroux M. L., Shapiro P. R., 1996, ApJS, 102, 191

Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581

Gnedin N. Y., 2000, ApJ, 535, 530

Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, ApJ, 672, 765

Greif T. H., Springel V., White S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011, ApJ, 737, 75

Gronwall C., et al., 2007, ApJ, 667, 79

Gronwall C., Bond N. A., Ciardullo R., Gawiser E., Altmann M., Blanc G. A., Feldmeier J. J., 2010, arXiv, arXiv:1005.3006

Guaita L., et al., 2010, ApJ, 714, 255

Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633

Guth A. H., 1981, PhRvD, 23, 347

Haardt F., Madau P., 2001, cghr.conf,

Haiman Z., Spaans M., 1999, ApJ, 518, 138

Haiman Z., Spaans M., Quataert E., 2000, ApJ, 537, L5

Haiman Z., Rees M. J., 2001, ApJ, 556, 87

Haiman Z., 2002, ApJ, 576, L1

Haiman Z., Cen R., 2005, ApJ, 623, 627

Hansen S. H., Haiman Z., 2004, ApJ, 600, 26

Hartmann L. W., Huchra J. P., Geller M. J., O'Brien P., Wilson R., 1988, ApJ, 326, 101

Hayashino T., et al., 2004, AJ, 128, 2073

Hayes M., Östlin G., Atek H., Kunth D., Mas-Hesse J. M., Leitherer C., Jiménez-Bailón E., Adamo A., 2007, MNRAS, 382, 1465

Hayes M., et al., 2010, Natur, 464, 562

Hayes M., Schaerer D., Östlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, ApJ, 730, 8

Henyey L. G., 1940, PNAS, 26, 50

Hill G. J., et al., 2008, ASPC, 399, 115

Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731

Hu E. M., McMahon R. G., 1996, Natur, 382, 231

Hu E. M., Cowie L. L., McMahon R. G., 1998, ApJ, 502, L99

Hu E. M., Cowie L. L., McMahon R. G., Capak P., Iwamuro F., Kneib J.-P., Maihara T., Motohara K., 2002, ApJ, 568, L75

Hu E. M., Cowie L. L., Capak P., McMahon R. G., Hayashino T., Komiyama Y., 2004, AJ, 127, 563

Hu E. M., Cowie L. L., Capak P., Kakazu Y., 2005, pgqa.conf, 363

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Iliev I. T., Scannapieco E., Shapiro P. R., 2005, ApJ, 624, 491

Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625

Iliev I. T., et al., 2006, MNRAS, 371, 1057

Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007, MNRAS, 376, 534

Iliev I. T., Shapiro P. R., Mellema G., Merz H., Pen U.-L., 2008, tera.conf, 31

Iliev I. T., Shapiro P. R., McDonald P., Mellema G., Pen U.-L., 2008, MNRAS, 391, 63

Iliev I. T., et al., 2009, MNRAS, 400, 1283

Inoue A. K., Iwata I., Deharveng J.-M., Buat V., Burgarella D., 2005, A&A, 435, 471

Iwata I., et al., 2009, ApJ, 692, 1287

Iye M., et al., 2006, Natur, 443, 186

Jakobsen P., Boksenberg A., Deharveng J. M., Greenfield P., Jedrzejewski R., Paresce F., 1994, Natur, 370, 35

Jarosik N., et al., 2011, ApJS, 192, 14

Kashikawa N., et al., 2006, ApJ, 648, 7

Kashikawa N., et al., 2011, ApJ, 734, 119

Kobayashi M. A. R., Kamaya H., Yonehara A., 2006, ApJ, 636, 1

Kobayashi M. A. R., Totani T., Nagashima M., 2007, ApJ, 670, 919

Kohler K., Gnedin N. Y., Miralda-Escudé J., Shaver P. A., 2005, ApJ, 633, 552

Kohler K., Gnedin N. Y., Hamilton A. J. S., 2007, ApJ, 657, 15

Kohler K., Gnedin N. Y., 2007, ApJ, 655, 685

Kollmeier J. A., Zheng Z., Davé R., Gould A., Katz N., Miralda-Escudé J., Weinberg D. H., 2010, ApJ, 708, 1048

Komatsu E., et al., 2011, ApJS, 192, 18

Kovač K., Somerville R. S., Rhoads J. E., Malhotra S., Wang J., 2007, ApJ, 668, 15

Lacey C., Cole S., 1994, MNRAS, 271, 676

Lai K., Huang J.-S., Fazio G., Cowie L. L., Hu E. M., Kakazu Y., 2007, ApJ, 655, 704

