Simulations of cosmic reionization

MILAN RAICEVIC

Simulations of cosmic reionization

Milan Raičević

Abstract

In this thesis we investigate numerically how ionizing photons emitted by stars in galaxies cause the reionization of the Universe, the transition during which most of the gas in the Universe from a mostly neutral, to a highly ionised state it is in today. To this end, we discuss and improve two techniques for the transport of ionising radiation across cosmological volumes, analyse the sources of ionising photons at high redshifts predicted by a semi-analytical galaxy formation model (GALFORM), and combine these to make consistent model of how reionization proceeds.

Our improvements to the hybrid characteristics (HC) radiative transport scheme are significant, making the code faster and more accurate, as demonstrated by our contribution to a code comparison paper (Iliev et al., 2009). Our improvements to the SimpleX radiative transport scheme allow for accurate and significantly better numerically converged calculations of the speeds of ionization fronts of cosmological HII regions. This is accomplished by a much more thorough analysis of how to properly model the density field on the unstructured density field in SimpleX.

The dependence of the ionizing emissivity of GALFORM galaxies on various parameters of the model is examined. We show that massive stars formed in abundance because of the assumed top-heavy stellar initial mass function during starbursts in the Baugh et al. (2005) model, triggered by galaxy mergers, are the dominant source of ionizing photons. We show that the luminosity functions predicted by this model are in good agreement with the most recent *Hubble Space Telescope* results at $z \gtrsim 8$. The model also demonstrates that most photons are produced in faint galaxies which are not yet seen in the current data.

We then combine the sources predicted by GALFORM with the SimpleX RT scheme to model inhomogeneous reionization including the effects of source suppression. We investigate how the morphology of reionization depends on the model for the sources, which may be crucial for future observations of this cosmic epoch.

Simulations of cosmic reionization

by Milan Raičević

A thesis submitted to the University of Durham
in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy.

Department of Physics
University of Durham
March 2010

Contents

List of Figures

List of Tables

Declaration

The work described in this thesis was undertaken between 2006 and 2010 while the author was a research student under the supervision of Dr. Tom Theuns in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or any other University.

The simulations from Chapter 2 have been used in the following paper:

- Iliev I.T., Whalen D., Mellema G., Ahn K., Baek S., Gnedin N.Y., Kravtsov A.V., Norman M., Raicevic M., Reynolds D.R., Sato D., Shapiro P.R., Semelin B., Smidt J., Susa H., Theuns T., Umemura M., "Cosmological Radiative Transfer Comparison Project II: The Radiation-Hydrodynamic Tests",Monthly Notices Of Royal Astronomical Society, 400, 1283

All figures have been produced by the author, except Figures 2.17, 2.22 and 2.23, produced by Ilian .T. Iliev et al.

Acknowledgements

This thesis is dedicated firstly to my supervisor, Tom Theuns, without whose help, support and endless patience in the face of an endless stream of my questions, none of this would have been possible. I can only hope for the opportunity to pass on everything I have learned with the same energy and enthusiasm with which Tom taught me. I am also hugely indebted to my co-supervisors: Richard Bower, who helped me get the gas flowing on AMR grids, and Cedric Lacey, who bravely took on the task of explaining to me how galaxies form. Both of them have been a constant source of support and encouragement I could thoroughly rely on. A special thank you goes to Lydia Heck, whose masterful command of all things HPC made this work possible.

A single page is hardly enough to mention all the people who have made my stay in Durham such an enjoyable experience. I was very fortunate spend most of my time here sharing an office with Raul Angulo, Phil Bett and Greg Davis, who instantly made me a part of the team. So many people were here to brighten up the usually rainy Durham days and share in various world travels: Yanchuan Cai, Andrew Cooper, Rob Crain, Peter Creasey, Nikolaos Fanidakis, Elise Jennings, Nigel Mitchell, Dave Murphey, Alvaro Orsi, Tim Rawle and anyone else that slipped my mind (I assure you, it was not intentional). For those rare moments away from the ICC, I was fortunate enough to hang out with Richard Stopford and the rest of the gang from the Animal House that was Old Belvedere. Out of the many wonderful people I had the opportunity to meet on the conference circuit, I must single out Jan-Pieter Paardekooper and Dan Whalen, whose company always makes going on the road even more fun.

A big thank you goes out to my family, especially my mother Smiljana, whose superhuman efforts to keep her mothering to a minimum I gratefully acknowledge. Finally, to my better nine tenths, Katarina Miljković, thank you for always having a ray of sunshine in your back pocket just for me.

Honourable mention goes to Red Bull, for giving me wings.

The work on this thesis has been supported by a grant from Microsoft Research Cambridge.

Chapter 1 | *Introduction*

The 20th century saw the transformation of cosmology from a physicist's hobby to a precision science with a well defined and predictive paradigm. This model is founded on two main ideas: i) the Universe expanded a finite time ago from an infinitesimally small size in an event dubbed the Big Bang (BB) and continues to expand to the present day and, ii) all the structure we observe in the Universe has grown from very small density inhomogeneities, seeded by inflation, under the influence of gravity. This picture is supported by 3 observational pillars: the expansion of the Universe, the BB nucleosynthesis and the presence of the cosmic microwave background (CMB).

This thesis focuses on one of the last unseen frontiers in the study of the Universe: the epoch of reionization (EoR). This is the period during which the radiation emitted from the first luminous objects in the Universe transformed the previously cold and neutral intergalactic hydrogen into the hot, highly ionized state in which we observe it today. Very little about this process is known, beyond that it had to happen. When did the reionization occur? How fast did it progress? What are the radiation sources responsible for it? These are some of the questions this thesis aims to address.

In the last decade, the theoretical study of reionization was spurred on by the promise of the direct observations of this epoch to come from a next generation of telescopes. Future data from radio interferometers such as LOFAR[1], MWA[2] and SKA[3] will be able to image the process by mapping the distribution of neutral hydrogen as a function of redshift. Furthermore, the next generation space telescope JWST[4] and ground based telescopes such as ELT[5] will be capable of observing galaxies at the time that the Universe was only a few hundred million years old, thus providing a direct view of possible sources ionizing the Universe.

[1] http://www.lofar.org
[2] http://www.haystack.mit.edu/ast/arrays/mwa/
[3] http://skatelescope.org/
[4] http://jwst.nasa.gov/
[5] http://www.eso.org/facilities/eelt/

For now, the study of reionization is largely limited to theory. The most promising tools in this field are numerical simulations that track radiative transfer of ionizing radiation over the cosmological density fields. A major part of this thesis is the improvement and application of a number of such radiative transfer methods.

The organisation of this chapter is as follows. In Section 1.1, we present a short overview of the current cosmological paradigm, from the initial density perturbations to the formation of first stars and galaxies. In Section 1.2, we discuss the current state of reionization studies, including the observational constraints and some of the numerical techniques being employed. In Section 1.3, we introduce the basic quantities in the theory of radiative transfer. Finally, Section 1.4 provides a summary of the topics and results presented in this thesis.

1.1 Cosmology and structure formation

In this section, we briefly review the concordance model for the evolution of the Universe and structure within it which provides the backdrop for the discussion of reionization. We do not intend to present an exhaustive treatment, instead focusing on only the key physical elements needed for the following discussions. There are many textbooks that provide an in-depth look into this subject (e.g. Peebles, 1993; Peacock, 1999; Coles & Lucchin, 2002), so the following overview is intentionally very brief.

1.1.1 The homogeneous Universe

Modern cosmology is based on two key theoretical and observational advances. One is the development of the General Theory of Relativity by Einstein (1916), which describes gravity in terms of the geometric properties of space-time intimately tied to the distribution of mass. The second is provided by Hubble & Humason (1931), who found that the other galaxies in the Universe recede from the Milky Way at velocities, v, proportional to the their distance, r. This is expressed in the form of the famous Hubble Law:

$$v = Hr, \tag{1.1}$$

where the proportionality constant is the *Hubble parameter*, with the units of velocity per unit distance. Common practise is to introduce the dimensionless Hubble parameter, h, defined to give $H = h \times 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The Hubble Law provided the first

observational confirmation of the *Cosmological principle*, which postulates that the Universe, on large enough scales, is both isotropic (i.e. looks the same in any direction) and homogeneous (i.e. looks the same from any location). The assumption underlying this principle is motivated primarily by philosophical arguments. The physical explanation, however, is provided by the theory of *inflation* (Guth, 1981; Linde, 1990), which gives concrete foundations to the Big Bang cosmology by tying it with quantum theory. The main assumption of inflation is that the Universe, when only $10^{-43}\mathrm{s} \lesssim t \lesssim 10^{-30}\mathrm{s}$ old, went through a stage of exponential expansion. This rapid expansion results in the apparent homogeneity of the Universe (the horizon problem in classic BB theory) and explains why the Universe appears to be nearly perfectly flat (resolving the flatness problem). It is also expected to blow up quantum-scale fluctuations from the earliest moments of the Universe to cosmological scales, which seed the tiny ripples seen in the cosmic microwave background (CMB). The theory of inflation provides remarkably accurate predictions for the observed CMB anisotropies, making it a key element of the Big Bang model.

The main tool for describing the evolution of the Universe as a whole was provided by Friedman (1922), whose solution of the general relativity field equations gives the relation for the evolution of the Hubble parameter in the Universe:

$$H^2(a) \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \tag{1.2}$$

where we have equated the Hubble parameter evolution with that of the *scale factor* $a(t)$. The scale factor describes how the physical distance between objects in the Universe depends on time and is usually normalised to unity at the present day. The Friedman equation, Eq. (1.2), therefore states that the relative scale of the Universe at any time depends on its contents, ρ being the total energy density and k, the geometric curvature. The critical density that yields a flat Universe ($k = 0$, obeying Euclidean geometry) is:

$$\rho_c(a) = \frac{3H^2(a)}{8\pi G}. \tag{1.3}$$

The relation of the actual mean density to the critical one defined in Eq. (1.3) determines the ultimate fate of the Universe (unless there is dark energy). In that case, when the mean density is super-critical the Universe is spatially finite ($k > 0$) and the gravitational force is sufficient to arrest its expansion potentially causing it to recollapse. Conversely,

the sub-critical Universe is spatially infinite ($k < 0$) and destined to expand ever faster indefinitely. In between the two, a flat Universe ($k = 0$, $\rho = \rho_c$) is the limiting case in which the expansion continues at all times, but at a constantly decreasing rate.

The energy density ρ in Eq. (1.2) consists of three components: $\rho = \rho_m + \rho_r + \rho_\Lambda$, where ρ_m is the matter, ρ_r radiation and ρ_Λ dark energy density. All of these components evolve differently with time: the matter density as $\rho_m(a) \propto a^{-3}$, caused by the increasing volume of the Universe, while radiation density evolves even faster, $\rho_r \propto a^{-4}$ due to both the increase in volume and redshifting of the emitted photon wavelength due to the expansion. In the ΛCDM paradigm, where CDM stands for cold dark matter (see below), the dark energy is assumed to be constant in time: $\rho_\Lambda \equiv \Lambda c^2/(8\pi G)$, where Λ is the so-called cosmological constant. Note that more general dark energy models are also possible.

The relative contribution of each component in Eq. (1.2) to the evolution of the Universe is commonly expressed as the ratio of the component density to the critical density at the present day, thus defining a set of cosmological variables dubbed the *density parameters*:

$$\Omega_{\mathrm{m},0} = \frac{\rho_{\mathrm{m},0}}{\rho_{\mathrm{c},0}}; \Omega_{\mathrm{r},0} = \frac{\rho_{\mathrm{r},0}}{\rho_{\mathrm{c},0}}; \Omega_{\Lambda,0} = \frac{\rho_{\Lambda,0}}{\rho_{\mathrm{c},0}}, \tag{1.4}$$

where index 0 denotes the value at the present day, when $a = a_0 = 1$ as mentioned previously. Dropping the said index for the sake of brevity, we can rewrite the Eq. (1.2) as:

$$H^2(a) = H_0^2(\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_\Lambda), \tag{1.5}$$

which states how the Hubble parameter (i.e. the scale size of the Universe) relates to the present day energy density content. Note that we have assumed that the Universe is flat, $k = 0$, which is what is suggested by the latest observations and predicted by inflation theory. This implies that $\Omega_m + \Omega_r + \Omega_\Lambda = 1$.

The matter component accounts for about 26% of the total energy density at the present epoch ($\Omega_\mathrm{m} \approx 0.26$). About 83% of all matter is thought to be in the form of some cold and dark, weakly interactive non-baryonic particles, the exact properties of which are still unknown and are commonly referred to as *dark matter*. The existence of dark matter was first inferred by Zwicky (1933) (who also coined the name) when he found that the dynamics of galaxies in massive clusters cannot be explained by their visible components alone. Further studies have shown (see e.g. Sofue & Rubin, 2001, for

a review) that the rotation curves of practically all galaxies do not exhibit a Keplerian $1/r^2$ drop-off that is expected when the mass within radius r is constant. The lack of this drop-off well beyond the radius of the galaxy suggests the presence of some unseen gravitating material. The existence of dark matter is also strongly suggested by CMB observations.

The remaining 17% of the total matter mass is the "ordinary" matter, baryons and fermions that are found in the atoms which form everything from largest galaxies to all life on Earth. This is the only matter component we can directly detect, as baryons both emit and interact with radiation across the electromagnetic spectrum. The primordial composition of the baryonic component is well predicted by the Big Bang nucleosynthesis theory (e.g. Alpher et al., 1948; Gamow, 1948): about 75% of mass consisted of hydrogen and about 25% was helium. Trace amounts of deuterium, lithium and beryllium were also present.

The majority of the energy density at present epoch is due to the cosmological constant, $\Omega_\Lambda \approx 0.74$. Even though Einstein introduced it to preserve the now discarded static Universe model, it was later reintroduced to reconcile the CMB and large-scale structure measurements (e.g. Efstathiou et al., 1990) and, later, the acceleration of the expansion at lower redshift found by Type Ia supernovae observations at $z \lesssim 1$ (Riess et al., 1998; Perlmutter et al., 1999). The exact nature of this dark or (vacuum) energy remains highly uncertain (see reviews by e.g. Carroll et al., 1992; Frieman et al., 2008, for the current state of the field). Finally, the radiation component (CMB photons, neutrinos and other light relativistic components) contributes about 0.01%, $\Omega_r \approx 10^{-4}$.

At the present day, the expansion is dominated by the cosmological constant. This was not always the case, as evidenced by the scale factor dependencies of the different components in Eq. (1.5). This divides the history of the expanding Universe into three broadly defined periods after inflation: i) radiation dominated, ii) matter dominated and iii) Λ dominated epochs. The transitions between them are not abrupt, but boundaries can be set by finding the redshift of equality between two components. Thus defined, the end of the radiation-dominated era is reached at "equality", $z_{\rm eq} = \Omega_{\rm m}/\Omega_{\rm r} - 1 \approx 2.6 \times 10^3$. The matter-dominated era continues until $z_\Lambda = (\Omega_\Lambda/\Omega_{\rm m})^{1/3} - 1 \approx 0.4$, when the cosmological constant begins to take over.

1.1.2 Structure formation

The previous section discussed the evolution of the Universe as a whole, without a clear connection to the formation of smaller structures (such as galaxies and groups of galaxies) we observe to exist. The mechanism that drives structure formation is thought to be *gravitational instability* that amplifies the tiny ripples in the density distribution introduced by inflation. A convenient way to describe density perturbations is by introducing the dimensionless density contrast: $\delta(\vec{r}) = \rho_m(\vec{r})/\bar{\rho}_m - 1$, where $\bar{\rho}_m$ is the mean matter density of the Universe and $\vec{r}$ is the spatial coordinate. It is also mathematically convenient to discuss the density constant in Fourier space, where the real space fluctuation $\delta(\vec{r})$ of wavelength λ is replaced by the corresponding Fourier fluctuation, $\delta(\vec{k})$ with the wave number $k = |\vec{k}| = 2\pi/\lambda$. The inflation theory predicts the primordial density fluctuation field to be a Gaussian, scale-invariant random field for which all Fourier modes are independent of each other. All statistical properties of such a field are fully defined by a power spectrum, $P(k)$.

When the density contrast is small, $\delta \ll 1$, the growth of fluctuations can be followed with linear theory. The evolution of fluctuations of different wavelengths λ generally depends on the epoch of the expansion history of the Universe during which the fluctuation enters the causally connected Universe, bounded by the particle horizon. The proper radius of the particle horizon is defined by the distance the light is able to travel since the Big Bang: $\lambda_{\rm H}(t) \sim ct \sim cH^{-1}$ (see e.g. Peebles, 1993, for a detailed discussion).

Once a density fluctuation reaches $\delta \sim 1$, the linear theory starts to break down. A useful tool for understanding the following non-linear collapse is provided by the spherical collapse model (Gunn & Gott, 1972; for a textbook presentation, e.g. Coles & Lucchin, 2002). In this model, regions of high density decouple from the Hubble flow under the pull of their self-gravity and collapse just like closed matter-dominated Universes ($\Omega > 1$). The collapse is halted after a violent relaxation process governed by the virial theorem (which provides a connection between kinetic and potential energy in dynamically stable systems) forming a dark matter halo. The properties of a halo can also be estimated from the virial theorem. The spherical collapse theory predicts that the halo will have an overdensity, defined as $\Delta(\vec{r}) = \rho_m(\vec{r})/\bar{\rho}_m$, of $\Delta_{\rm coll} \approx 180$ (when $\Omega_{\rm m} = 1$), independent of halo mass and redshift[6]. The mass of the halo is then:

[6]Only true in the Einstein-de Sitter Universe.

$M_{\rm vir} = (4\pi/3)\bar{\rho}_m \Delta_{\rm coll} r_{\rm vir}^3$, where $r_{\rm vir}$ is the halo radius defined by its mass and predicted overdensity. The baryon gas collapsing into the halo is shock-heated to a temperature called the virial temperature:

$$T_{\rm vir} = \frac{\mu m_{\rm H} v_{\rm vir}^2}{2k}, \tag{1.6}$$

where μ is the mean molecular weight, $m_{\rm H}$ the mass of a hydrogen atom, k the Boltzmann constant and $v_{\rm vir} = \sqrt{GM_{\rm vir}/r_{\rm vir}}$ is the halo circular velocity at radius $r_{\rm vir}$. Of interest for some of the discussion in this thesis is the relation between the halo mass and virial temperature: (e.g. Peacock, 1999):

$$M_{\rm vir} \approx 2 \times 10^6 {\rm M}_\odot/h \left(\frac{T_{\rm vir}}{10^3{\rm K}}\right)^{3/2} \left(\frac{1+z}{10}\right)^{-3/2} \left(\frac{\mu}{1.22}\right)^{-3/2}. \tag{1.7}$$

The halos are the cradle for the formation of stars and galaxies (e.g. White & Rees, 1978; White & Frenk, 1991), which in turn produce the ionizing radiation that causes reionization and keeps the Universe ionized after it. We discuss the theory of galaxy formation in more detail in Chapter 4. For now, it is sufficient to say that the luminosity produced in halos is, among other things, a function of the halo mass, Eq. (1.7).

The dark matter is assumed to be pressureless and its non-linear evolution is fully determined by gravity. On the other hand, when considering baryonic matter, non-zero thermal pressure must be taken into account. The thermal pressure prevents the growth of density fluctuations in the baryonic component if the sound crossing time of the fluctuation λ/c_s is smaller than the free-fall time: $\lambda \leq \lambda_J \equiv c_s \pi^{1/2} (G\rho_m)^{-1/2}$, where $c_s = \sqrt{\gamma kT/(\mu m_{\rm H})}$ is the sound speed and γ is the ratio of specific heats (equal to $5/3$ for ideal, mono-atomic gas). The Jeans scale λ_J defines the Jeans mass, $M_J = (4\pi/3)\rho_m(\lambda_j/2)^3$, which can also be written as (e.g. Barkana & Loeb, 2001):

$$M_J \approx 6 \times 10^7 {\rm M}_\odot/h \left(\frac{T}{10^4{\rm K}}\right)^{3/2} \left(\frac{1+z}{10}\right)^{-3/2} \left(\frac{\mu}{1.22}\right)^{-3/2}. \tag{1.8}$$

Prior to reionization, the gas temperature evolves in two stages. In the first stage, it remains coupled to the CMB temperature due to the scattering of residual free electrons until redshift $1 + z_{\rm dec} \approx 137(\Omega_b h^2/0.022)^{2/5}$ (e.g. Peebles, 1993). At the present day, the CMB radiation has the spectrum of an almost ideal black body with temperature $T_{\rm CMB,0} = 2.73\,{\rm K}$. The temperature of the CMB (therefore also the gas temperature prior to $z_{\rm dec}$) evolves as $T_{\rm CMB} = T_{\rm CMB,0}(1+z)$ due to the expansion. The Jeans mass remains independent of redshift in this period. After $z_{\rm dec}$, the gas cools adiabatically, $T \propto (1+z)^2$,

resulting in $T \approx 30$ K at $z = 10$ with the corresponding Jeans mass at $\Delta = 1$ being $M_J \approx 6 \times 10^3\,\mathrm{M}_\odot/h$, from Eq. (1.8). The Jeans scales are relevant for photo-ionization feedback, which we discuss later.