Larson D., et al., 2011, ApJS, 192, 16

Laursen P., Sommer-Larsen J., 2007, ApJ, 657, L69

Laursen P., Sommer-Larsen J., Andersen A. C., 2009, ApJ, 704, 1640

Laursen P., 2010, arXiv, arXiv:1012.3175

Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, ApJ, 728, 52

Le Delliou M., Lacey C., Baugh C. M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S. L., 2005, MNRAS, 357, L11

Le Delliou M., Lacey C. G., Baugh C. M., Morris S. L., 2006, MNRAS, 365, 712

Lehnert M. D., et al., 2010, Natur, 467, 940

Leitherer C., et al., 1999, ApJS, 123, 3

Loeb A., Rybicki G. B., 1999, ApJ, 524, 527

Loeb A., 2009, JCAP, 3, 22

Lowenthal J. D., Hogan C. J., Green R. F., Caulet A., Woodgate B. E., Brown L., Foltz C. B., 1991, ApJ, 377, L73

Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648

Madau P., Rees M. J., Volonteri M., Haardt F., Oh S. P., 2004, ApJ, 604, 484

Maio U., Dolag K., Ciardi B., Tornatore L., 2007, MNRAS, 379, 963

Maio U., Ciardi B., Yoshida N., Dolag K., Tornatore L., 2009, A&A, 503, 25

Maio U., Ciardi B., Dolag K., Tornatore L., Khochfar S., 2010, MNRAS, 407, 1003

Malhotra S., Rhoads J. E., 2002, ApJ, 565, L71

Malhotra S., Rhoads J. E., 2004, ApJ, 617, L5

Malhotra S., Rhoads J. E., 2006, ApJ, 647, L95

Maselli A., Ferrara A., Ciardi B., 2003, MNRAS, 345, 379

Maselli A., Ciardi B., Kanekar A., 2009, MNRAS, 393, 171

Matsuda Y., et al., 2004, AJ, 128, 569

McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043

Meier D. L., 1976, ApJ, 207, 343

Meier D. L., Terlevich R., 1981, ApJ, 246, L109

Meiksin A., 2005, MNRAS, 356, 596

Meiksin A. A., 2009, RvMP, 81, 1405

Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006, NewA, 11, 374

Mesinger A., Furlanetto S., 2007, ApJ, 669, 663

Mesinger A., Furlanetto S. R., 2008, MNRAS, 386, 1990

Mesinger A., 2010, MNRAS, 407, 1328

Miller A. D., et al., 1999, ApJ, 524, L1

Miniati F., Ferrara A., White S. D. M., Bianchi S., 2004, MNRAS, 348, 964

Miralda-Escude J., Rees M. J., 1998, ApJ, 497, 21

Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1

Mo H., van den Bosch F. C., White S., 2010, gfe..book,

Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127

Mortlock D. J., et al., 2011, Natur, 474, 616

Nagamine K., Ouchi M., Springel V., Hernquist L., 2010, PASJ, 62, 1455

Neufeld D. A., 1990, ApJ, 350, 216

Neufeld D. A., 1991, ApJ, 370, L85

Nilsson K. K., Fynbo J. P. U., Møller P., Sommer-Larsen J., Ledoux C., 2006, A&A, 452,

Nilsson K. K., et al., 2007, A&A, 471, 71 L23

Nilsson K. K., Östlin G., Møller P., Möller-Nilsson O., Tapken C., Freudling W., Fynbo J. P. U., 2011, A&A, 529, A9

Ono Y., et al., 2011, arXiv, arXiv:1107.3159

Osterbrock D. E., 1962, ApJ, 135, 195

Osterbrock D. E., 1989, agna.book,

Östlin G., Hayes M., Kunth D., Mas-Hesse J. M., Leitherer C., Petrosian A., Atek H., 2009, AJ, 138, 923

Ostriker J. P., Vishniac E. T., 1986, ApJ, 306, L51

Ouchi M., 2008, ASPC, 399, 46

Ouchi M., et al., 2008, ApJS, 176, 301

Ouchi M., et al., 2009, ApJ, 696, 1164

Ouchi M., et al., 2010, ApJ, 723, 869

Paardekooper J.-P., Pelupessy F. I., Altay G., Kruip C. J. H., 2011, A&A, 530, A87

Padmanabhan T., 1993, sfu..book,

Pakmor R., 2010, PhD thesis, TU Munich.