The number density of halos is best obtained from numerical simulations that solve the equations of gravitational collapse (we will use such simulations extensively throughout this thesis). Alternatives are analytically derived functions that approximate the results of computer simulations as simpler one-dimensional functions. While failing to provide the spatial distribution of halos, these analytical techniques are extremely useful as they allow for much faster exploration of the relevant parameters. The original method for this purpose was developed by Press & Schechter (1974). In their approach, the abundance of halos at redshift z is obtained from linear theory by applying the spherical collapse theory to the density field. The method provides the co-moving number density of halos, $\mathrm{d}n$, with mass between M and $M + \mathrm{d}M$ in the form (Peebles, 1993):

$$M\frac{\mathrm{d}n}{\mathrm{d}M} = \left(\frac{2}{\pi}\right)^{1/2} \frac{\mathrm{d}(\ln\sigma)}{\mathrm{d}(\ln M)} \frac{\rho_{m,0}}{M} \nu_c e^{-\nu_c^2/2}, \tag{1.9}$$

where $\rho_{m,0}$ is the mean mass density at present day, $\sigma = \sigma(M)$ is the standard deviation of the density contrast smoothed over a window containing mass M, and $\nu_c = \Delta_{\mathrm{coll}}/\sigma$. Since its introduction, the analytic approach has been improved by many authors (e.g. Bond et al., 1991; Lacey & Cole, 1994; Sheth & Tormen, 2002; Reed et al., 2006).

Equation (1.9) also demonstrates one of the main features of current structure formation theory: *the hierarchical buildup of structure*. The standard deviation increases with decreasing mass in the window, leading to the (on average) earlier collapse of small fluctuations. The buildup of structure proceeding from less to more massive halos is evidenced by, for example, the fact that the galaxy sized halos are observed at redshifts $z \lesssim 10$ while the larger, cluster sized halos (which hold hundreds of galaxies) do not assemble until $z \lesssim 1$.

1.2 Reionization: transition to an ionized Universe

The formation of cosmic structure leads to the formation of radiation sources that transform the neutral gas in the intergalactic medium to its highly ionized present day state

during the period called reionization[7]. A number of review articles about this epoch have been written in recent years (see e.g. Barkana & Loeb, 2001; Ciardi & Ferrara, 2005; Loeb, 2006; Ciardi, 2007). The name itself suggests that during this epoch the Universe becomes highly ionized for the second time. Indeed, for the first ≈ 350000 years, the Universe was hot and dense enough for the collisions between particles to prevent the formation of neutral atoms. Only at redshift $z \approx 1100$ has the Universe cooled down significantly enough (to about $T \approx 3000K$) for the neutral atoms to form. This event is known as the cosmological *recombination* (see Sunyaev & Chluba, 2008, for a recent review) and it marks the first major phase transition in the ionization state of the Universe. Recombination significantly decreased the Thompson electron scattering optical depth, freeing most of the photons that were tightly coupled to matter before it. This radiation constitutes the cosmic microwave background.

After recombination, the Universe continues to expand and cool with the majority of the baryonic matter in the neutral state. The transition back to the highly ionized state begins with the formation of the first luminous structures capable of producing ionizing radiation (photons with energies $h\nu \geq 13.59$ eV needed to ionize hydrogen). The exact nature of the radiation sources and their relative contribution to the total ionizing photon budget are one of the many uncertainties in the understanding of reionization. A probable candidate can be found in the first stars (also called Population III stars) formed in the Universe (e.g. Abel et al., 2002a; Bromm & Larson, 2004). These stars are expected to form out of primordial composition gas condensing in small mass halos, which are the first to form in the hierarchical structure formation scenario. In order for the gas to collapse, it must be able to efficiently cool in order to reduce its thermal pressure. In general (e.g. Rees & Ostriker, 1977; Schneider et al., 2002), the cooling is efficient when $t_{\rm cool} \ll t_{\rm ff}$, where $t_{\rm cool} = 3nkT/2\Lambda(n,T)$ is the cooling time, $t_{\rm ff} = (\pi/G\rho)^{1/2}$ is the free-fall time, n is the number density of gas with the mass density ρ and $\Lambda(n,T)$ is the net radiative cooling rate per unit volume in units of erg $\mathrm{cm}^{-3}\,\mathrm{s}^{-1}$. In the primordial gas of temperatures $T \lesssim 10^4$K, the most efficient cooling mechanism is the radiative de-

[7]Note that all mentions of reionization in this thesis refer exclusively to the reionization of cosmic hydrogen. The (complete) reionization of helium is, while related, a different problem from the one discussed here: it occurs at a significantly later time ($z \approx 3$) and is most probably driven by radiation from quasars which are necessary to produce significant amount of high energy photons ($h\nu \gtrsim 54$ eV) needed to twice ionize helium.

excitation of collisionally excited *molecular hydrogen* (e.g. Abel et al., 2000; Bromm et al., 2002). The conditions needed for effective molecular hydrogen cooling are found in halos with virial temperatures $T_{\rm vir} \sim 10^3{\rm K}$ (e.g. Thoul & Weinberg, 1996; Yoshida et al., 2003; Bromm & Larson, 2004) corresponding to virial mass of, according to Eq. (1.7), $M_{\rm vir} \sim 10^6{\rm M}_\odot/h$. The typical mass of the stars formed, although still uncertain, is expected to be $\sim 100M_\odot$, much more massive than typical present day stars ($\sim 1M_\odot$). Their primordial gas composition (i.e. practically zero-metallicity) means that their luminosity is fuelled by proton-proton nuclear burning.

The first stars are expected to be hotter, and therefore produce more ionizing radiation, than their more metal-enriched counterparts. For example, Bromm et al. (2001a) find that a Pop III star with mass $\approx 300M_\odot$ resembles a black body of temperature $\approx 10^5{\rm K}$ and produces about an order of magnitude more H and HeI ionizing photons than a Pop II star of comparable mass. The case for Pop III stars being the dominant sources of ionizing radiation for the epoch of reionization is complicated by the fact that the intergalactic H_2 is easily dissociated by the soft-UV radiation in the Lyman-Werner bands. This suppression is very efficient and just a single generation of Pop III stars can significantly suppress the subsequent star formation in their host halos as well as their surroundings (Omukai & Nishi, 1999; Nishi & Tashiro, 2000; Oh & Haiman, 2002; Susa & Umemura, 2004a; Ahn et al., 2008). On the other hand, new H_2 molecules can form in gas compressed by shock fronts (e.g. Ferrara, 1998; Ricotti, 2002) or colliding atomic gas clouds (Cen, 2005). This complex feedback loop makes estimates of Pop III stars contribution to the ionizing photon budget uncertain. The Pop III era ends when the metallicity reaches $Z/Z_\odot \sim 10^{-3.5}$ and the formation of normal, lower mass Population II stars begins (e.g. Omukai, 2000; Bromm et al., 2001b; Bromm & Larson, 2004; Fang & Cen, 2004).

In halos with virial temperatures $T_{\rm vir} \gtrsim 10^4{\rm K}$, the gas can cool through the collisional excitation and radiative de-excitation of atomic hydrogen and helium and begin the formation of galaxies as we know them today. Pop II stars found in them may also be a dominant source of ionizing radiation. This scenario allows the application of the more extensive knowledge of galaxy formation at lower redshifts to the epoch of reionization. A number of works have found the Pop II stars with standard or somewhat top-heavy initial mass functions (IMF) to be capable of single-handedly completing reionization

within the redshift range suggested by current observations (e.g. Gnedin, 2000a; Ciardi et al., 2003; Sokasian et al., 2003; Iliev et al., 2006a; Benson et al., 2006). Conversely, there is still controversy about whether the currently observed population of galaxies at $z \lesssim 6$ is sufficient to keep the Universe in its highly ionized state (e.g. Bolton & Haehnelt, 2007; Bouwens et al., 2008b; Gnedin, 2008; Pawlik et al., 2009), which may or may not be important for the reionization itself. These uncertainties suggest that the contribution of quasars or some more exotic sources (e.g. Rees, 1986; Tegmark et al., 1993; Dijkstra, 2006; Srbinovsky & Wyithe, 2007; Loeb, 2009), though thought to be sub-dominant (e.g. Madau, 1999; Faucher-Giguere et al., 2008), cannot be completely discounted.

The issue of deciding which sources of ionizing radiation dominate during EoR is further complicated by the highly uncertain *escape fraction*, $f_{\rm esc}$, i.e. the fraction of photons produced that reaches the intergalactic medium, from galaxies or quasars. The observations used to infer the escape fraction of UV radiation from our own galaxy (e.g. Bland-Hawthorn & Maloney, 1999) as well as local starbursts (e.g. Hurwitz et al., 1997; Heckman et al., 2001; Inoue et al., 2006) and Lyman-break galaxies at $z \sim 3$ (Shapley et al., 2006) find $f_{\rm esc} \approx 0.01 - 0.1$. These values are supported by some hydrodynamical simulations (e.g. Fujita et al., 2003; Razoumov & Sommer-Larsen, 2006, 2007; Gnedin, 2008). The escape fraction may depend strongly on the porosity of the interstellar medium within the galaxy, caused by the density distribution or, for example, the presence of supernovae driven winds (e.g. Ciardi et al., 2002; Clarke & Oey, 2002). Some of the latest models that attempt to include these density inhomogeneities in simulations are finding a significantly larger escape fraction ($f_{\rm esc} \sim 0.5$) in high-z galaxies (e.g. Wise & Cen, 2009; Razoumov & Sommer-Larsen, 2009). In this thesis, $f_{\rm esc}$ is treated as a free parameter.

If the sources of ionizing radiation are stars or quasars, they are associated with dark matter halos and therefore spatially clustered. This leads to a highly inhomogeneous progression of reionization. The overdense gas around the sources is generally the first to be ionized. The low density gas becomes ionized when the radiation breaks into the underdense voids. This reionization progression is dubbed inside-out, as it proceeds from high to low density regions. The growth of HII regions continues in this fashion until they begin to overlap in the voids. The reionization is completed once the overdense regions further away from the sources are finally ionized, as they are circumvented in the earlier stages due to their higher recombination rate. In this final stage, the reionization

therefore proceeds outside-in. This scenario is supported by both analytical and numerical works (e.g. Gnedin, 2000b; Miralda-Escude et al., 2000; Ciardi et al., 2003; Furlanetto et al., 2004; Finlator et al., 2009a), but it is not universally accepted. More recent high resolution simulations suggest the exclusively inside-out reionization progression (e.g. Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007; Zahn et al., 2007, see also Chapter 5).

Yet another aspect that may be important for understanding reionization is the back-reaction of the ionization field onto the radiation sources. This photo-ionization feedback affects the formation and evolution of galaxies in two ways. First, the gas is heated to $T \sim 10^4\mathrm{K}$ (e.g. Miralda-Escude & Rees, 1994; Hui & Haiman, 2003; Tittley & Meiksin, 2007) which, according to Eq. (1.8) raises the mean density ($\Delta = 1$) Jeans mass to $M_J \sim 10^9 \mathrm{M}_\odot/\mathrm{h}$. This so-called *Jeans filtering* (e.g. Shapiro et al., 1994; Gnedin & Hui, 1998) prevents the baryons from collapsing into halos below the Jeans mass. The increase in temperature may also evaporate the already collapsed gas out of the halo (e.g. Thoul & Weinberg, 1996; Barkana & Loeb, 1999; Shapiro et al., 2004; Iliev et al., 2005a; Okamoto et al., 2008; Pawlik et al., 2009) back into the IGM. Note that the characteristic mass below which the halos are affected by the increase in temperature of the gas (e.g. Gnedin, 2000b; Hoeft et al., 2006; Okamoto et al., 2008) is still under considerable debate. Second, the ionization of neutral gas in the halo can significantly affect the halo when the main cooling channel is the atomic hydrogen lines (e.g. Efstathiou, 1992; Wiersma et al., 2009). All of these processes result in the suppression of the gas cooling and star formation rate within the affected halo.

1.2.1 Observational constraints

The epoch of reionization is still beyond the direct reach of current telescopes. There are, however, a number of indirect clues that place some constraints on the epoch. Here we present a short overview of the available data. For more details, we refer the reader to the reviews by Fan et al. (2006) or Furlanetto et al. (2006). The first constraint is provided by the observation of the absorption of light from high redshift quasars caused by the intervening gas. The objects containing traces of residual neutral hydrogen produce a forest of absorption features blueward of the Lyman-α line in the spectrum of background quasars (see Rauch, 1998, for a review). Gunn & Peterson (1965) (GP) predicted

that, in the case of a neutral IGM, the quasar spectra would show near complete absorption instead of a series of absorption lines. The tentative evidence of the GP trough was first observed by Becker et al. (2001) in a quasar at $z \approx 6.28$ found in the SDSS[8] sample. The spectra of quasars of slightly lower redshifts do not exhibit the trough feature, which suggested a significant change in the ionization state of hydrogen at $z \approx 6$, i.e. the end of the reionization era. Similar results were found with a bigger quasar sample by Fan et al. (2002).

The second constraint is provided by the CMB observations. The CMB photons scatter off the free electrons produced by reionization. This smooths the CMB temperature anisotropies and generates a polarisation signal on angular scales $\lesssim 10^{\circ}$, corresponding to the scale of the horizon at reionization (e.g. Zaldarriaga, 1997; Hu & White, 1997; Komatsu et al., 2009). Analysis of the 7-year WMAP[9] CMB data suggests that, assuming an instantaneous complete reionization model, the redshift of reionization is $z = 10.5 \pm 1.2$ (Komatsu et al., 2010). The CMB may also hold signatures of the inhomogeneous nature of reionization in anisotropies on much smaller scales (e.g. Bruscoli et al., 2000; Benson et al., 2001; Santos et al., 2003).

The quasar data only probes the very end of reionization, because strong absorption appears in the spectra even for trace neutral fractions, $x_{\rm HI} \equiv n_{\rm HI}/n_{\rm H} \gtrsim 10^{-4}$. On the other hand, the current CMB polarisation data provides only an integrated constraint, since the electron scattering optical depth is largely insensitive to the reionization history (i.e. the evolution of $x_{\rm HI}$ with redshift). Much stronger constraints are expected to come from the results of the upcoming 21 cm telescopes (Scott & Rees, 1990; Madau et al., 1997; Furlanetto et al., 2006). This next generation of radio telescopes (e.g. LOFAR, MWA) will provide the statistical measurements of the 21 cm signal, but we may have to wait for the following generation (e.g. SKA, still in planing stages) to obtain the more detailed maps of neutral hydrogen as a function of redshift, giving a direct look into the reionization process.

[8] http://www.sdss.org/
[9] http://map.gsfc.nasa.gov/

1.2.2 Numerical simulations of reionization

Numerical simulations are currently the most promising tool for the theoretical study of the epoch of reionization. While analytic and semi-analytic models provide significant glimpses into the process (e.g. Madau, 1999; Miralda-Escude et al., 2000; Barkana & Loeb, 2004; Furlanetto et al., 2004; Zahn et al., 2007; Choudhury et al., 2008), only full numerical simulations can combine the inhomogeneous cosmological density field with full 3D radiative transfer without any assumed symmetries which is needed to model the growth of HII regions during reionization self-consistently. The techniques used have significantly evolved in the past 5 to 10 years, enough for the current state of the field to warrant a review, recently provided by Trac & Gnedin (2009).

The numerical simulations bring their own set of limitations to the already large set of uncertainties about the reionization process. For studying the large-scale properties of reionization, the cosmological density field should be represented by a simulation box of $\geq 100\,\mathrm{Mpc}/h$ side length for two major reasons. First, this box size is needed to guarantee that the simulation is representative of the cosmic mean, i.e. that it provides an accurate sampling of the halo mass function at redshifts of interest (Barkana & Loeb, 2004; Iliev et al., 2006a). Second, the HII regions during reionization are expected to be several tens of Mpc in size (e.g. Furlanetto et al., 2004) so a large simulation box size is needed to provide a significant statistic. In itself, the large box size is not an issue - the main problem comes from the need to resolve all the halos where stars can form, which, if Pop III stars are included, means $M_{\mathrm{halo}} \sim 10^6\,\mathrm{M}_\odot/h$. The minimum resolved mass must be even lower to include all the density field inhomogeneities that affect the recombination rate, for which the relevant scale is given by the pre-heating Jeans mass. Current state-of-the-art simulations of dark matter only resolve halos down to $M \approx 10^8\mathrm{M}_\odot/h$ (which generally includes all the halos capable of atomic line cooling at the redshifts of interest) in the $100\,\mathrm{Mpc}/h$ box sizes (Shin et al., 2008; Iliev et al., 2008; Trac et al., 2008). Hydrodynamical simulations that would include the gas physics at the same resolution are, at this time, impossible. To combat the very high computational cost of these simulations, the majority of the numerical models of reionization represents the density field by a series of static snapshots computed beforehand (e.g. Ciardi et al., 2003; Iliev et al., 2006a; McQuinn et al., 2007, see Chapter 5). This approximation is justified as the cosmological HII regions expand much faster than the underlying density evolves (Shapiro & Giroux,

1987) and it allows for a single density field to be reused with different assumptions about the other components of the model (e.g. the source population). Note that some groups follow a more consistent path of evolving the density concurrently with the ionization field (e.g. Trac & Cen, 2007; Baek et al., 2009), though the potential benefits of this approach compared to the former one have not been explored in detail. The unresolved density fluctuations which may have a large effect on the recombination rate (which is $\propto n_{\mathrm{H}} n_e$, i.e. the square of the density) are often ignored (e.g. Ciardi et al., 2003; Trac & Cen, 2007) or represented by so-called clumping factors that are calibrated using higher mass resolution, but smaller box size simulations (e.g. Iliev et al., 2005a; McQuinn et al., 2007; Iliev et al., 2007; Kohler et al., 2007, see Chapter 3 for a more detailed discussion).

Regardless of the previously discussed issues with the density field representation, the most severe challenge in numerical reionization simulation must be the accurate treatment of the radiative transfer of ionizing photons. In general, the radiative transfer equation that is solved to obtain the ionizing intensity throughout the simulation box, which we present in the next section, has seven dimensions: three spatial, two directional and one for both frequency and time. The issue is further complicated by the non-local influence of radiation e.g. similarly to gravitation, the radiation from a single source may reach every volume element in the simulation box. Unlike gravity, the radiation interaction between the two volume elements depends on the state of the elements between them. Finally, the very large number of sources found in the boxes required for reionization simulations significantly increases the computational cost, especially for some traditional methods where this cost scales linearly with the number of sources (i.e. every source interacts with every volume element, see Chapter 2 for a more detailed look at such methods).

To make these problem tractable, a number of simplifying assumptions can be used. This naturally lead to the development of a wide range of radiative transfer methods designed specifically for reionization simulations. In general, the computational cost is decreased by reducing the dimensionality of the radiative transfer (RT) equation and by lessening the steep, linear scaling of the cost with the number of sources. The dimensionality is commonly reduced by tracing only a few relevant radiation frequencies (e.g. assuming that all photons have the hydrogen ionization energy, $\approx 13.6\,\mathrm{eV}$) and by solving the radiative transfer equation along discrete directions (so-called ray trac-

ing). Monte Carlo RT methods (e.g. Maselli et al., 2003; Semelin et al., 2007; Altay et al., 2008) trace discrete packets of radiation energy along randomly cast rays and reduce the cost by claiming that the ionization field is sufficiently resolved with a set number of such packets irrespective of the number of sources. A popular method of adaptive ray-tracing (Abel & Wandelt, 2002b) keeps the angular resolution of rays constant by splitting them periodically with increasing distance from the source. The schemes using this technique decrease the linear scaling by grouping the sources (Razoumov & Cardall, 2005), limiting the splitting of rays after some travelled distance (McQuinn et al., 2007) or merging near-parallel rays (Trac & Cen, 2007). Similarly, the rays may be emitted in a set number of directions from each volume element to its neighbours and grouped before being emitted again in direction-conserving fashion (e.g. Ritzerveld et al., 2003; Pawlik & Schaye, 2008; Paardekooper et al., 2009, discussed in more detail in Chapter 3). An altogether different method, introduced by Gnedin & Abel (2001), considers moments of the radiation field. In this way, the radiative transfer equation is simplified to a system of differential equations for the photon density and flux, tied together with a term called the Eddington tensor. This method is explicitly independent of the number of radiation sources and its implementations generally differ in the way the Eddington tensor is computed (Aubert & Teyssier, 2008; Finlator et al., 2009b; Petkova & Springel, 2009). Note that the linear scaling of computational cost with the number of sources is not an insurmountable obstacle as demonstrated by the C^2-ray scheme (Mellema et al., 2006). This method reduces the computational cost of radiative transfer by properly modeling the time evolution of the optical depth along rays which allows the use of very long computation time steps compared to other previously mentioned methods (this is further elaborated in Chapter 2). Many of these codes have been compared by Iliev et al. (2006b, 2009) who found them to produce generally the same results in a series of simple tests, suggesting that a certain level of maturity of the radiative transfer codes has been reached.

Even with all the mentioned improvements, the cost of solving the radiative transfer equation is still much higher than that of, e.g. gravitational interaction. This results in the spatial resolution of grids used to advance the ionization field to be significantly worse than that of the underlying density field. Indeed, practically all the large-scale reionization schemes interpolate the density field to a uniform Cartesian grid of mod-

est resolution (compared to that of the density) before evolving the ionization field (e.g. Ciardi et al., 2003; Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007; Shin et al., 2008; Iliev et al., 2008). This in turn further smooths the density field, decreasing the importance of recombinations. Several methods have been implemented to operate on some type of spatially adaptive density field representation, e.g. smoothed particle hydrodynamics (SPH, Susa, 2006; Semelin et al., 2007; Altay et al., 2008; Pawlik & Schaye, 2008; Petkova & Springel, 2009), adaptive mesh refinement (AMR, Rijkhorst et al., 2006; Wise & Abel, 2007; Reynolds et al., 2009) or more exotic grids (e.g. Voronoi tessellation, Ritzerveld et al., 2003; Paardekooper et al., 2009, see Chapter 3). Still, almost none of these methods have so far been used to model large scale reionization in the required box sizes ($\approx 100\,\mathrm{Mpc}/h$, but see Chapter 5).

1.3 Basic concepts of radiative transfer

As a major part of this thesis concerns the techniques for numerical radiative transfer, it is useful to introduce the main quantities in the theory of radiative transfer at this stage[10] . For a more detailed discussion, we suggest the textbooks by Rybicki & Lightman (1986), Chandrasekhar (1950) or Castor (2004).

In essence, RT theory describes the propagation of electromagnetic radiation energy through any kind of medium. The basic variable is the *specific intensity* at frequency ν:

$$I_\nu = \frac{\mathrm{d}E}{\mathrm{d}A\mathrm{d}t\mathrm{d}\Omega\mathrm{d}\nu}, \tag{1.10}$$

which is defined as the amount of energy $\mathrm{d}E$ passing through surface $\mathrm{d}A$ in time $\mathrm{d}t$ and in frequency range $\mathrm{d}\nu$ in the direction encompassed by the solid angle $\mathrm{d}\Omega$. The specific intensity in the cgs system has the units of $\mathrm{erg\,s^{-1}cm^{-2}ster^{-1}Hz^{-1}}$. In a general radiation field, I_ν is a function of seven variables:

$$I_\nu \equiv I_\nu(x, y, z, \theta, \phi, t), \tag{1.11}$$

where x, y and z denote spatial coordinates, θ and ϕ direction angles and t time. The spherically averaged intensity crossing surface $\mathrm{d}A$ defines the *mean intensity*:

$$J_\nu = \frac{1}{4\pi}\int I_\nu \mathrm{d}\Omega, \tag{1.12}$$

[10]Note that this whole discussion neglects effects of the cosmological expansion on radiation propagation.

which is commonly used in equations that describe interaction of radiation with matter, for example, in the definition of the photo-ionization rate we introduce in Eq. (2.12). Telescopes detect the *net flux* of electromagnetic radiation, i.e. the amount of energy passing through surface $\mathrm{d}A$. The net flux is related to the specific intensity as:

$$F_\nu = \int I_\nu \cos\theta \mathrm{d}\Omega, \tag{1.13}$$

where $\cos\theta = \vec{n} \cdot \vec{n}'/(|\vec{n}||\vec{n}'|)$ is the cosine of the angle between the normal vector of surface $\mathrm{d}A$, $\vec{n}$, and the directional vector of solid angle $\mathrm{d}\Omega$, $\vec{n}'$.

It can be shown that, in the case of no energy gains or losses, the specific intensity along a single ray remains constant (see e.g. Rybicki & Lightman, 1986). This means that the intensity of radiation at position $\vec{r} = x\vec{e_x} + y\vec{e_y} + z\vec{e_z}$ travelling in the direction given by unit vector $\vec{n} = \sin\theta\cos\phi\vec{e_x} + \sin\theta\sin\phi\vec{e_y} + cos\theta\vec{e_z}$ is the same as the intensity at a point $s = c\tau$ away after time τ, where c is the speed of light and s the distance light travels in τ. Using Eq. (1.11) we get:

$$I_\nu(\vec{r} + \vec{n}c\tau, \vec{n}, t + \tau) = I_\nu(\vec{r}, \vec{n}, t). \tag{1.14}$$

After the left-hand side is Taylor-expanded, the right-hand side transferred to the left and the whole expression divided by $c\tau$, we obtain:

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \vec{n}\nabla I_\nu = 0, \tag{1.15}$$

which is the radiative transfer equation in the case of no energy gains or losses. These two effects are, in general, described by the emission and absorption coefficients, j_ν and k_ν, respectively which give the amount of radiation energy of frequency ν gained or lost per unit time per unit solid angle per unit volume. The energy gains along path element $\mathrm{d}s = c\mathrm{d}\tau$ usually[11] do not depend on the value of I_ν along the path, $\mathrm{d}I_\nu = j_\nu \mathrm{d}s$, but the losses do, $\mathrm{d}I_\nu = -k_\nu I_\nu \mathrm{d}s$. Taking the change in the energy into account leads us to the default form of the *radiative transfer equation* (RTE):

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \vec{n}\nabla I_\nu = j_\nu - k_\nu I_\nu. \tag{1.16}$$

When solving the radiative transfer along specific rays and assuming that the intensity remains constant during time, Eq. (1.16) transforms from a partial differential

[11] There are some astrophysical systems that feature stimulated emission, e.g. naturally occurring masers in the ISM (e.g. Strelnitski, 1997).

equation to an ordinary differential equation along the ray:

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = j_\nu - k_\nu I_\nu. \tag{1.17}$$

Note that this formulation ignores the finite speed of light. In the case of point sources and no scattering, $j_\nu = 0$, which leads to the simple solution:

$$I_\nu(s) = I_\nu(0)\exp(-k_\nu s), \tag{1.18}$$

i.e. the solution is that of pure attenuation. This relation will be our main tool for evolving the ionization fields in methods we present in both Chapters 2 and 3.

1.4 Thesis overview

In this section, we provide a short overview of the contents of subsequent chapters.

Chapter 2: we discuss the Hybrid characteristics (HC) radiative transfer method (Rijkhorst et al., 2006), implemented in the adaptive mesh refinement code FLASH (Fryxell et al., 2000). After introducing both FLASH and HC, we present the improvements and fixes we introduced to the method. The most important of these is the inclusion of both spatial and temporal photon-conservation (the terminology is introduced in this chapter), which results in computational efficiency improvement measured in orders of magnitude. We briefly discuss the physics of ionization fronts before presenting the results of the radiative hydrodynamics tests we performed for the second Cosmological Code Comparison project (Iliev et al., 2009). These tests show our improved HC method to be in excellent agreement with other RT codes taking part in the project.

Chapter 3: we introduce another radiative transfer method, SimpleX (Ritzerveld et al., 2003). This scheme solves the radiative transfer equation by transporting photon packets between neighbouring Voronoi tessellation cells. The main features of the scheme are an adaptive grid for representing the density field (we construct the tessellation directly on the particles from N-body runs, thus matching their spatial resolution) and a computational cost that does not depend on the number of sources. We show that the original SimpleX implementation, specifically the way it sampled the density field, introduced significant numerical errors, as evidenced by the results of Test 4 in Iliev et al. (2006b). We introduce a different density field sampling, directly proportional to the local density and corresponding to an unweighted random selection of N-body particles,

and show that it results in SimpleX finding much better agreement with other codes in the aforementioned Test 4. The rest of the chapter focuses on how to accurately compute the recombination rate on a coarse RT grid that does not resolve all the density field data. First, we show that the clumping factor must be a function of both the local overdensity and the spatial window over which it is computed in order to avoid severe overestimates of the total recombination rate. Second, we argue for a different interpretation of the ionized fraction in highly optically thick cells in order to obtain the correct recombination rates. The importance of this change is demonstrated in the case of an HII region propagating through a cosmologically expanding uniform density field. The SimpleX method with all the changes discussed in this Chapter is used to evolve the ionization field in our reionization simulations presented in Chapter 5.

Chapter 4: We introduce GALFORM, a semi-analytic scheme for modelling the formation and evolution of galaxies (Cole et al., 2000). We use the GALFORM scheme to provide the source model for the full reionization simulations of Chapter 5. We focus on the specific implementation introduced by Baugh et al. (2005) and examine the role of different parameters in the model on the production of ionizing photons during reionization. We find that the largest contribution comes from the assumed top-heavy IMF during merger-triggered bursts of star formation. In the second part of the Chapter, we compare the predictions of the Baugh et al. (2005) model, and some variations of it, to the observed population of Lyman-break galaxies at $z \sim 6 - 10$ (Bouwens et al., 2008b, 2009a,b). We find a good agreement between the observed data and the predictions of the default Baugh et al. model, in terms of both the luminosity function and the Schechter fit parameters.

Chapter 5: We present our method for simulating large scale reionization, constructed by coupling the SimpleX and GALFORM schemes, both discussed in the two previous chapters. The key element is the communication between the two codes that allows for the inhomogeneous reionization to be taken into account when computing the evolution of galaxies. The resulting method can compute the evolution of the ionization field during reionization with spatial resolution comparable to (or better than) that of other state-of-the-art codes, in a fraction of the CPU time. This makes it ideal for exploring the currently largely unconstrained parameter space of reionization. In the first half of the Chapter, we show that our results are converged for the chosen RT grid resolution. The

second part discusses some of the first results we obtained addressing the issues of: i) the topology of reionization and its dependence on the local overdensity, ii) the morphology of HII regions in a source model dominated by starbursts and iii) the importance of source suppression for the progression of reionization.

Chapter 2

Radiative hydrodynamics with FLASH

2.1 Introduction

Many problems in astrophysics, e.g. the formation and evolution of stars and galaxies, evolution of supernova remnants, stellar winds and other outflows just to name a few, require accurate numerical treatment of multi-dimensional fluid flows. While some systems can be successfully modelled in two dimensions or with imposed symmetries, a majority of problems requires a general, three dimensional approach. Without any simplifying assumptions, the numeric solution of the astrophysical fluid flows in three dimensions requires the memory and CPU cycles available only with the use of modern supercomputers. Even then, the brute force approach is impractical and sophisticated numerical techniques must be developed. The models are further complicated by the addition of radiative transfer. The radiation field can have a large impact on astrophysical fluid flows, yet its characteristic temporal and spatial scales are often very different from the hydrodynamic ones, evident in the orders of magnitude difference between the speed of light and the speed of sound in the cosmic gas. Examples of the complex interaction between radiation and fluid flow can be found in the study of ionization fronts, that will be discussed in detail throughout this chapter.

The subject of this chapter is a radiative hydrodynamics scheme constructed by adding radiative transfer into the adaptive mesh refinement (hereafter, AMR) hydrodynamics code FLASH (Fryxell et al., 2000). The radiative transfer method in question, Hybrid characteristics (Rijkhorst, 2005; Rijkhorst et al., 2006), efficiently traces rays across the block structured AMR grid employed by FLASH and computes the intensity of radiation at every computational cell. In the following section, we discuss the basics of

FLASH, focusing mostly on the AMR grid structure. In Section 2.3, we describe the Hybrid characteristics (hereafter, HC) RT scheme as it was implemented in FLASH. Section 2.4 introduces the basic physics of ionization fronts. Section 2.5 discusses the errors arising in RT schemes due to the necessary spatial and temporal discretization of the RT equation, commonly called photon conservation issues as well as the steps we take to avoid them in HC. Finally, in Section 2.6 we present the results of the radiative transfer tests we performed with our updated version of HC for the second RT Code Comparison Project paper (Iliev et al., 2009). Such a precise radiative hydrodynamics method will allow us to study several problems relevant to the epoch of reionization, e.g. the gas dynamical evolution of HII regions around first stars or the photo-evaporation of minihalos encountered by cosmological ionization fronts.

2.2 FLASH: AMR hydrodynamics

2.2.1 Euler equations

Inviscid astrophysical fluid flows can be described with a system of partial differential equations called *Euler equations* (e.g. Landau & Lifshitz, 1959):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \vec{v} = 0 \tag{2.1}$$

$$\frac{\partial \rho \vec{v}}{\partial t} + \nabla \cdot (\vec{v} \otimes (\rho \vec{v})) + \nabla p = 0 \tag{2.2}$$

$$\frac{\partial E}{\partial t} + \nabla \cdot (E + p)\vec{v} = 0, \tag{2.3}$$

where ρ is the fluid mass density, $\vec{v}$ is the fluid velocity vector, p is the pressure and $E = \frac{1}{2}\rho v^2 + \rho e$ is the total energy per unit volume, with e being the internal energy per unit mass. This system is closed once the relation between pressure, density and internal energy is supplied, given by the equation of state. In essence, these equations represent three conservation laws for the fluid motion: mass, momentum and energy in Eqs. (2.1), (2.2) and (2.3) respectively.

There are two fundamental approaches to solving the Euler equations numerically. In the *Eulerian description*, the computational mesh is fixed in space. The fluid flow is modelled by computing the fluxes of mass, momentum and energy across a static grid.

There is a large degree of freedom in choosing the shape of the grid. It can be structured (e.g. uniform Cartesian mesh; Norman, 2000) or unstructured (e.g. Delaunay triangulation; Xu, 1997). Structured grids are not necessarily limited to a single spatial resolution: we will discuss adaptive computational meshes in the next section. On the other hand, in the *Lagrangian description*, the mesh moves along with the fluid flow. The mass of every grid cell is constant in time i.e. there is no mass flux between cells. Commonly used Lagrangian methods are mesh-free particle based methods grouped under the name smoothed particle hydrodynamics (hereafter SPH; Monaghan, 1992). Finally, a combination of the two approaches leads to *arbitrary Lagrangian Eulerian* methods. Here, the grid is allowed to move, but the mass in the cell is not constant, i.e. the advection fluxes need to be computed (e.g. Gnedin, 1995; Pen, 1998; Springel, 2009). The strengths and weaknesses of many of these methods have been tested in various comparison papers (e.g. Frenk et al., 1999; Agertz et al., 2007; Tasker et al., 2008).

2.2.2 Adaptive mesh refinement

In general purpose Eulerian schemes, Cartesian grids are most commonly used, due to their conceptual simplicity and easy implementation. While a uniform mesh is the obvious choice, its use is often impractical if the dynamic range of the problem is too large. To properly model such systems, the mesh must resolve the smallest spatial scales relevant to the problem, resulting in a highly redundant number of cells on larger scales. To circumvent this issue, the so-called adaptive mesh refinement (AMR) grids can be used (Berger & Oliger, 1984; Berger & Colella, 1989). In these schemes, the problem is initially represented by a coarse grid. Higher resolution meshes are placed where necessary, according to some criteria, until all the features in the problem have been resolved. Special care is taken to ensure that the fluxes between cells with different resolution are correct and all quantities are conserved. At every time step, the position and scale of the "interesting" areas may change and an AMR scheme must be able to adapt the mesh as needed, by both refining and derefining the grid to match the new state of the system.

The main distinction between AMR schemes is in the way the mesh cells are grouped. In the *cell-structured AMR*, every mesh cell can be independently refined if the refinement criterion is met (e.g. Raga et al., 2000; Teyssier, 2002; Lim & Mellema, 2003). On the

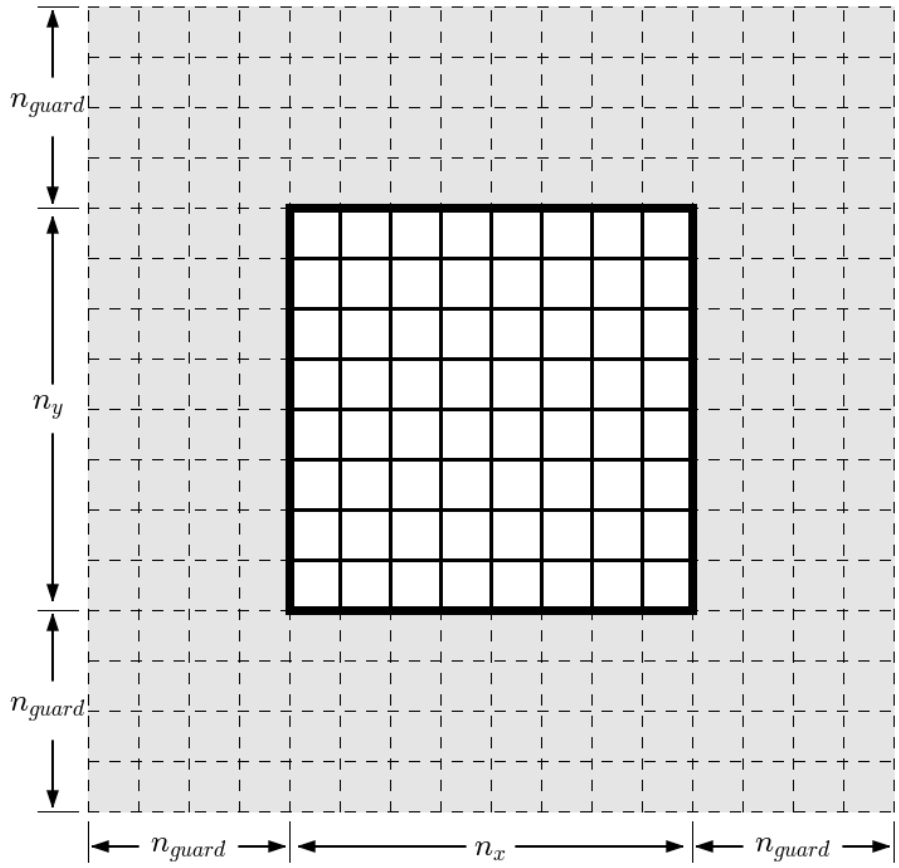

Figure 2.1: A single PARAMESH block in 2 dimensions. The number of cells along each coordinate, n_x and n_y, are the input parameters of the code and fully define the shape of all blocks. Around the central cells of the block is a layer of guard cells, n_{guard} cells thick. The guard cell layer is used to provide the boundary conditions and holds the state of the fluid in neighbouring blocks.

other hand, in the *block-structured AMR*, groups of mesh cells are refined and de-refined together (e.g. Berger & Colella, 1989; Bryan & Norman, 1997, 1998; Fryxell et al., 2000). The FLASH hydrocode uses a block-structured AMR mesh provided by the PARAMESH package, which we discuss in the following section.

2.2.3 PARAMESH: the AMR technique used in FLASH

The FLASH code uses the PARAMESH AMR module (MacNeice et al., 2000) to construct the computational mesh. PARAMESH is a block-structured AMR scheme in which all

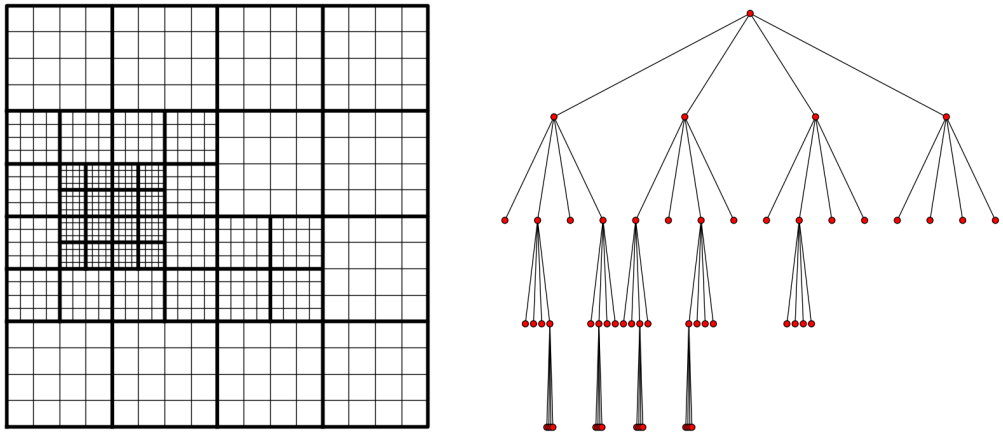

Figure 2.2: An example AMR 2D grid (left) and its corresponding quadtree data structure (right). The grid shows only the leaf blocks, nodes on the quadtree that have no child nodes and where all the fluid data is stored. The distance of the block from the root node on the quadtree gives its refinement level (e.g. the highest level of refinement in this grid is 5). Each block has $n_x = n_y = 4$.

the blocks have the same shape. This is the main difference from the original AMR implementation (Berger & Colella, 1989) where the blocks are allowed to have varying shapes. A PARAMESH block consists of $n_x \times n_y \times n_z$ central cells, where n_x, n_y and n_z are integers that give the number of cells along every spatial coordinate. Around the central cells is a layer of *guard cells*, n_{guard} cells thick, which are used as boundary conditions for the central cells and are set by the fluid states in neighbouring blocks. The required number of guard cells, n_{guard}, is given by the interpolation scheme used in specific physics modules (e.g. PPM hydro requires $n_{guard} = 4$). An example of a single block in 2D is given in Fig. 2.1.