Partl A. M., Maselli A., Ciardi B., Ferrara A., Müller V., 2011, MNRAS, 414, 428

Partridge R. B., 1974, ApJ, 192, 241

Partridge R. B., Peebles P. J. E., 1967, ApJ, 147, 868

Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, MNRAS, 394, 1812

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

Peacock J. A., 1999, coph.book,

Peebles P. J. E., 1993, ppc..book,

Pentericci L., et al., 2011, ApJ, 743, 132

Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, MNRAS, 381, 1053

Perlmutter S., et al., 1999, ApJ, 517, 565

Petkova M., Springel V., 2009, MNRAS, 396, 1383

Petkova M., Springel V., 2011, MNRAS, 412, 935

Pierleoni M., Maselli A., Ciardi B., 2009, MNRAS, 393, 872

Pierleoni M.et al., 2012, in prep.

Pirzkal N., Malhotra S., Rhoads J. E., Xu C., 2007, ApJ, 667, 49

Pritchard J. R., Loeb A., Wyithe J. S. B., 2010, MNRAS, 408, 57

Pritchard J. R., Loeb A., 2011, arXiv, arXiv:1109.6012

Pritchet C., Hartwick D., 1989, JRASC, 83, 318

Rai\vcević M., Theuns T., 2010, arXiv, arXiv:1011.3823

Razoumov A. O., Norman M. L., Abel T., Scott D., 2002, ApJ, 572, 695

Razoumov A. O., Sommer-Larsen J., 2006, ApJ, 651, L89

Razoumov A. O., Sommer-Larsen J., 2010, ApJ, 710, 1239

Rhee G. F. R. N., Webb J. K., Katgert P., 1989, A&A, 217, 1

Rhoads J. E., Malhotra S., 2001, ApJ, 563, L5

Rhoads J. E., et al., 2003, AJ, 125, 1006

Rhoads J. E., et al., 2004, ApJ, 611, 59

Ricotti M., Shull J. M., 2000, ApJ, 542, 548

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2004, ApJ, 607, 665

Santos M. G., Cooray A., Haiman Z., Knox L., Ma C.-P., 2003, ApJ, 598, 756

Santos M. R., 2004, MNRAS, 349, 1137

Samui S., Srianand R., Subramanian K., 2009, MNRAS, 398, 2061

Schaerer D., 2003, A&A, 397, 527

Schaerer D., Hayes M., Verhamme A., Teyssier R., 2011, A&A, 531, A12

Schenker M. A., Stark D. P, Ellis R. S., Robertson B. E., Dunlop J. S., McLure R. J., Kneib J. -., Richard J., 2011, arXiv, arXiv:1107.1261

Schneider R., Ferrara A., Salvaterra R., Omukai K., Bromm V., 2003, Natur, 422, 869

Schneider R., Omukai K., Inoue A. K., Ferrara A., 2006, MNRAS, 369, 1437

Semelin B., Combes F., Baek S., 2007, A&A, 474, 365

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., Erb D. K., 2006, ApJ, 651, 688

Shapley A. E., 2011, ARA&A, 49, 525

Shibuya T., Kashikawa N., Ota K., Iye M., Ouchi M., Furusawa H., Shimasaku K., Hattori T., 2011, arXiv, arXiv:1112.3997

Shimasaku K., et al., 2006, PASJ, 58, 313

Shin M.-S., Trac H., Cen R., 2008, ApJ, 681, 756

Siana B., et al., 2010, ApJ, 723, 241

Sokasian A., Abel T., Hernquist L. E., 2001, NewA, 6, 359

Songaila A., Cowie L. L., 2010, ApJ, 721, 1448

Spitzer L., 1978, ppim.book,

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., 2005, MNRAS, 364, 1105

Stark D. P., Ellis R. S., Richard J., Kneib J.-P., Smith G. P., Santos M. R., 2007, ApJ, 663, 10

Stark D. P., Ellis R. S., Ouchi M., 2011, ApJ, 728, L2

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2000, ApJ, 532, 170

Steidel C. C., Pettini M., Adelberger K. L., 2001, ApJ, 546, 665

Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, arXiv, arXiv:1101.2204

Stern D., Yost S. A., Eckart M. E., Harrison F. A., Helfand D. J., Djorgovski S. G., Malhotra S., Rhoads J. E., 2005, ApJ, 619, 12

Sunyaev R. A., Chluba J., 2008, ASPC, 395, 35

Sunyaev R. A., Zeldovich I. B., 1980, ARA&A, 18, 537

Taniguchi Y., et al., 2005, PASJ, 57, 165

Tasitsiomi A., 2006, ApJ, 645, 792

Telfer R. C., Zheng W., Kriss G. A., Davidsen A. F., 2002, ApJ, 565, 773

Thomas R. M., et al., 2009, MNRAS, 393, 32

Tilvi V., Malhotra S., Rhoads J. E., Scannapieco E., Thacker R. J., Iliev I. T., Mellema G., 2009, ApJ, 704, 724