The full mesh is constructed hierarchically, starting from a few blocks that represent the whole simulation volume. A block is refined by being split in half along each coordinate to form new blocks (4 or 8 new blocks in 2D and 3D, respectively) that cover the same volume as the original block. Each block is a node of the tree data structure (quad or octree tree in 2D or 3D, respectively), where the newly spawned *child* blocks are connected to the block from which they are spawned, the *parent* block. The data values in the

new blocks are interpolated from the values in the parent block in a so-called *prolongation* procedure. The opposite step, *restriction*, is used when derefining a region. The distance of each block from the root node on the tree graph gives its *refinement level*. The solvers operate only on *leaf* blocks, blocks that do not have any children. In PARAMESH, neighbouring leaf block cannot differ by more than one refinement level. Figure 2.2 shows an example AMR grid in two dimensions and its corresponding quad tree structure.

The FLASH code was written for use on distributed memory machines. In order for the code to run optimally on such machines, all the computational work must be evenly distributed among all the used processors. The domain decomposition in FLASH is also based on blocks. A Morton space filling curve (Morton, 1966) is used to map the multi-dimensional distribution of blocks to a one dimensional array of indices. The key property of space filling curves is that spatially close regions will also be placed close on the 1D mapping on average. Once the workload of each block is estimated, the 1D mapping can be chopped in pieces with approximately the same work balance and distributed among processors. Since the blocks that are neighbours on the space filling curve are also spatial neighbours on average, the inter-processor communication is minimised. The idea of using the Morton space-filling curve for domain decomposition was first proposed by Warren & Salmon (1993). It is not the only space-filling fractal that can be used for load balancing, e.g. the Gadget code (Springel, 2005a) uses a Peano-Hilbert curve for the same purpose.

2.2.4 Hydrodynamics solver

In Cartesian coordinates, the system of Euler equations, Eqs. (2.1) - (2.3), can be written in a general form as:

$$\frac{\partial \vec{U}}{\partial t} + \frac{\partial \vec{F^{\alpha}}(\vec{U})}{\partial x^{\alpha}} = 0, \tag{2.4}$$

where $\vec{U} \equiv (\rho, \rho\vec{v}, E)$ is the *state vector* of the fluid and $\vec{F^{\alpha}}$ are the corresponding fluid *fluxes* along each axis. Another transformation gives:

$$\frac{\partial \vec{U}}{\partial t} + \mathcal{A}^{\alpha}\frac{\partial \vec{U}}{\partial x^{\alpha}} = 0, \tag{2.5}$$

where we have introduced the Jacobian matrices:

$$\mathcal{A}^{\alpha} = \frac{\partial \vec{F}^{\alpha}(\vec{U})}{\partial \vec{U}}. \tag{2.6}$$

The system of Euler equations is *hyperbolic*, since the Jacobians in Eq. (2.6) have real and distinct eigenvalues (called characteristic velocities) and the corresponding eigenvectors form a complete set in R^3.

Godunov (Godunov, 1959) proposed a way to numerically solve Eq. (2.4) in the Eulerian paradigm. On finite volume cells, the fluid variables are cell-centred quantities, given at time t_n, the beginning of the time step. The evolution of U in these discrete steps is found by averaging Eq. (2.4) over the volume of each cell and the time step, $\Delta t = t_{\mathrm{n+1}} - t_{\mathrm{n}}$ (e.g. Castor, 2004):

$$V \frac{\Delta \langle \vec{U} \rangle}{\Delta t} = - \int \int_S \vec{F}^{\alpha} \cdot \vec{n} \mathrm{d}A, \tag{2.7}$$

where $\langle \vec{U} \rangle$ is the cell-averaged state vector and $\vec{F}^{\alpha} \cdot \vec{n}$ is the flux of fluid quantities through a cell edge, averaged over the time step. If we can obtain all fluxes from the value of $\langle \vec{U} \rangle$ at t_n, we automatically get its value at $t_{\mathrm{n+1}}$.

Since the fluid variables are assumed to be constant over each cell, the edges between cells will mark a discontinuity in U. This discontinuity defines the initial conditions for a *Riemann problem*, a generalisation of the shock tube problem (e.g. Sod, 1978) with non-zero initial fluid velocities. The essence of the Godunov method is calculating the fluxes in Eq. (2.7) by solving the evolution of the Riemann problem at every cell edge during the time step. Without going into the details of the numerical solution of the Riemann problem, we must point out that it is, in general, very computationally intensive. Still, much of the structure in the solution is lost due to the averaging of the fluid variables in discrete cells. This suggest that approximate solutions to the Riemann problem can be used in numerical hydrocodes to obtain satisfactory results. Here we must point out the largest benefit of basing the numerical technique on the Riemann solver: using a Riemann solution explicitly introduces the non-linearity in the difference equations of Eq. (2.4) and provides a tool for precise resolution of the sharp shock fronts that may arise in the flow. This is a very important feature for hydrocodes used to model astrophysical fluid flows, where the shocks are a rule rather than the exception.

The FLASH hydro solver uses an extension to the Godunov method, the *piecewise parabolic method* (hereafter, PPM). The method was described in detail by Woodward & Colella (1984) and Colella & Woodward (1984). The PPM method combines the values of the cell and its two neighbours on each side to estimate cubical polynomials used to

extrapolate the variables to cell boundaries. When integrated over space, these polynomials are required to reproduce the average state in the cells and may be modified to produce monotonicity and contact discontinuity steepening. The PPM algorithm is very well suited for modeling both smooth flows and flow discontinuities. For more details on the method and its implementation in FLASH see Colella & Woodward (1984) and Fryxell et al. (2000), respectively.

2.2.5 The FLASH code architecture and other physics modules

The FLASH code is designed to be highly modular. It is easy for the user to choose which physics modules to use for any given problem or to write and include new ones. A python script is used to put together the source codes of all the chosen physics modules and resolve any inconsistencies in the setup.

Many physical modules are provided in the default FLASH distribution. We mentioned only the two most basic ones: PARAMESH for the grid construction and PPM for solving the hydrodynamics equations. There are modules for different materials equations of state, for nuclear burning, simple gravitational potentials or self-gravity and many more. In the next section, we will discuss a radiative transfer technique written specifically for FLASH that we use in our work.

2.3 Hybrid Characteristics: RT for FLASH

This section discusses the Hybrid characteristics scheme (Rijkhorst, 2005; Rijkhorst et al., 2006) for modeling the transport of ionizing radiation over the PARAMESH grids. The ray tracing in HC borrows the best features from both long and short characteristics schemes (discussed in Section 2.3.1), resulting in a precise and fully parallel scheme, described briefly in Section 2.3.2. The chemistry equations that describe the interaction of the gas with incoming ionizing radiation are presented in Section 2.3.3.

2.3.1 Long versus short characteristics ray-tracing

Let us examine the propagation of ionizing radiation through a static hydrogen-only medium. Along a single line of sight, or ray, the general partial differential form of the radiative transfer equation (RTE), Eq. (1.16), transforms to an ordinary differential

equation, Eq. (1.17), making the ray a *characteristic* of the general RTE. If the emissivity, j_ν can be ignored, e.g. in case the radiation field is dominated by a point source, the intensity along the ray is simply obtained by integrating Eq. (1.17):

$$I_\nu(r) = I_\nu(0)e^{-\tau_\nu(r)}, \tag{2.8}$$

where r is the distance from the source and we introduced τ_ν, the optical depth at frequency ν:

$$\tau_\nu(r) = \sigma_\nu N_{\rm HI}(r), \tag{2.9}$$

where σ_ν is the hydrogen photo-ionization cross-section and $N_{\rm HI}$ is the column density of neutral hydrogen between the source and r. From Eqs. (2.8) and (2.9), it is clear that the intensity of radiation at any point can be obtained if the column number density up to that point is known. The column density of neutral hydrogen is:

$$N_{\rm HI}(r) = \int_0^r n_{\rm HI}(r'){\rm d}r', \tag{2.10}$$

which is just the integral of the HI number density field along the line of sight. In numerical simulations, the density field is represented by discrete value cells. Equation (2.10) is then discretised as:

$$N_{\rm HI}(r) = \sum_{\rm i=1}^{N_{\rm cut}} n_{\rm HI,i}\Delta l_i, \tag{2.11}$$

where $N_{\rm cut}$ is the number of cells cut by the ray up to r, $n_{\rm HI,i}$ is the neutral hydrogen number density in the i-th cell along the path and Δl_i is the length of the path section going through that cell. The sum of all sections Δl_i is the distance r, $r = \sum_{\rm i=1}^{N_{\rm cut}} \Delta l_i$.

In a computer model, Eq. (2.11) is solved through *ray tracing*, an umbrella term for algorithms used to compute the paths of particles or waves through interacting media. In RT, the most straightforward ray tracing technique is the *long characteristics*[1] scheme, where a single line is drawn between the source and the target and the intersection of the line with cells is computed to obtain Δl_i for each cell on the intersected list. An example is given in the left panel of Fig. 2.3. While this scheme is very precise, it is also rather inefficient, as the cells close to the source are crossed by rays many times.

[1]Characteristics is a historical name for ray-tracing methods used in numerical RT field, due to the connection with the method of characteristics technique for solving PDE. While it is an appropriate term for the long characteristics method, where every ray is a characteristic, it is purely descriptive for more complex ray tracing schemes, such as short characteristics.

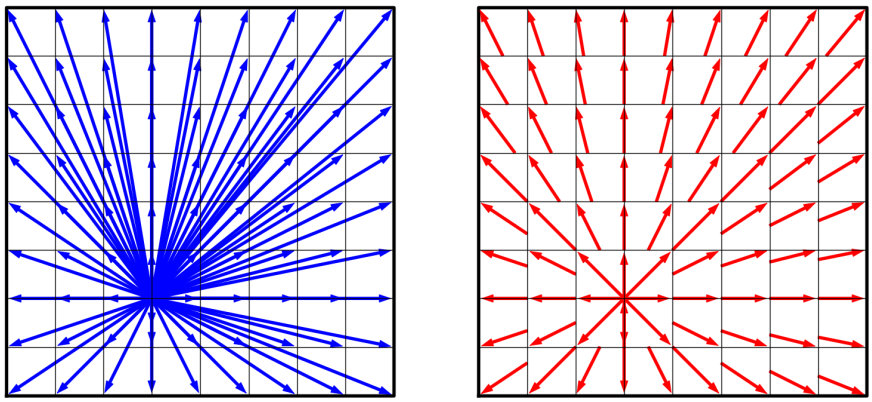

Figure 2.3: Examples of two approaches to ray tracing. *Left*: the long characteristics method. The rays are simply drawn from the source to every cell on the mesh. While this method is precise, it also sports many redundant calculations since the cells near the source are crossed many times. *Right*: the short characteristics method. Each cell is crossed only once. The column density at each cell uses the interpolated value of the upwind cells. This scheme is less precise, but more computationally efficient.

Their contribution to the column density is calculated for every ray, resulting in many redundant calculations.

This calculation redundancy is removed in the *short characteristics* scheme. There, every cell is traversed only once and the total column density in each cell is computed by interpolating the values from upwind cells. The interpolation technique necessarily introduces some numerical diffusion, but the number of operations is significantly smaller than in the long characteristics method. This is demonstrated in the right panel of Fig. 2.3. The main drawback is that the grid must be traversed in a specific order, since upwind values must be available for each cell. This, in turn, complicates the parallelization of the algorithm.

2.3.2 Tracing rays over the FLASH AMR grid

We discussed previous ray tracing schemes with the implicit assumption that the computational grid is uniform and fully stored in the available computer memory. This is

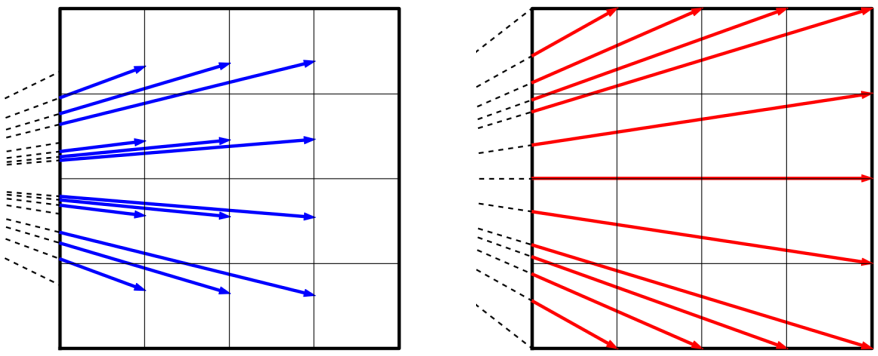

Figure 2.4: Local ray tracing in each block. *Left*: Local contribution to the ray that end at the cells in this block. *Right*: Values at the faces of cells away from the source on the opposite side of the block. The contribution of the block to every ray crossing it is computed by interpolation from these face values.

not true for the grid in FLASH, as discussed in Section 2.2.3. The Hybrid characteristics scheme was designed to perform efficient ray tracing over the PARAMESH distributed AMR grid. As the name might suggest, HC borrows the best features from both long and short characteristics methods: the long characteristics paradigm is employed to ensure a fully parallel algorithm and high precision, while the short characteristics idea of interpolation is used to increase the efficiency of the calculation and minimise memory use and inter-processor communication.

If we use the long characteristics scheme, a single ray will, in general, traverse more than one AMR block. Since computing the contribution of cells in one block to $N_{\rm HI}$ along the ray is independent from the contributions of others (i.e. Eq. (2.11) can be summed in any order), the tracing of each *ray segment* in a block can be done in parallel. To obtain $N_{\rm HI}$ along the whole ray, the individual contributions from all ray segments must be summed. The rays can be assembled one at a time, but then only the blocks that are crossed by the ray do any work which results in many wasted CPU cycles. Alternatively, every block can compute the contribution along every segment of every ray that crosses it. This again leads to much more work being done in blocks close to the source, in turn complicating load balancing. To avoid this issue, the HC scheme employs the interpo-

lation idea of short characteristics, but on the block level - every block only computes the ray segments leading from the source to the cell edges on the opposite side of the block. The contribution of the block to each ray is then obtained by interpolating these *face values* to the desired ray section length. Note that in all instances of ray tracing, the fast voxel traversal algorithm of Amanatides & Woo (1987) is used to efficiently find the list of traversed cells or blocks and the lengths of individual ray segments in each.

The algorithm for computing the column densities to every cell from a single source is then as follows:

1. Compute the *local* contribution to $N_{\rm HI}$ for every block. These are the "ends" of each full ray, computed by drawing a ray segment from the block edge closest to the source, along the ray direction and to the edge of every cell in the block. An example is given in the left panel of Fig. 2.4. This is a fully parallel, well load-balanced process.

2. Compute the *face values*: local column densities to the block edges opposite the source, used to calculate the contribution of this block to each ray traversing it. These are computed by drawing block ray segments from the source to the corners that touch the opposite side of the block of each block cell furthest away from the source. Right panel of Fig. 2.4 gives an example. Similar to the previous step, this is a local calculation, thus fully parallel.

3. Exchange the face values between processors. After a global communication (a "gather" operation), every processor ends up with all the face values from all the blocks. This data is used to construct full rays in parallel.

4. On every processor, construct a *global block map*, an array of integers that gives the global spatial distribution of blocks. In practise, this is a uniform grid with cell size equal to the smallest block. Each cell stores the unique index of the block that covers its spatial position. For example, in 3D, a block at highest refinement level, $RL_{\rm max}$, will be represented by a single cell on the global block map, while a block at $RL_{\rm max}-1$ will be represented by 8 cells and so on. Each processor first creates its local block map, with the global map obtained through the reduction of all local ones. The global block map is used to obtain the *list of blocks* that are cut by each ray.

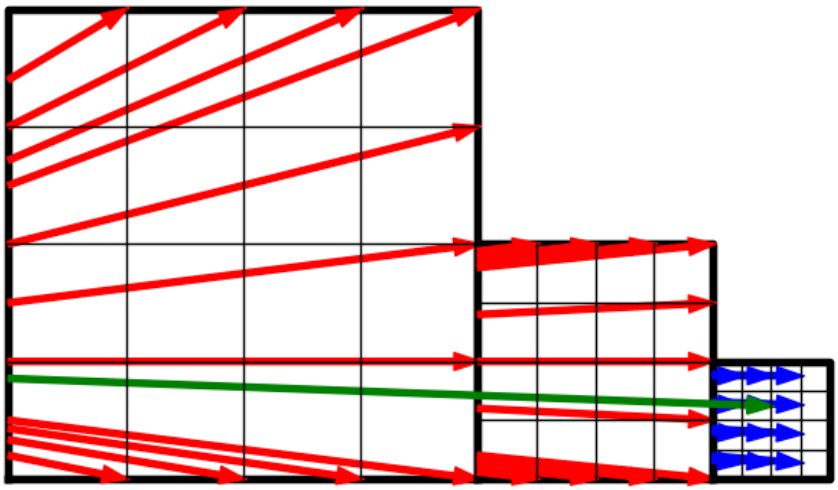

Figure 2.5: A single ray (green arrow) constructed by traversing blocks with varying level of refinement. The contribution of each ray segment is obtained by interpolation from face values at that block (red arrows), until the last block, where the column density to the cell itself is added (blue arrow). In the HC schemes, the traversal of the AMR grid is handled with ease, due to all operations in the ray being local to blocks.

5. Assembly of the whole ray. Once all the face values and the global block map have been exchanged, the contribution of ray segments can be summed to obtain $N_{\rm HI}$ along the whole ray. In this step, every ray is traced again, this time through the global block map, to obtain the list of intersected blocks. The contribution of each block to $N_{\rm HI}$ is obtained by linear interpolation from the face values closest to the intersecting ray segment. Finally, the local contribution to $N_{\rm HI}$, computed in step 1, is added to the ray. An example of a single ray construction is shown in Fig. 2.5.

The final product of the procedure is the column density between the source and the face of every cell at position $\vec{r}$, $N_{\rm HI}(\vec{r})$, which goes into Eq. (2.8) to obtain the intensity of radiation from the source at each cell. The HC scheme provides high precision, comparable to that of the long characteristics method, yet the number of operations is closer to short characteristics. It also exploits the full resolution provided by the AMR grid, and is well balanced for parallel execution, with only a few global communications per each step. For more details about the scheme, we refer the reader to the original paper (Rijkhorst et al., 2006).

2.3.3 Ionization of hydrogen gas

Currently, we use HC to model the transport of ionizing radiation through a medium consisting of hydrogen only. Ionizations may influence the fluid flow by heating the gas, as well as by changing the material composition, thus affecting the equation of state. In hydrogen-only gas, the ionization rate depends on the radiation field as (Osterbrock, 1989):

$$\Gamma = \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J_\nu}{h\nu} \sigma_\nu {\rm d}\nu, \tag{2.12}$$

where J_ν is the mean radiation intensity at frequency ν, σ_ν is the hydrogen photo-ionization cross section at frequency ν and $\nu_{\rm th}$ is the Lyman limit frequency giving the minimum photon energy necessary to ionize hydrogen, $h\nu_{\rm th} = 13.59\,{\rm eV}$. For example, the mean intensity J_ν from an isotropically radiating point source is:

$$4\pi J_\nu = \frac{L_\nu}{4\pi r^2} e^{-\tau_\nu(r)}, \tag{2.13}$$

where L_ν is the luminosity of the source and the $(4\pi r^2)^{-1}$ term represents the geometrical dilution of a spherical radiation field. The optical depth τ_ν to every cell is calculated using the ray tracing procedure described in Section 2.3.2.

The photo-ionization cross-section for hydrogen is given by (Osterbrock, 1989):

$$\sigma_\nu = \sigma_{\nu_{\rm th}} \left(\frac{\nu_0}{\nu}\right)^4 \frac{e^{4-(4/\tan(\epsilon)/\epsilon)}}{1-e^{-2\pi/\epsilon}}, \tag{2.14}$$

for $\nu \geq \nu_{\rm th}$. Here, $\sigma_{\nu_{\rm th}} = 6.3 \times 10^{-18}{\rm cm}^2$ is the cross-section at frequency $\nu_{\rm th}$ and $\epsilon = \sqrt{\nu/\nu_{\rm th} - 1}$. Throughout this thesis, we use an appropriate approximation to Eq. (2.14):

$$\sigma_\nu = \sigma_{\nu_{\rm th}} \left(\frac{\nu}{\nu_{\rm th}}\right)^{-3}. \tag{2.15}$$

Note that we do not follow the radiation at different frequency bins to obtain τ_ν. For a known source spectrum, Eq. (2.15) gives a direct relation between $\tau_{\nu_{\rm th}}$ and the optical depths for any other frequency. This allows us to tabulate the ionization rate from Eq. (2.12) as a function of $\tau_{\nu_{\rm th}}$ which is the only optical depth we compute through ray tracing. This approach is both faster and more memory efficient than explicitly following different frequency bins, while at the same time giving the exact solution.

If the gas has sufficiently high temperature or velocity, collisions between energetic particles can also cause ionizations. The collisional photo-ionization rate we employ is

(Theuns et al., 1998):

$$\Gamma_{\rm coll} = 1.17 \times 10^{-10} T^{0.5} e^{-157809.1/T} \left(1 + \left(\frac{T}{10^5 {\rm K}}\right)^{0.5}\right)^{-1}. \tag{2.16}$$

Free electrons can recombine with the hydrogen ions, forming neutral atoms. If the free electron is captured directly to the ground state, a new ionizing photon is produced. This in turn means that the material away from the source can have non-zero ionizing emissivity. In order to simplify the calculation, we assume the on-the-spot approximation (Spitzer, 1978; Mihalas & Mihalas, 1984; Osterbrock, 1989), i.e. all the ionizing photons produced by recombinations to the ground state are assumed to be absorbed locally. This approximation is implemented with a modified recombination rate, called case-B recombination rate:

$$\alpha_B(T) = \alpha_B(10^4 {\rm K}) \left(\frac{T}{10^4 {\rm K}}\right)^{-0.7}. \tag{2.17}$$

Here, the case-B recombination rate at $T = 10^4\,{\rm K}$ is $\alpha_B(10^4{\rm K}) = 2.57 \times 10^{-13} {\rm cm}^3/{\rm s}$.

Putting together the processes from Eqs. (2.12) - (2.17) leads to the rate equation for neutral hydrogen number density:

$$\frac{{\rm d}n_{\rm HI}}{{\rm d}t} = \alpha_B(T) n_{\rm HII} n_{\rm e} - \Gamma n_{\rm HI} - \Gamma_{\rm coll} n_{\rm HI} n_{\rm e}. \tag{2.18}$$

Here, $n_{\rm HI}$ and $n_{\rm HII}$ are the number densities of neutral and ionized hydrogen respectively, and $n_{\rm e}$ is the number density of free electrons.

The rate of change of thermal energy per unit volume is given by:

$$\rho \frac{{\rm d}u}{{\rm d}t} = n_{\rm H}^2 \left(\mathcal{H}(n_{\rm H}, n_{\rm HI}) - \mathcal{C}(T, n_H, n_{\rm HI}, n_{\rm HII}, n_e)\right), \tag{2.19}$$

where ρ is the mass density of hydrogen, $n_{\rm H} = \rho/m_{\rm u}$ is the number density of hydrogen and m_u is the atomic mass constant (very close to proton mass, $m_u \approx m_p$) and $\mathcal{H}$ and $\mathcal{C}$ are the heating and cooling rates, respectively. The heating rate includes only the contribution from photo-ionization heating, $\mathcal{H} = n_{\rm HI}\varepsilon/n_{\rm H}^2$, where the rate is given by:

$$\varepsilon = \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J_\nu}{h\nu} h(\nu - \nu_{\rm th}) \sigma_\nu {\rm d}\nu. \tag{2.20}$$

Effectively, every ionization inputs $h(\nu - \nu_{\rm th})$ energy towards heating the medium. For the cooling rate, $\mathcal{C}$, we use a sum of collisional ionization, collisional excitation, recombinations and Bremsstrahlung cooling rates (see Table B2 from Theuns et al., 1998). We assume the ideal gas equation of state: the total internal energy in the gas is $U = 3/2NkT$,

where $k = 1.38 \times 10^{-16}\,\mathrm{erg K^{-1}}$ is the Boltzmann gas constant and N is the number of particles. Noticing that $N = \frac{m}{\mu m_u}$, where m is the mass of all particles, and μ is the mean molecular weight, we get the connection between internal energy per unit volume and temperature:

$$u = \frac{3}{2}\frac{k}{\mu} n_H T. \tag{2.21}$$

The pressure of the gas is:

$$p = \frac{k}{\mu} n_H T. \tag{2.22}$$

The mean molecular weight depends on the ionization state of the gas as $\mu = (1 + x)^{-1}$, where $x = n_{\mathrm{HII}}/n_{\mathrm{H}}$ is the ionized hydrogen fraction. When the material is completely neutral, $\mu = 1$, meaning there is only one type of particle in the medium, while when it is completely ionized, $\mu = 1/2$.

In every time step, the system of equations (2.18) and (2.19), called *chemistry equations*, are solved in every cell to obtain the full evolution of the ionization and temperature fields. The chemistry equations are stiff[2] and special care must be taken to integrate them. We discuss this issue in more detail in Section 2.5.

We conclude the HC scheme discussion with a list of major improvements we introduced to the scheme:

1. Changed the chemistry solver to make the code photon conserving (Abel et al., 1999; Mellema et al., 2006). This change allows for much longer time steps to be used in the RT calculations, making the scheme orders of magnitude faster. See Section 2.5.

2. Adapted the fast voxel traversal ray-tracing scheme used by HC (Amanatides & Woo, 1987; Rijkhorst, 2005) to allow a source to be placed anywhere. In the original implementation, the source had to be placed in a cell corner. Furthermore, the original implementation suffered from round-off issues that produced incorrect results even when the source is placed correctly. Our implementation allows sources to be placed outside of the box, a feature used in one of the radiative hydrodynamics tests (Section 2.6.3).

[2]Stiff differential equations are broadly defined as equations for which some types of numerical solutions are unstable unless the integration step is taken to be very short. This is due to the very different scales of equation terms.

3. Fixed bugs that were making the accurate calculation of temperature impossible. Due to these issues, HC scheme did not perform Test 2 in the first RT code comparison paper (Iliev et al., 2006b). The correct temperature calculation is crucial for radiative hydrodynamics (Section 2.6).

2.4 Ionization front basics

For the rest of this chapter we will focus on a commonly encountered problem in astrophysics: propagation of ionization fronts. Ionization fronts (of I-fronts for short) occur when ionizing radiation emitted by an energetic source (e.g. an O star) propagates through the cold neutral medium (e.g. interstellar medium) and produces an HII region. The I-front is an interface, only a few photon mean free paths thick, that marks the boundary between ionized and neutral gas as well as a discontinuous jump in temperature and pressure (Mihalas & Mihalas, 1984). As a first approximation, we assume an infinitely sharp I-front (i.e. its thickness is zero, the source-side gas to be completely ionized and the gas on the other side to be completely neutral) propagating through a static hydrogen-only medium, leading to the I-front "jump condition":

$$u_{\rm I} = \frac{F}{n_{\rm H}}. \tag{2.23}$$

Here, $u_{\rm I}$ is the speed of the I-front, F is the flux of ionizing radiation through the front interface and $n_{\rm H}$ is the hydrogen number density. Equation (2.23) states that the I-front speed is given by the rate at which newly ionized atoms pass through the front interface.

If the front is expanding around an isotropically-radiating point source that produces $\dot{N}_\gamma$ ionizing photons per unit time, the flux F at the distance r from the source is:

$$F(r) = \frac{1}{4\pi r^2}\left[\dot{N}_\gamma - \frac{4\pi}{3}r^3\alpha_B(T)n_H^2\right]. \tag{2.24}$$

Here, the flux is obtained by decreasing the number of photons that arrive from the source to distance r by the number of recombinations in the sphere of that radius, where we used the case-B recombination rate, introduced in Eq. (2.17). The position of the front as a function of time evolves as (Spitzer, 1978):

$$r_I = r_S\left(1 - e^{-t/t_{\rm rec}}\right)^{1/3}, \tag{2.25}$$

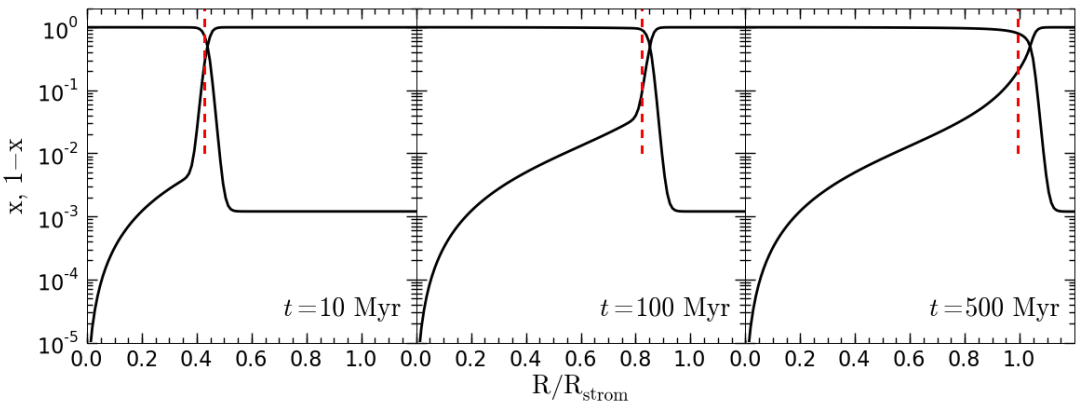

Figure 2.6: The ionized (x) and neutral ($1-x$) fractions as a function of distance to the source in Test 1 of the RT Code Comparison (Iliev et al., 2006b) obtained with our final HC scheme (including all updates discussed in this chapter). The three panels give results at three different times: 10 Myr, 100 Myr and 500 Myr, from left to right respectively. The vertical red dashed line marks the position of the front from the analytic solution of Eq. (2.25). The HC I-front position (defined as x=0.5, the intersection of x and $1-x$ lines) is slightly in front of the analytic solution. This is expected as the HC front is not infinitely sharp. For the same reason, the HC I-front extends slightly beyond the Strömgren sphere, $R_{\rm strom}$.

where

$$t_{\rm rec} = \frac{1}{\alpha_B(T) n_H}, \tag{2.26}$$

is the recombination time and

$$r_S = \left[\frac{3\dot{N}_\gamma}{4\pi\alpha_B(T) n_{\rm H}^2} \right]^{1/3}, \tag{2.27}$$

is the "Strömgren" radius (Strömgren, 1939), which is reached when the number of ionizations is balanced by the number of recombinations.

In reality, an I-front always has a finite thickness. Fig. 2.6 shows the results of RT Code Comparison Test 1 (Iliev et al., 2006b) we obtained with our final version of HC, which includes all the updates that we discuss in this chapter. Test 1 involves an I-front expanding around a point source of monochromatic ionizing photons with frequency $\nu_{\rm th}$

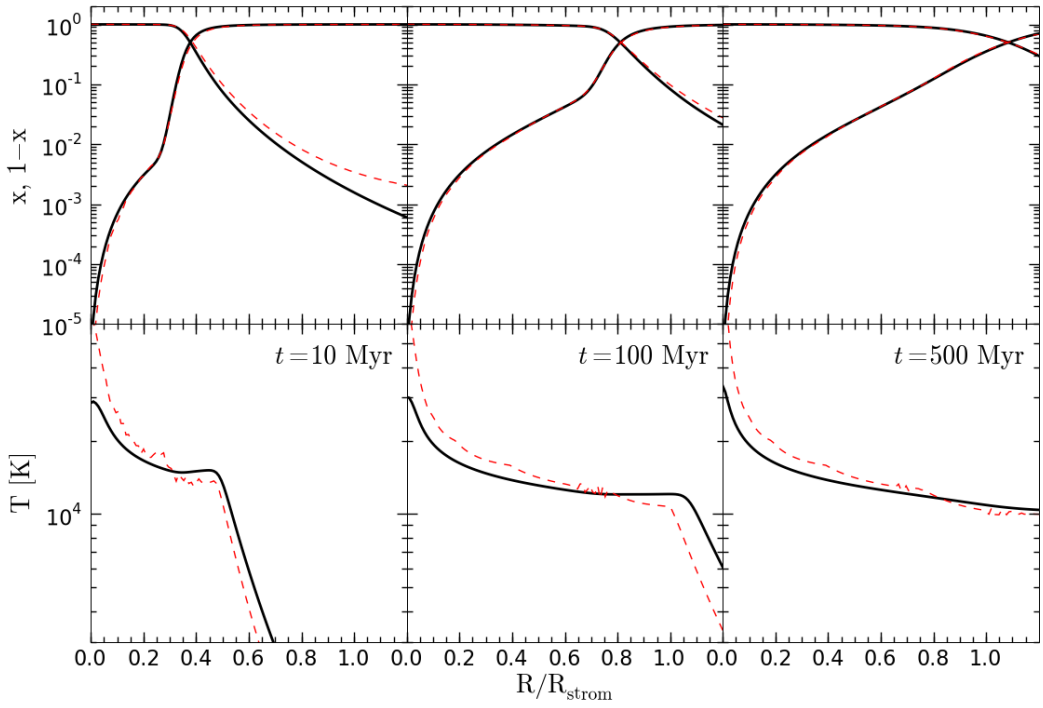

Figure 2.7: *Top*: Same as Fig.2.6 for the Test 2 results with our final HC scheme. *Bottom*: The temperature at the same distance from the source and at same times: 10, 100 and 500 Myr. The HC results are plotted with black solid lines. The C^2-ray results are plotted for reference (red dashed lines). In comparison to the results of Test 1 (Fig. 2.6), the I-front is significantly wider, due to the use of a harder source spectrum.

through a static (gas velocities $\vec{v} = 0$) uniform ($n_{\rm H}(\vec{r}) = {\rm const}$) density and constant temperature ($T(\vec{r}, t) = {\rm const}$) medium. The gas is initially completely neutral, i.e the ionized fraction $x = 0$ in the whole volume. While the I-front clearly has a finite thickness and the ionized fraction $x > 0$ always, Eq. (2.25) still provides a good estimate of the front's position when compared to the simulation data (where we defined the front position to be $x = 0.5$). The right panel of Fig. 2.6 depicts the distribution of x with distance from the source R inside the HII region, once the Strömgren radius, Eq. (2.27), is reached. The results of or HC code are in excellent agreement with other codes performing the same test in Iliev et al. (2006b), as well as with the analytic solution presented in Pawlik & Schaye (2008).

The constant temperature approximation is not good when the massive O and B stars

produce the ionizing radiation. These sources are approximately black body spectra, with non-negligible energy output at frequencies $\nu > \nu_{\rm th}$. Photo-ionization increases the thermal energy of the gas, as per Eq. (2.20). Thus, the presence of the I-front marks the boundary between regions with very different temperatures, as shown in Fig. 2.7. There, we show our HC scheme results of the Test 2 of the RT Code Comparison (see Iliev et al., 2006b, for exact numerical parameters). Test 2 has the same setup as Test 1, shown in Fig. 2.6, but now the source has a black body spectrum (with $T_{\rm source} = 10^5\,{\rm K}$) and the gas temperature changes due to photo-ionization heating and gas cooling, i.e. we solve Eq. (2.19) along with (2.18). Figure 2.7 shows the effect of multi-frequency transport on the shape of the HII region (bottom panels): the higher energy photons have longer mean free paths (smaller cross-sections) which allows them to penetrate further into the neutral medium, increasing the front's thickness. The higher energy photons are very effective at heating the gas, as demonstrated by the fact that the temperature field expands further than the I-front position. The recombination rate also depends on the temperature, but only weakly, as in Eq. (2.17). Therefore, this dependence does not significantly affect the shape of the front.

Finally, the static medium approximation does not universally hold. The change in temperature following photo-ionization inevitably leads to the departures from pressure equilibrium. At the I-front interface, the conservation equations defined by the Euler equations (2.1)-(2.3) must apply. As the I-front marks a discontinuity, we can describe it with the same "jump conditions" used for shock fronts (Mihalas & Mihalas, 1984; Castor, 2004). Assuming that the flow proceeds in one dimension, the first two Euler equations lead to:

$$\rho_1 u_1 = \rho_2 u_2 = m_u F, \tag{2.28}$$

$$p_1 + \rho_1 u_1^2 = p_2 + \rho_2 u_2^2. \tag{2.29}$$

Here, index 1 marks the fluid conditions in front of the shock, while 2 marks the same behind the shock. The energy equation jump conditions are not given since they are not used in the following text. The $m_u F$ term in Eq. (2.28) is an addition to the usual shock front jump conditions (e.g. Landau & Lifshitz, 1959), making clear that the front speed is not defined by the gas flow but by the radiation flux.

Historically, I-fronts have been classified according to their speed with respect to the gas and the change in gas density at the front interface (Kahn, 1954; Axford, 1961;

Spitzer, 1978). To derive the critical speeds, we must remember that the ionization fronts are not isothermal: due to the input of radiation energy, the sound speeds at each side of the front can be markedly different. The speed of sound in the ideal gas is given by $C = \sqrt{\gamma p/\rho}$. Using this definition (with $\gamma = 1$, Spitzer, 1978) and Eq. (2.28), we can obtain the general solution of Eq. (2.29):

$$\frac{\rho_2}{\rho_1} = \frac{(C_1^2 + u_1^2) \pm \sqrt{(C_1^2 + u_1^2)^2 - 4u_1^2 C_2^2}}{2C_2^2}. \tag{2.30}$$

The ratio of densities must be a real number, restricting the possible values u_1 can take. It may either exceed u_R, given by:

$$u_R = C_2 + \sqrt{C_2^2 - C_1^2} \approx 2C_2, \tag{2.31}$$

or it must be less than u_D, given by:

$$u_D = C_2 - \sqrt{C_2^2 - C_1^2} \approx \frac{C_1^2}{2C_2}. \tag{2.32}$$

If $u_I \geq u_R$, the I-front is called "R-type", where R refers to the rarefied gas, as u_I always exceeds u_R for sufficiently low ρ_1. Using the negative sign in Eq. (2.30) leads to a "weak R-type" front, where the relative change of density is small. This is the most commonly found R-type in astrophysical problems. The opposite, "strong R-type", obtained using the positive sign in Eq. (2.30), requires some special mechanism to maintain large front velocity in the face of large changes in density at the front. When $u_I \leq u_D$, the I-front is D-type, with D referring to the dense gas. With respect to the ionized gas, this type of I-front can be either subsonic, "weak D-type" obtained from the positive sign in (2.30) or supersonic, "strong D-type" from the negative sign in (2.30). Front velocities between two critical values $u_D < u_I < u_R$ are possible but, in that case, the I-front must be preceded by a shock that compresses the gas entering the front sufficiently to slow it down and guarantee it transitions to D-type. This evolution of I-fronts in full radiative hydrodynamics calculations will be seen in the tests we present in Section 2.6.

2.5 Photon conservation

Assume a radiation field around a point source of luminosity L_ν in a medium with the frequency-dependent optical depth τ_ν. Combining equations (2.12) and (2.13), the photo-

ionization rate is:

$$\Gamma_{\rm local}(r) = \frac{1}{4\pi r^2}\int_{\nu_{\rm th}}^{\infty}\frac{L_\nu}{h\nu}e^{-\tau_\nu(r)}{\rm d}\nu. \tag{2.33}$$

Equation (2.33), used in the original implementation of HC, is only accurate in a single point in space and a single moment in time. To numerically solve the relevant RT and chemistry equations, space and time must be discretised into finite-sized cells and time steps. With discrete cells, the use of Eq. (2.33) will provide an accurate estimate of Γ only if all the cells are optically thin and the time step is very short, e.g. a fraction of the light crossing time. To satisfy these conditions, the mesh resolution must be very high and the time step orders of magnitude shorter than the hydrodynamics CFL criterion (e.g. Bodenheimer et al., 2007). This makes the RT calculation very costly, if not entirely impossible. If these conditions are not satisfied, the RT scheme may violate the conservation of energy - this issue is called "photon conservation". The loss of photon conservation results in incorrect propagation speed of ionization fronts, as we will demonstrate shortly.

In this Section, we examine the roots of the photon non-conservation, found in numerical implementation of Eq. (2.33). We separate the discussion on the issues stemming from spatial (Section 2.5.1) and temporal (Section 2.5.2) discretization. This updated code is used for radiative hydrodynamics tests presented in Section 2.6.

2.5.1 Effect of spatial discretization

Assume that the density field is discretised in finite-sized cells of side length Δl. The photo-ionization rate computed with Eq. (2.33) is a good approximation for the whole cell only if the cell is optically thin, $\Delta\tau_\nu = \sigma_\nu n_{\rm HI}\Delta l \ll 1$. In this case, the photo-ionization rate is approximately constant through the cell. In contrast, when $\Delta\tau_\nu$ is not small, different regions in a given cell should experience different photo-ionization rates and Eq. (2.33) is no longer representative of the actual number of ionizations occurring in that cell.

Abel et al. (1999) proposed setting Γ directly to the analytic solution to RTE along a ray passing through that cell, given by Eq. (2.8). If $\dot{N}_{\nu,0}$ photons per second enter the cell with optical depth $\Delta\tau_\nu$, $\dot{N}_{\nu,0}e^{-\Delta\tau_\nu}$ photons are going to pass through it. All the $\dot{N}_{\nu,0}(1-e^{-\Delta\tau_\nu})$ that stay in the cell result in ionizations, giving the ionization rate. We adopted this idea in the form presented in Mellema et al. (2006).

Consider a spherical shell in the region with constant hydrogen density n_H with a point source of luminosity L_ν at its centre. Let r be the radius of the inner shell edge and $r+\Delta r$ the radius of the outer one, thus making Δr the thickness of the shell. The volume of the shell is:

$$V_{\rm shell} = \frac{4\pi}{3}\left((r+\Delta r)^3 - r^3\right). \tag{2.34}$$

The number of hydrogen atoms in the shell is $N_{\rm shell} = n_H V_{\rm shell}$. The total number of ionizations per unit time in the shell is $\dot{N}(r) - \dot{N}(r+\Delta r)$, where $\dot{N}(r)$ and $\dot{N}(r+\Delta r)$ are the number of photons that reach and leave the shell per unit time, respectively. The ionization rate is then given simply as:

$$\Gamma = \frac{\dot{N}(r) - \dot{N}(r+\Delta r)}{n_{\rm HI} V_{\rm shell}}. \tag{2.35}$$

Let τ_ν be the optical depth between the source and the inner shell edge and $\Delta\tau_\nu$ the optical depth within the shell. Then, we can rewrite Eq. (2.35) as:

$$\Gamma = \int_{\nu_{\rm th}}^{\infty} \frac{L_\nu e^{-\tau_\nu}}{h\nu} \frac{1-e^{-\Delta\tau_\nu}}{n_{\rm HI} V_{\rm shell}} {\rm d}\nu. \tag{2.36}$$

Putting $\Delta r = \Delta l$, Eq. (2.36) can be used to compute the ionization rate in a Cartesian cell.