Tilvi V., et al., 2010, ApJ, 721, 1853

Tolman R. C., 1934, rtc..book,

Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050

Totani T., Kawai N., Kosugi G., Aoki K., Yamada T., Iye M., Ohta K., Hattori T., 2006, PASJ, 58, 485

Trac H., Cen R., 2007, ApJ, 671, 1

Trac H., Cen R., 2008, AIPC, 990, 445

Trac H., Gnedin N. Y., 2009, arXiv, arXiv:0906.4348

Tumlinson J., Shull J. M., 2000, ApJ, 528, L65

Unno W., 1952, PASJ, 4, 100

Valls-Gabaud D., 1993, ApJ, 419, 7

van de Voort F., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 2782

Vanzella E., et al., 2011, ApJ, 730, L35

Verhamme A., Schaerer D., Atek H., Tapken C., 2007, ASPC, 380, 97

Verhamme A., Schaerer D., Maselli A., 2006, A&A, 460, 397

Wang J.-X., Malhotra S., Rhoads J. E., Zhang H.-T., Finkelstein S. L., 2009, ApJ, 706, 762

Weidinger M., Møller P., Fynbo J. P. U., Thomsen B., 2005, A&A, 436, 825

Wise J. H., Cen R., 2009, ApJ, 693, 984

Wolf C., Wisotzki L., Borch A., Dye S., Kleinheinrich M., Meisenheimer K., 2003, A&A, 408, 499

Wolf C., et al., 2010, AAS, 42, #410.10

Wood K., Loeb A., 2000, ApJ, 545, 86

Wyithe J. S. B., Loeb A., 2003, ApJ, 586, 693

Wyithe J. S. B., Loeb A., 2005, ApJ, 625, 1

Wyithe J. S. B., Cen R., 2007, ApJ, 659, 890

Wyithe J. S. B., Mould J., Loeb A., 2011, ApJ, 743, 173

Yajima H., Umemura M., Mori M., Nakamoto T., 2009, MNRAS, 398, 715

Yajima H., Choi J.-H., Nagamine K., 2011, MNRAS, 412, 411

Yajima H., Li Y., Zhu Q., Abel T., 2011, arXiv, arXiv:1109.4891

Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003, ApJ, 592, 645

Zahn O., Zaldarriaga M., Hernquist L., McQuinn M., 2005, ApJ, 630, 657

Zahn O., Mesinger A., McQuinn M., Trac H., Cen R., Hernquist L. E., 2011, MNRAS, 414, 727

Zanstra H., 1949, BAN, 11, 1

Zhang Y., Meiksin A., Anninos P., Norman M. L., 1998, ApJ, 495, 63

Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2010, ApJ, 716, 574

Zheng Z., Cen R., Weinberg D., Trac H., Miralda-Escude J., 2010, arXiv, arXiv:1010.3017

Zwicky F., 1937, ApJ, 86, 217

Acknowledgments

A lot of people have played important roles throughout my PhD, making it a wonderful learning experience. Firstly, I would like to thank MPA and IMPRS for giving me this excellent opportunity and helping me achieve it. My sincere gratefulness to Benedetta for being my perfect supervisor - all that I needed her to be at the right times, giving me the space I needed while keeping me on track. I would also like to thank Simon White, Martin Asplund and Henk Spruit for being my thesis committee members, for all their timely advice and support. Mark, with a lot of interesting discussions, has helped me understand and appreciate the Lyα line a lot better. Many thanks to Umberto, Marco, Luca, Jamie and Antonella for providing me with the codes and simulations without which the projects in this thesis would not have been possible. Also thanks to everyone I have met and had discussions with at meetings/summer school/conferences for making them a very productive, stimulating and enjoyable experience and thanks to MPA for permitting me to attend them. Last but not the least, thanks to the support staff at MPA/MPE for making my moving to a new country so easy.

On the personal front, this thesis is dedicated to my family - Appa, Amma, Pappu and Velliamma - this journey which started long ago would not have been possible without their love, inspiration and support. My amazing friends - Mansura, Shiuli, Cibin, Stephanie, Janina, Rajyalakshmi, Wenting, Francesca, Laura - thanks for being there and keeping me sane. To all the wonderful people I have met in Munich - thanks for making this a brilliant period of my life.

Finally, thanks to Astronomy...