In the limit $\Delta r \to 0$, for which $\Delta\tau_\nu \to 0$, one retrieves Eq. (2.33), backing our earlier claim that Eq. (2.33) is appropriate for use in optically thin cells. The left panel of Fig. 2.8 shows the dependence of the ionization rate on the optical depth of the cell, $\Delta\tau_\nu$. Using Eq. (2.33) in optically thick cells overestimates the ionization rate, which directly results in I-fronts moving too fast (right panel of Fig. 2.8). This means that, for a given problem, the results may depend on mesh resolution. Increasing the resolution improves agreement with the expected result (black crosses), but provided a sufficiently short time step is used. In general, a treatment for the temporal discretization, which will be discussed in the following section, is also needed.

2.5.2 Effect of temporal discretization

The ionization rate of Eq. (2.36) is only correct if $\tau_\nu \approx {\rm const}$ and $\Delta\tau_\nu \approx {\rm const}$ during the calculation time step, Δt. To guarantee this, RT schemes are usually forced to employ very short time steps e.g. a fraction of the light crossing time $\Delta l/c$ (Abel et al., 1999) or a few ionization times, Γ^{-1} (Bolton et al., 2004). Increasing the scheme computational

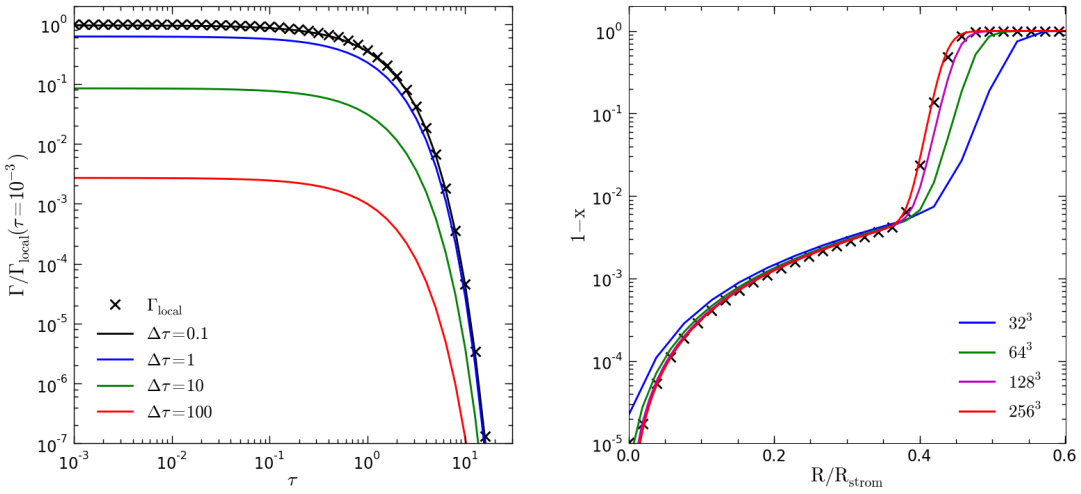

Figure 2.8: *Left:* the dependence of the photo-ionization rate, Γ computed with Eq.(2.36), on the optical depths τ and $\Delta\tau$. All values of Γ were computed assuming $r \gg \Delta r$ and a monochromatic source. The X symbols give $\Gamma_{\rm local}$ computed with Eq. (2.33). For easier comparison, all curves have been normalised to $\Gamma_{\rm local}(\tau = 10^{-3})$. The solid curves, from top to bottom, show the results of Eq. (2.36) with increasing cell optical depth, $\Delta\tau$. When $\Delta\tau$ is low (i.e. $\Delta\tau \ll 1$), the results of Eqs (2.36) reproduce the results of Eq. (2.33), but when $\Delta\tau$ increases into the optically thick regime ($\Delta\tau > 1$), Γ decreases. Effectively, Γ will be overestimated in optically thick cells if one employs Eq. (2.33), leading to too fast I-fronts as shown in the right panel. *Right:* The neutral fraction at t = 10 Myr in Test 1 on varying resolution computational grids using the original, non-photon conserving HC implementation. The calculation time step was chosen to minimise the temporal issues (discussed in Section 2.5.2). As the cell size is increased (by reducing the spatial resolution), the I-front speed increases: this is consistent with the prediction based on the left panel plot. The optical depths $\Delta\tau \approx 0.5, 1, 2, 4$ for grids with 256, 128, 64 and 32 cells along a box size respectively.

efficiency by taking longer time steps is possible if an estimate for optical depth evolution during Δt is included. This idea is the basis of the C^2-ray RT code (Mellema et al., 2006).

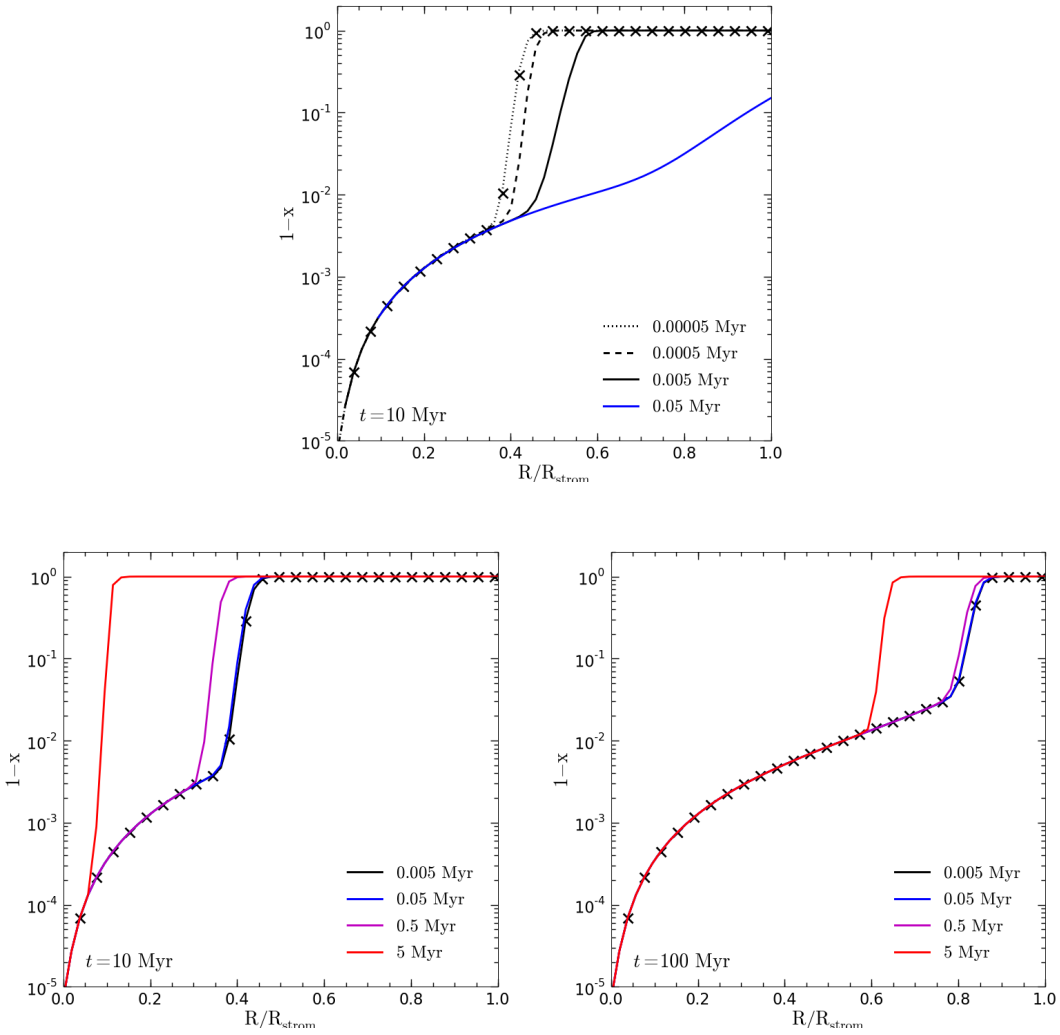

Figure 2.9: The effect of temporal photon conservation on the Test 1 results. Resolution is 64^3 cells and time step is varied by order of magnitude. *Top:* Γ, computed with Eq. (2.36), is constant during a time step, and as a consequence the speed of the I-front is overestimated for longer time steps. *Bottom left:* The same as in the top panel, but with time-dependant Γ in the cell only (only $\Delta\tau = f(t)$, see text). With this change the time steps may be much longer, yet very long time steps now result in the I-front moving too slowly, as evidenced by the lagging fronts at Δt = 0.5 Myr and 5 Myr. This effect is less important at later times, e.g. *bottom right* panel, when the front slows down due to the distance from the source.

Time-dependent τ in the RT cell

We will discuss the evolution of the two optical depths in Eq. (2.36) separately. The evolution of the optical depth in the cell, $\Delta\tau_\nu$, is easier to include in the calculation as it only depends on variables local to the cell. Given that $\Delta\tau_\nu = \sigma_\nu(1-x)n_H\Delta l$, $\Delta\tau_\nu$ can vary during a time step only when the cell ionized fraction x changes, since both the number density and cell size are constant during Δt. For a constant incoming flux, the correct evolution of x in a cell is obtained by integrating Eq. (2.18) with the ionization rate of Eq. (2.36) with $\Delta\tau_\nu = \Delta\tau_\nu\left(x(t)\right)$. The original HC implementation uses the DORIC solver (e.g. Frank & Mellema, 1994) to integrate the chemistry system of equations, Eqs. (2.18) and (2.20). This solver is built around the work of Schmidt-Voigt & Koeppen (1987) who showed that Eq. (2.18) can be solved analytically if Γ, n_e and T are assumed to be constant in time. The solution is:

$$x(t) = x_{\rm eq} + (x_0 - x_{\rm eq})e^{-t/t_i}, \tag{2.37}$$

where

$$t_i = \left(\Gamma + n_e\Gamma_{\rm coll} + n_e\alpha_B(T)\right)^{-1} \tag{2.38}$$

and

$$x_{\rm eq} = \frac{\Gamma + n_e\Gamma_{\rm coll}}{\Gamma + n_e\Gamma_{\rm coll} + n_e\alpha_B(T)}. \tag{2.39}$$

The general solution to the chemistry equations system can then be obtained by iterating on n_e. The assumption of constant Γ means we cannot use the DORIC package, because we want to have $\Gamma = \Gamma(t)$ explicitly. We therefore replaced the original set of solver routines with new ones based around a backward differentiation ODE solver provided in the LSODE (Hindmarsh, 1980), a package commonly used to solve stiff systems. This implementation allows for explicitly time-dependent ionization and heating rates. We present the effect of time-evolving $\Delta\tau_\nu$ in Fig. 2.9. All panels present the Test 1 results (see Section 2.4) of the HC method using Eq. (2.36) and varying the time step Δt. The top panel shows the results at $t = 10$ Myr with ionization rate being constant during a time step, while the bottom two show the results with Γ varying only due to the change in $\Delta\tau_\nu$ ($\tau_\nu = \text{const}$ during each time step). A constant ionization rate (top panel) clearly leads to wrong results for $\Delta t \gtrsim 5 \times 10^{-5}$ Myr. Once the time-dependent Γ is included, the scheme produces correct results for $\Delta t \lesssim 0.05$ Myr (bottom panels). For even longer

time steps, the I-front begins to lag behind: at such high Δt, more than one cell is crossed at a time, meaning τ_ν cannot be treated as a constant. This is further evidenced by the fact that this delay is more prominent at early times (bottom left panel), than at later times (bottom right panel) as the I-front speed decreases with increasing distance from the source (see Eq. (2.25)).

Time-dependent τ between the RT cell and the source

If an I-front is allowed to cross multiple cells in Δt, a cell that did not see any radiation flux, due to high τ_ν at the beginning of the step, may see a much more intense flux at some later time during Δt. If τ_ν is kept constant in this case, the flux will be constant and the I-fronts will propagate slower than expected as evidenced in Fig. 2.9 (bottom two panels). Therefore, for an RT scheme to be truly independent from the time step Δt, the change in τ_ν during a time step must be included.

The C^2-ray code (Mellema et al., 2006) is the only characteristics-based RT scheme that explicitly handles this issue. Their method uses the short characteristics paradigm because of the specific order in which the cells are crossed (see Section 2.3.1). In their scheme, they do not have the explicit time-dependence of optical depths. Instead, they use the time-averaged optical depths in each cell, $\langle \Delta\tau_\nu \rangle = \frac{1}{\Delta t} \int_{\Delta t} \Delta\tau_\nu(t) \mathrm{d}t$, when integrating the chemistry equations. To construct τ_ν in short characteristics schemes, all the cells between the cell and the source must be crossed in order. This means that $\langle \Delta\tau_\nu \rangle$ for all cells along the ray are known at the time when the optical depth in a given cell is to be evaluated. In that case, the time-averaged optical depth between the source and the cell is known: $\langle \tau_\nu \rangle = \sum_{\mathrm{crossed}} \langle \Delta\tau_\nu \rangle$, a sum of time-averaged optical depths of all cells crossed by the ray.

The HC ray-tracing scheme is much more similar to long characteristics, i.e. there is no specific order in which the cells are crossed. In addition, we found that using the time-averaged optical depths may result in incorrect temperature evolution for long time steps, a problem that may appear in C^2-ray as well (Garrelt Mellema, private communication). Therefore, we propose using the ionization front tracking method (hereafter IFT; e.g. Alvarez et al., 2005) for computing the change of τ_ν during a time step. In principle, the IFT method solves the jump condition presented in Eq. (2.23) for a (non-uniform) static density field. Along every ray, the incoming flux at the front interface is then given

by:

$$F(r) = \frac{1}{4\pi r^2}\left[\dot{N}_\gamma - \sum_{j=0}^{i-1} \alpha_B(T) n_{\mathrm{H,j}}^2 \Delta V_j, \right], \tag{2.40}$$

where the recombinations rate in a sphere with radius r, Eq. (2.24), is replaced by the sum of rates in spherical shells of density corresponding tho j-th cell cut by the ray and with volume $\Delta V_j = 4\pi(r_{j+1}^3 - r_j^3)/3$, where $r_{j+1} - r_j = \Delta r_j$ is the ray length in cell j.

To test the validity of using IFT, we wrote a 1D long characteristics code, with the exact same chemistry module as in HC. The IFT procedure is done as follows. At the beginning of every time step, we set the IFT front position so that both IFT and HC find the same column of ionized gas along the direction of the ray well beyond the position of both fronts. Once the position of the IFT front is set, we solve the jump condition Eq. (2.23) along the ray with the flux computed with Eq. (2.40). We tabulate positions of the IFT front in equal substeps during Δt. This table is used in the chemistry solver to provide the change in τ_ν since the beginning of the time step. Because the IFT front is sharp, its position fully defines the optical depth along the ray: at $t = 0$, $\tau = \sigma_{\nu_{\mathrm{th}}} n_{\mathrm{HI}} l$, where l is the distance between the front and the end of the ray, while at some later time $t = \delta t$, $\tau = \sigma_{\nu_{\mathrm{th}}} n_H (l - \delta l)$, where δl is the distance the front travelled in δt. Since the initial optical depth defines l, the change in τ obtained through raytracing is given by the change of the IFT front position, δl, at every time during the time step: $0 \leq \delta t \leq \Delta t$.

Figure 2.10 shows the results of Test 1, as in Fig. 2.9, but now with the evolution of τ_ν computed using IFT. While the IFT is not a perfect estimate, it does allow for solutions close to the correct one to be computed at even longer time steps. As we suggested before, an estimate of evolving τ_ν is crucial the I-front crosses a multiple of cells in one time step, for example when the front is close to the source. The importance of temporal photon conservation is clear also when the temperature evolution is being computed (i.e. if the energy Eq. (2.20) is solved, as shown in Figure 2.11). The reason for the extreme spikes in the temperature is the overestimated heating rate, itself a consequence of the constant optical depths. Only a few cells closest to the source are affected in this case and the temperature approaches the correct value for cells further away from the source. This purely numerical effect may be important in radiative hydrodynamics applications, where these temperature spikes will result in unphysical pressure forces. The problem is largely avoided when the time evolution of $\Delta\tau_\nu$ is introduced - the time evolution of τ_ν

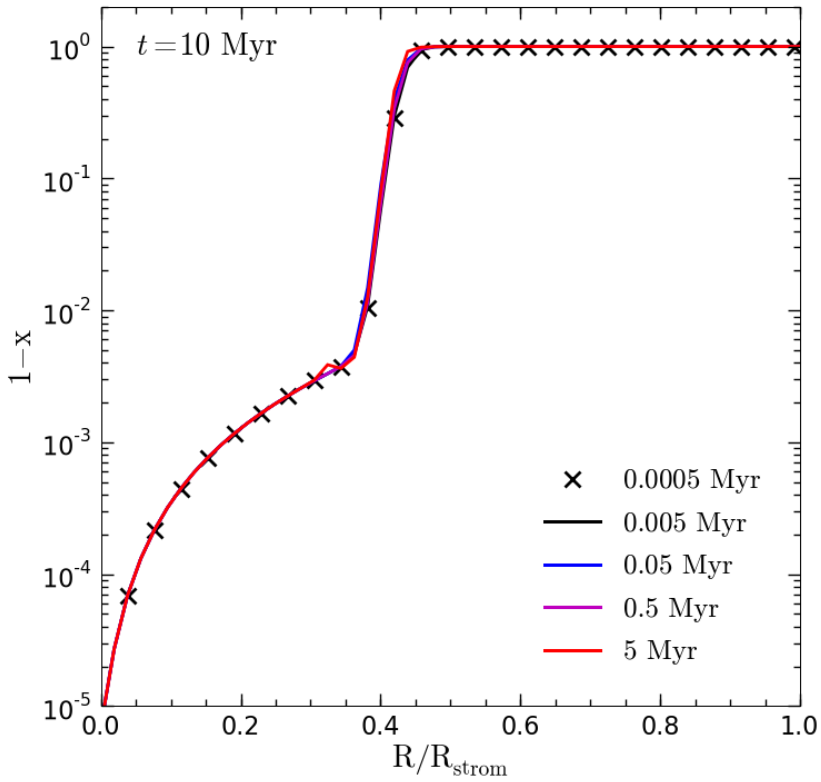

Figure 2.10: Test 1 with full photon conservation, using time-dependant Γ both in the cell and in front of it (see text). The results are well converged for all used time steps (better than temporal conservation in the cell-only case, Fig. 2.9).

is not crucial for two reasons: the effect of constant τ_ν is that the rates are actually lower than they should be ($e^{-\tau_\nu}$ term in Eq. (2.36)) and the temperature spikes are located in the cells nearest to the source.

The new chemistry solver we developed for HC uses Eq. (2.36) to achieve spatial photon conservation and implements the temporal photon conservation in each cell with time dependent $\Delta\tau_\nu$. The change of optical depth in front of the cell using the IFT solution is not yet implemented at this point. We have found it only to be important for very long time steps. In practise, especially for the tests we are to present in the next section, these time steps are longer than the hydrodynamics time step. We demonstrate this at the end of Section 2.6.1, where we show that our results are converged irrespective of

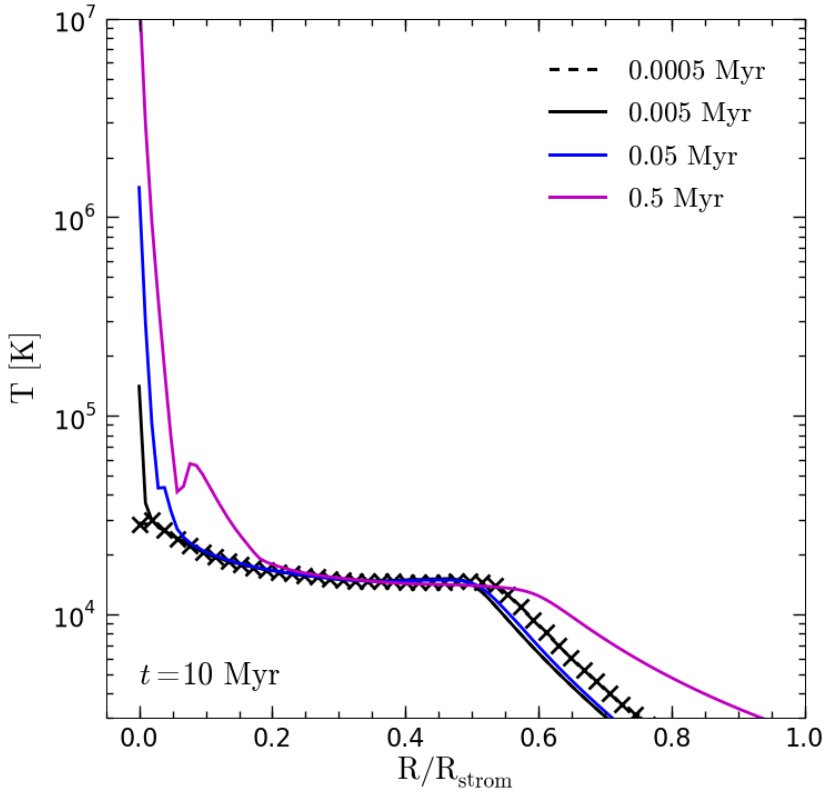

Figure 2.11: The importance of temporal photon conservation on the state of the temperature in Test 2. The black crosses give the result of the run with full temporal photon conservation (both in the cell and between the cell and the source), which the coloured lines do not include. The $\Delta t = 5 \times 10^{-4}$Myr case (dashed black lines) agrees with the full photon conservation case. The temperature is overestimated if the time step is too long. This effect may be important in radiative hydrodynamics applications, since the temperature around the source produces the pressure difference that drives the gas flow.

the chosen RT time step. To achieve this result without implementing the IFT solution, we use a simple trick: we choose the initial time step to produce the correct evolution in the early R-type stages and let it increase as dictated by the FLASH internal hydrodynamics criteria. At the later D-type stage, the I-front speed matches the gas flow speed, which guarantees that it does not cross more than one cell, i.e. the I-front travels at the

same speed as the gas. We note however that the short time step with which we start the simulations is 100 times longer than what would be necessary without any treatment of temporal photon conservation.

We have not performed detailed timing tests, because the improvements that the photon conserving chemistry solver we introduced are anything but subtle. The chemistry calculation itself is more computationally expensive, but the ability to take 5 orders of magnitude longer time steps more than makes up for it: for example, the original HC implementation had to compute Test 1 for several CPU days (while still not getting the exact solution due to not using Eq. (2.36)), where as we can do this computation in less than 1 CPU hour. We present the use of our improved HC scheme in the radiative hydrodynamics tests in the next section. As a final note, we discuss when the implementation of the time-dependent Γ is important.

Time-dependent ionization rate as a function of optical depth

In all the previously discussed tests, the RT cells are mildly optically thick ($\Delta\tau_\nu \approx 1-10$). Now we ask the question of how important the temporal photon conservation is in other optical thickness regimes. The main consequence of the time-evolving optical depth in a single cell is that it takes longer for that cell to become fully ionized, due to Γ decreasing with the increasing ionized fraction. Assume a single RT cell that is being ionized by a constant incoming mean intensity, i.e. τ_ν and L_ν are constant in Eq. (2.36). In that case, the photo-ionization rate in the cell will only depend on its own optical depth, $\Delta\tau_\nu$, which in turn depends only on the ionized fraction of the cell. Assuming a simple case of no recombinations, the rate equation (2.18) transforms to:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \Gamma(1-x). \tag{2.41}$$

Hence, if Γ is constant in time, $x = 1 - \exp(-\Gamma t)$. The time to fully ionize a cell is clearly infinite. However, for physical insight it is sufficient to see how long it takes for a cell to become highly ionized, e.g. up to $x = 0.99$, the time we label as T_{ion}.

We consider two versions of Eq. (2.41) that results in two different ionization times: one which assumes the photon-conserving time-evolving $\Gamma = \Gamma(x = f(t))$, for which we label the ionization time as T^t_{ion}, and one which assumes a Γ constant in time, set to the value at $x = 0$, $\Gamma = \Gamma(x = 0)$, labelled as $T^{\mathrm{const}}_{\mathrm{ion}}$. We solve Eq. (2.41) numerically in both

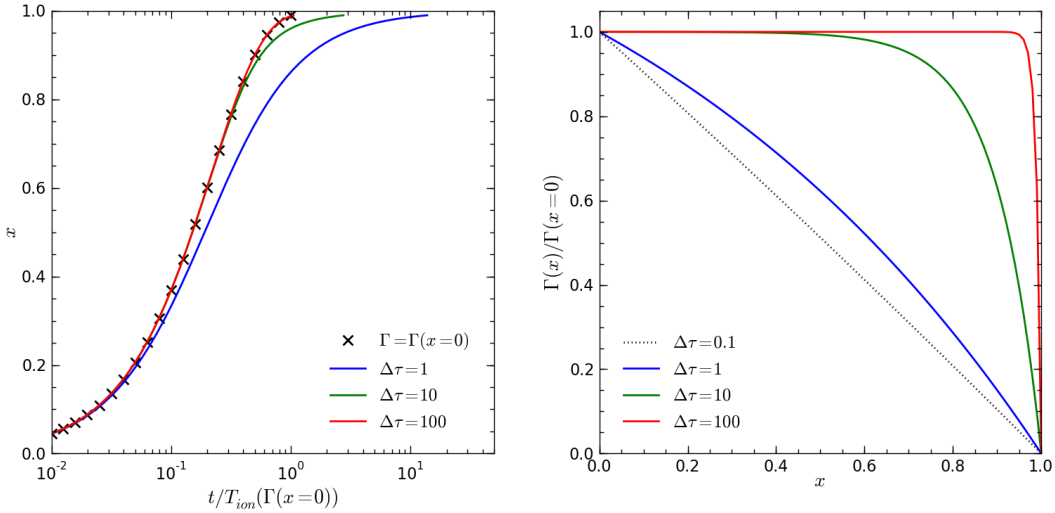

Figure 2.12: Importance of including time-dependent Γ in the evolution of x in a cell. *Left:* The evolution of the ionized fraction during cell ionization (ignoring recombinations) for cells with different initial optical depths of neutral cells, $\Delta\tau$. Time is quoted in units of the "ionization time" computed assuming Γ does not change with evolving x (i.e. it has the same value as when $x = 0$), $T_{\rm ion}(\Gamma = \Gamma(x = 0))$ (see text for elaboration). The black crosses give the evolution of x assuming said constant Γ. The solid lines give the evolution when Γ varies with $\Delta\tau$, which in turn depends on x. Line colors correspond to different initial optical depths: blue, green and red for $\Delta\tau$ = 1, 10 and 100, respectively. *Right:* Dependence of Γ on x. While in lower optical depth cells (e.g. dotted black and solid blue lines) Γ varies for all values of x, in optically thick cells (solid red line), it remains mostly constant. This is reflected in the time it takes to ionize the cell, as shown in the left panel.

cases, with the latter also tested against the analytic solution. We look at the evolution of x with time in these two cases for cells of various optical depths, $\Delta\tau_\nu$ in the left panel of Fig. 2.12. A larger optical depth of a cell necessarily means more neutral absorbers in the cell, which results in a longer ionization time. To remove this dependence, we quote the time in units of $T_{\rm ion}^{\rm const}$ for each $\Delta\tau_\nu$. This allows us to discuss the change of $T_{\rm ion}$ caused by the change of Γ in relative terms, i.e. how much longer does it take to ionize a cell when the photo-ionization rate is allowed to change due to the increase of the ionized

fraction. The black crosses show the evolution of x with constant Γ - it is independent from $\Delta\tau_\nu$ because $T_{\rm ion}^{\rm const}$ for each single $\Delta\tau_\nu$ is used as a unit of time in that run. The solid lines show the evolution of x when Γ changes with x. In that case, we see that $x = f(t)$ now depends on the cell initial optical depth $\Delta\tau_\nu$. The delay in achieving "complete" ionization ($x = 0.99$) increases with decreasing optical depth, e.g. when $\Delta\tau = 1$ (blue line), $T_{\rm ion}^{t} \approx 15 \times T_{\rm ion}^{\rm const}$ (remember that $T_{\rm ion}$ is the time it takes to reach $x = 0.99$), while for $\Delta\tau = 100$ (red line), $T_{\rm ion}^{t} \approx 1.1 \times T_{\rm ion}^{\rm const}$.

The reason for this behaviour can be seen in the right panel of Fig. 2.12, where we show Γ as a function of x in cell of different initial $\Delta\tau_\nu$. For optically thin cells ($\Delta\tau_\nu \lesssim 1$, blue solid and black dotted lines), Γ evolves almost linearly with x, meaning that it begins to drop as soon as any fraction of the cell is ionized. Conversely, for optically thick cells (red and green solid lines), Γ stays constant for a significant range of x. The actual drop in Γ happens only when $\Delta\tau$ begins to transition from optically thick to thin regimes. Plainly put, the more optically thick the cell is initially, the longer it will stay optically thick (remember that $\Delta\tau \propto x$) thus making Γ computed at $x = 0$ to be the correct value longer. This fact directly translates to the differences in the two ionization times shown in the left panel of Fig. 2.12.

These results point to a conclusion that the importance of temporal photon conservation, measured in the delay of ionization due to the change in opacity, decreases with increasing optical depth. It is interesting to point out that it is very important in optically thin cells - while optically thin cells may solve the problem of spatial photon conservation (as discussed in Section 2.5.1), this is not a sufficient fix for the temporal issues. The opposite is true for highly optically thick cells, where the spatial component is crucial (as shown in the left panel of Fig. 2.8), yet time-dependence of $\Delta\tau$ is not. All cells make the transition to the optically thin regime as $x \to 1$ where Γ becomes highly variable, but if this transition happens late (as for $\Delta\tau = 100$ case in Fig. 2.12), the speed of I-fronts crossing the cell is not significantly affected. For reionization simulations in general, RT cells are extremely optically thick (e.g. the minimum optical depth of a cell in the 50 Mpc/h box we use in Chapter 5 is ~ 100, while the maximum ~ 1000), due to the large scales considered. Keeping in mind the conclusions we reached here leads to simpler radiative transfer schemes, where longer time steps can safely be taken without implementing temporal photon conservation. One such scheme is the SimpleX scheme we

use to simulate reionization in the following chapters (see Chapter 3 for more on that scheme).

2.6 Radiative hydrodynamics tests

In some problems, such as the epoch of reionization, the I-fronts are R-type (e.g. Shapiro & Giroux, 1987; Gnedin, 2000b; Ciardi et al., 2003; Iliev et al., 2006a; McQuinn et al., 2007; Trac et al., 2008) and the slow response of gas to the I-front motion allows for the assumption of a static medium. However, a number of astrophysical problems involve slow, D-type I-fronts where the gas motion influences the behaviour of the I-front. Modeling these problems requires numerical methods that fully couple radiation transport with hydrodynamics, e.g. the photo-evaporation of minihalos (e.g Shapiro et al., 2004)

The computational cost of radiation hydrodynamics calculations necessitates the use of many simplifying assumptions. Thus, it is very important to validate schemes for their reliability and accuracy. The approach we take in this section is to compare our results with various other schemes on a set of well-defined simple astrophysical problems. For that purpose, we took part in the second RT Code Comparison project, results of which have been published in Iliev et al. (2009).

The first paper of this comparison (Iliev et al., 2006b, hereafter Paper 1) focused on the fixed density field tests, two of which we already presented in Section 2.4. It was found that all codes produce reliable results for I-front positions, but not all methods are appropriate for problems where a precise determination of the evolution of gas temperature is required, such as the radiative hydrodynamics problems we present here. The chosen test problems are relatively simple, but still trigger the commonly encountered problems in an astrophysical setting. The comparison was made between 10 different codes, see Iliev et al. (2009) (hereafter Paper 2) for more details about each. All the test results had to be supplied on a regular 128^3 Cartesian grid. All the codes that use AMR grids, one of those being our method, have been asked to restrict the refinement and use grids as close as possible to the required data format.

In the rest of this section, we present the three test problems defined for this code comparison. We focus on the results we obtained with the FLASH-HC method, discussing them in more detail than was done in Paper 2, presenting a direct comparison

with some other codes only at a few places. For a more detailed discussion of the differences between the codes, we refer the reader to Paper 2. Note that the test numbering scheme assumes that the tests presented here are a continuation of the set of tests presented in Paper 1.

2.6.1 Test 5: classical HII region expansion

This test problem is the expansion of an HII region in an initially uniform medium around a point source of UV radiation. Irrespective of the simplicity of the setup, it is very relevant for the study of I-fronts, as the front will transition from a weak R-type in the early stages of evolution to a D-type. While analytic solutions can be constructed for early and late stages of the problem, the I-front type transition itself does not have one.

The numerical parameters for the problems are given as follows: the computational box size $L_{\rm box} = 15\,{\rm kpc}$, initial gas number density $n_H = 10^{-3}\,{\rm cm}^{-3}$ (we assume the gas is pure hydrogen), the ionizing photon emission rate $\dot{N}_\gamma = \int_{\nu_{\rm th}}^{\infty} L_\nu/(h\nu){\rm d}\nu = 5 \times 10^{48}\,{\rm photons\,s}^{-1}$, initial gas velocity is zero and its environment temperature $T_e = 100\,{\rm K}$. The source is a 10^5 K black body placed in the $(x_s, y_s, z_s) = (0, 0, 0)$ corner of the box. The boundaries that contain the source origin are reflective while the rest are permissive. The total running time is $t_{\rm sim} = 500\,{\rm Myr}$. Note that this is the exact same setup as in Tests 1 and 2 of Paper 1, only with a larger simulation box.

Figure 2.13 shows the values of 4 relevant variables in the Test 5 results we obtained with FLASH-HC at three different times: 10 Myr, 200 Myr and 500 Myr. These output times showcase three distinct phases in the evolution of the front. At early times, the front is weak R-type: the ionized region propagates without almost no response from the gas (compare left panels of two top rows, x and $n_{\rm H}$, respectively in Fig. 2.13). The expansion of the ionized region begins to stall after a recombination time, $t_{\rm rec} = 122.4\,{\rm Myr}$ (Eq. (2.26)) for the given value of n_H. If the medium were static, the I-front would stop propagating once the Strömgren radius is reached, $r_S = 5.4\,{\rm kpc}$ (Eq. (2.27) when all the emitted ionizing photons are lost to recombinations in the HII region). In the non-static case, the large pressure difference between ionized and neutral gas (due to ionization heating) drives the gas out of the HII region, which decreases the recombination rate and increases the Strömgren sphere, allowing the I-front to expand further. Still, the

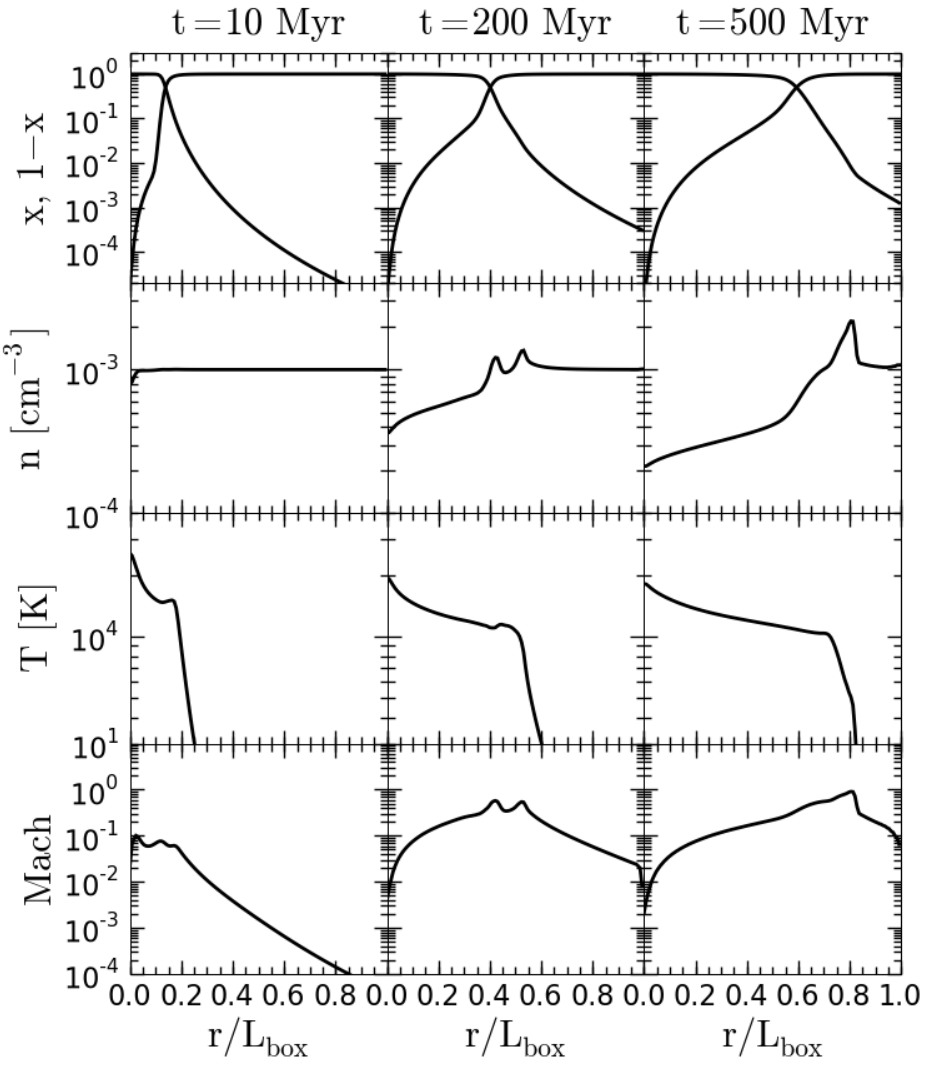

Figure 2.13: Evolution of fluid variables in Test 5 as given by FLASH-HC. Top to bottom rows: neutral and ionized fraction, number density, temperature and Mach number. Columns are the results at three different output times. See text for discussion.

front slows down enough to start the transition from R-type to D-type at $\approx t_{\rm rec}$. The transition is complete by $t \approx 4t_{\rm rec} \approx 500\,{\rm Myr}$, where the I-front is D-type and preceded by a weak shock (third panel second and last rows of Fig. 2.13, $n_{\rm H}$ and Mach number).

All the codes find more or less the same evolution of all variables (see Figs. 11-14 in Paper 2, FLASH-HC results given with pink dashed line). The position of the I-front,

defined to be at $x = 0.5$ (i.e. at the intersection of x and $1 - x$ curves in top rows of Fig. 2.13) and its velocity varies by only a few percent between all codes. Most of the differences are caused by a different handling of the energy equation, Eq. (2.20), and of hard photons with long mean free paths. An interesting flow feature seen by almost all codes are the double peaks in the density distribution at intermediate times (see middle panel in second row of Fig. 2.13). The cause of the double peaks is heating due to high energy photons (see Paper 2 for a detailed discussion). In short, while the I-front position is behind the shock front at $t = 200\mathrm{Myr}$, the high energy tail of the spectrum is able to penetrate the shock. The high energy photons are unable to maintain a high ionized fraction (it is $x \approx 0.1$ at $0.5L_{\mathrm{box}}$), but they still can inject a significant amount of heat, as evidenced by a temperature jump at around $0.45L_{\mathrm{box}}$ visible in the second panel of third row of Fig. 2.13. This temperature increase causes the shock front to split in two by producing opposite photo-evaporative flows in the front frame. The high pressure in the inner part of the HII region causes the inner peak to join its outer counterpart by $t = 500\mathrm{Myr}$. The only code that does not show this feature is ENZO-RT, which is the only scheme taking part in the project that does not implement multi-frequency transport in some form (cyan dashed line in middle panel of Fig. 14 in Paper 2). We also note that all codes find spherically symmetric and dynamically stable fronts (see Fig. 9 in Paper 2). This will not be the case in the other two tests.

In this code comparison, the performance of the contributing schemes was not tested. Still, we wish to point out the performance increase that our improved photon-conserving scheme brings (Section 2.5). Figure 2.14 shows the stability of our Test 5 result with respect to varying the time step. Very long time steps give the same answer as shorter ones at a fraction of the CPU time. The hydrodynamics time step (CFL) is $\Delta t_{\mathrm{hydro}} \sim 3-4$ Myr for the largest part of the calculation and our implementation of HC gives a perfectly accurate result when using that time step (red line in Fig. 2.14). The inclusion of both spatial and temporal photon conservation allows us to solve the tests we are presenting here several orders of magnitude faster than with the original implementation. For example, the results we supplied for Test 5 were obtained in only 4 CPU hours (15 minutes on 16 processors). The results are also robust for changes in spatial resolution, both with the use of AMR of uniform meshes.

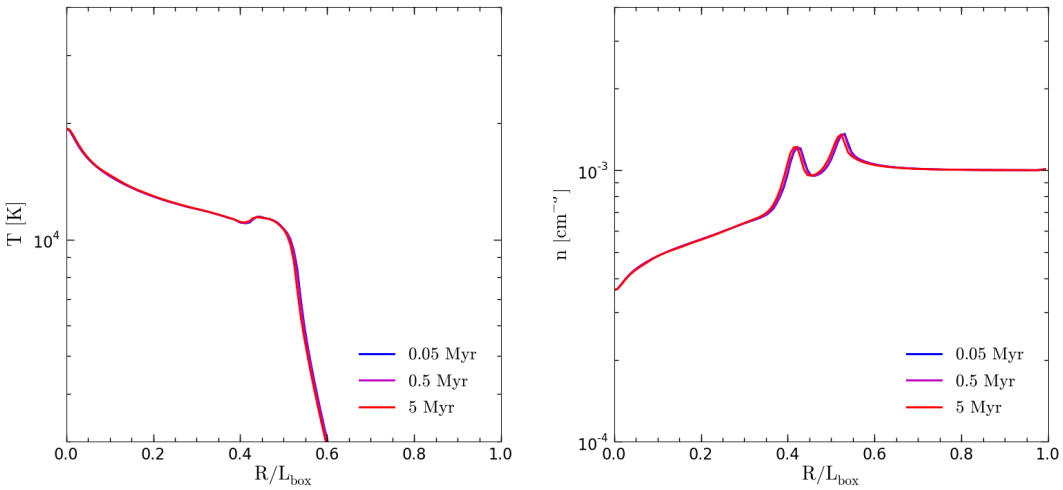

Figure 2.14: Test 5 results of our photon-conserving scheme with varying time step. All flow variables are robust with respect to the calculation time step at all times: here we show temperature (left) and density (right) at 200 Myr. See Section 2.5 for the discussion about photon conservation.

2.6.2 Test 6: I-front propagating in a $1/r^2$ density profile

Test 6 is the propagation of an I-front around a point source located at the centre of spherically symmetric density field with a constant density core surrounded by a steeply decreasing density profile. The (number) density field is defined as:

$$n_H(r) = \begin{cases} n_0 & \text{if } r \leq r_0 \\ n_0 \left(\frac{r}{r_0}\right)^2 & \text{if } r > r_0, \end{cases} \tag{2.42}$$

where n_0 and r_0 are the number density and radius of the central core, respectively. The density profile of Eq. (2.42) was chosen to mimic the conditions for I-fronts breaking out of typical cosmological structure with an embedded radiation source, e.g. minihalos or dwarf galaxies. Again, an analytic solution for the full hydrodynamic problem does not exist, but this type of density profile has been extensively studied with semianalytic and numerical methods (e.g. Franco et al., 1990).

The evolution of the I-front in the core is identical to Test 5, presented in the previous section. The further behaviour of the front depends on whether or not it converts to D-type before it falls onto the slope. If the initial Strömgren radius $r_S < r_0$, the I-front will

begin the transition to a D-type before the slope is reached. In general, when travelling down the density slope a D-type front may accelerate and become R-type once again. The numerical simulations of Whalen & Norman (2008) showed that, for density profiles similar to Eq. (2.42), an initially D-type front remains D-type on the slope for the lifetime of the source, meaning the transition back to R-type is not sudden. If, on the other hand, $r_S > r_0$, the front will remain R-type and "flash" ionize the whole density field. In such a fully ionized and nearly isothermal medium, large pressure gradients will naturally arise wherever there are steep density gradients. This pressure gradient will drive the gas down the density slope r^{-2} at a roughly constant speed (Franco et al., 1990).

The exact numerical parameters for Test 6 were chosen to model an I-front that does transition to a D-type: the computational box length $L_{\rm box} = 0.8\,{\rm kpc}$, $n_0 = 3.2\,{\rm cm}^{-3}$, $r_0 = 91.5\,{\rm pc}$ (the Strömgren radius corresponding to the density in the core is $r_{{\rm S},0} = 67.5\,{\rm pc}$), initial ionized fraction $x = 0$, ionizing photons emission rate $\dot{N}_\gamma = 10^{50}\,{\rm photons\,s}^{-1}$ and initial environment temperature $T_e = 100{\rm K}$. Source position and boundary type are the same as in Test 5. Total running time is $t_{\rm sim} = 75\,{\rm Myr}$, with outputs at times t = 1, 3, 10, 25 and 75 Myr. Note that this setup is not in hydrostatic equilibrium. This fact is ignored for the sake of simplicity, as the neutral gas pressure forces are much weaker than the ones driven by photo-ionization heating.

Figure 2.15 shows the fluid variables as a function of distance from the source in FLASH-HC Test 6 results at three different times: t =3, 10 and 25 Myr. The left columns ($t = 3$ Myr) show the already formed shock front as it approaches the edge of the constant density core. Note the non-zero Mach numbers on the far side of the box: these are due to the pressure forces that arise because the hydrostatic equilibrium was not imposed. By $t = 10\,{\rm Myr}$ (middle panels), the I-front has propagated deep into the density slope, $r_I \approx 3r_0$. The temperature in Fig. 2.15 (third row, middle panel) shows a small plateau of $T \approx 10^4\,{\rm K}$ in front of the radiation driven temperature field, which is due to heating produced by the preceding shock front. Finally, at $t = 25{\rm Myr}$ (right-hand side plots), the front is about to leave the simulation box. The heating caused by the shock front is now evident, as it increases with increasing front speed. All of these flow features are seen in the results of the other codes as well (see Figs. 25-29 in Paper 2).

Slices through the data show a somewhat worse agreement between codes than the spherical averages. Figure 2.16 shows the slices through the density, ionized fraction

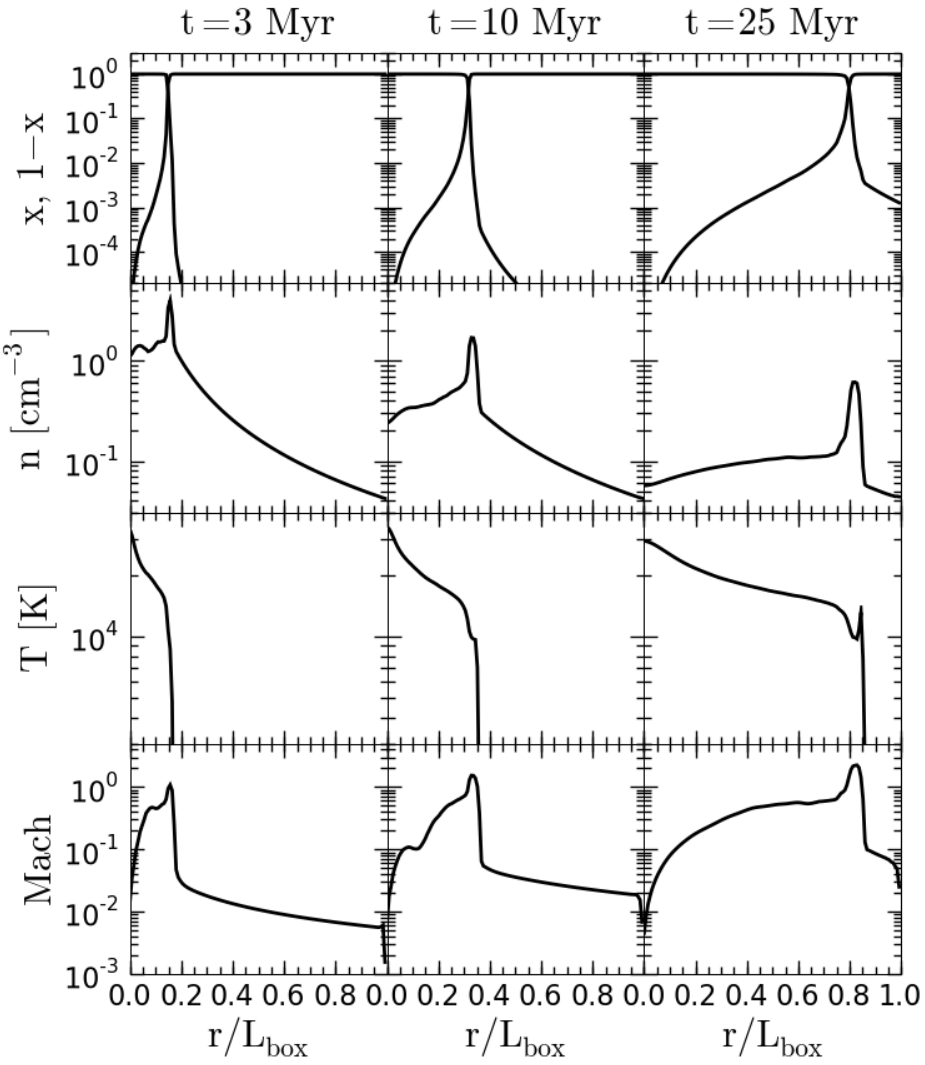

Figure 2.15: Evolution of fluid variables in Test 6 as obtained by FLASH-HC. Top to bottom rows are: neutral and ionized fraction, number density, temperature and Mach number. Columns left to right give results at times t = 3, 10 and 25 Myr. See text for discussion.

and temperature of FLASH-HC at $\mathrm{t} = 25\,\mathrm{Myr}$. There appears to be some weak flow instability, yet its cause is not clear. Our results are not unique in finding deviations from expected spherical symmetry, as is seen in Fig. 2.17. The root of the instabilities is most likely not physical. Paper 2 discusses three known instabilities in the I-front related flows. First, Vishniac-type (Vishniac, 1983) instabilities might arise if the neutral

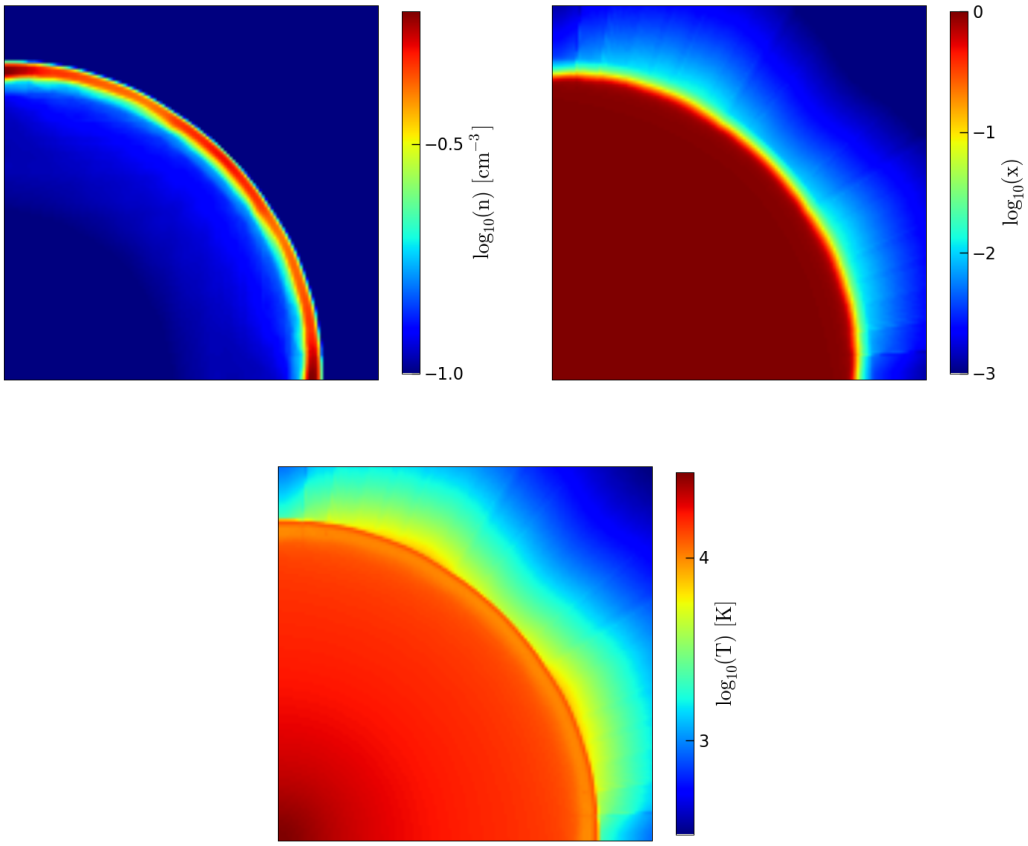

Figure 2.16: Slices along $z = 0$ through density (top left), ionized fraction (top right) and temperature (bottom) at $t = 25\mathrm{Myr}$ as obtained by our FLASH-HC scheme. The results show a presence of weak departures from sphericity in the shock front (top left panel; note that the colorbar shows the restricted data range to point out the issue), which results in varying optical depth along the rays and allows high energy photons to penetrate deeper into the neutral medium along some directions (as seen in top right and bottom panels). See text for discussion of potential causes of this instability.

shocked shell preceding a D-type I-front is able to cool by radiation. Still, this is not the likely case as Whalen & Norman (2008) have shown that atomic hydrogen cooling alone (as used in this test setup) is not enough to instigate such instabilities. Second, I-fronts on descending density profiles may instigate Rayleigh-Taylor instabilities (e.g. Sharp,

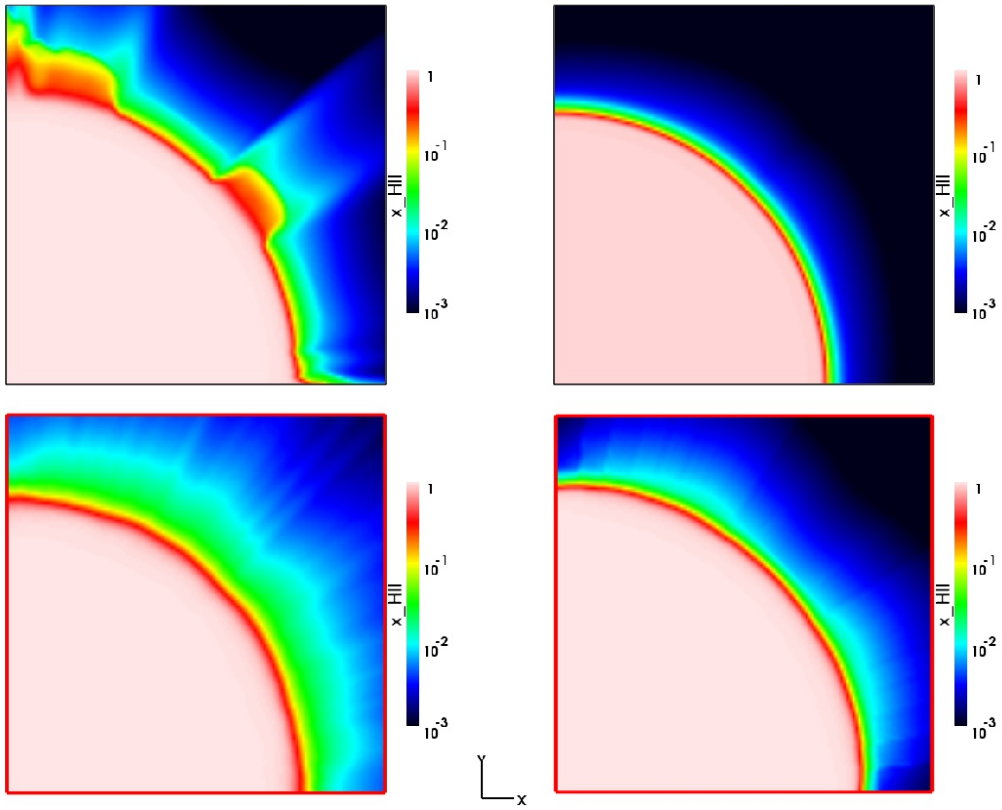

Figure 2.17: Slices along $z = 0$ through the ionized fraction, x, at $\mathrm{t} = 25\mathrm{Myr}$ in 4 codes: clockwise from top left, C^2-ray+Capreole, C^2-ray+TVD, FLASH-HC and LICORICE. FLASH-HC results are not alone in finding some departures from spherical symmetry, though not for the same reasons (see text for discussion). The features seen in x are generally mirrored in the density and temperature.

1984), but the $1/r^2$ used for this test is not known to produce them. Third, small density inhomogeneities may trigger a shadowing instability when they are passed by an R-type front (Williams, 1999). These inhomogeneities should not exist in a density field defined by Eq. (2.42), but may appear due to the discretization of the density field on the computational mesh. In our initial conditions, we assign density to each cell by computing the value of Eq. (2.42) at cell centres, resulting in a departure from spherical symmetry between neighbouring cells. To test if this is indeed the source of the instability, we performed a higher resolution run of Test 6 (256^3 uniform grid) as well as computed the density as a volume weighted mean in each cell, i.e. we computed the densities in each

cell on a much finer grid and use the average value. Neither of these approaches had any effect on the presence of the instabilities seen in Fig. 2.16. The same two approaches were used by Capreole+C^2-Ray (the top left panel of Fig. 2.17, shows the most pronounced instabilities), but also without any significant effect.

As for numerical causes, a likely suspect is the "carbuncle" instability (Quirk, 1994). This is a somewhat unpredictable problem found in Riemann solver-based techniques, such as FLASH and Capreole, arising when a shock travels parallel to a coordinate axis. The case is strengthened by the fact that the TVD solver (Trac & Pen, 2004), which is not known to suffer from carbuncle problems, gives an undisturbed result when coupled with C^2-ray (top right panel of Fig. 2.17). LICORICE (bottom left in Fig. 2.17) shows the irregularities consistent with the spurious fluctuations in the density field which are unavoidable in SPH methods such as the one it employs. Finally, as the 1D profiles presented in Fig. 2.15 and Figs. 25-29 in Paper 2 show, the presence of these instabilities does not significantly affect the global flow properties.

2.6.3 Test 7: photoevaporation of a dense gas clump

This test involves a plane-parallel front encountering a uniform spherical clump in a constant low-density background. This problem is relevant in, for example, the photoevaporation of dense clumps in planetary nebulae (Mellema et al., 1998) or minihalos during reionization (Shapiro et al., 2004; Iliev et al., 2005b). Depending on the column density in the clump, an incoming R-type front may either flash ionize it without significant speed decrease or transition to a D-type and possibly halt completely until the clump is evaporated (as simulated by Bertoldi, 1989 and Iliev et al., 2005b, respectively). Following the same logic for deriving the Strömgren radius in Eq. (2.27), when a plane-parallel front with given flux F encounters a slab of neutral hydrogen with number density n_H, it will penetrate it up to length l_s before recombinations balance the ionizations. The equilibrium is achieved when: $FS = \alpha_B n_H^2 l_s S$, where S is a surface area. This leads to l_S being:

$$l_s = \frac{F}{\alpha_B n_H^2}. \tag{2.43}$$

If hydrodynamics are neglected, the clump is able to trap the front if $2r_{\rm clump}/l_s > 1$, where $r_{\rm clump}$ is the radius of the clump and the length l_S is computed using the number density of the clump (Shapiro et al., 2004).

The numerical parameters for Test 7 are as follows: the constant ionizing flux incident on the $x = 0$ plane is $F = 10^6\,\mathrm{photons\,s^{-1}cm^{-2}}$, the background hydrogen number density and temperature $n_{\rm out} = 2\times10^4\,\mathrm{cm^{-3}}$ and $T_{\rm out} = 8000\,\mathrm{K}$, respectively while the initial clump density and temperature are $n_{\rm clump} = 200 \times n_{\rm out} = 0.04\,\mathrm{cm^{-3}}$ and $T_{\rm clump} = 40\,\mathrm{K}$. This choice of parameters sets the clump in pressure equilibrium before the radiation flux is applied. This column density is sufficient to trap the front and force its transition to a D-type, as we will show shortly. The simulation box length is $L_{\rm box} = 6.6\,\mathrm{kpc}$, the clump has radius $r_{\rm clump} = 0.8\,\mathrm{kpc}$ and is centred at $(x_c, y_c, z_c) = (5, 3.3, 3.3)\,\mathrm{kpc} = (97, 64, 64)\,\mathrm{cells}$. All boundary conditions are outflow. Note that these are the same parameters used in Test 3 presented in Paper 1, where this test was performed on a static medium.

For the FLASH-HC run, we model the plane-parallel front by placing a source of radiation far out of the box, at $(x_S, y_S, z_S) = (-850, 3.3, 3.3)\,\mathrm{kpc}$. The source luminosity is set to produce F at the $x = 0$ box edge. The distance of the source is chosen so that every ray that hits the cell with indices $(0, j, k)$ at $x = 0$, leaves the box from cell $(127, j, k)$, where the grid has 128 cells along each coordinate.

Figure 2.18 shows the initial trapping phase, as seen in FLASH-HC results at $t = 1\,\mathrm{Myr}$. The incoming flux produces a thin ionized shell at the source side and a clear shadow behind the clump. As the gas is essentially static early on, the results are almost the same as in Test 3 (see Paper 1.). There are some small differences as evidenced by the non-zero Mach number (right panel in Fig. 2.18). All the codes agree very well at this stage, with the biggest discrepancies seen in SPH codes that have slightly more diffuse shadows (see Fig. 31 in Paper 2). The discrete assignment of densities to grid cells, already mentioned in Test 6, may introduce grid artifacts that can affect the ablation flow at later times if the fix, also discussed in Test 6, is not applied. Two other grid codes in the comparison, ZEUS-MP and Capreole+C^2-ray, were forced to implement this fix, but we found it does not affect our results.

By t= 10 Myr, the front begins its transition to D-type front and the heated gas at the source side of the clump starts to photoevaporate in the form of a supersonic wind flowing away from the clump. As the front eats through the clump, more and more layers of neutral gas are heated and blown away. Figure 2.19 shows the results we obtained with FLASH-HC. A notable feature in our results, not seen in other codes, is the presence of

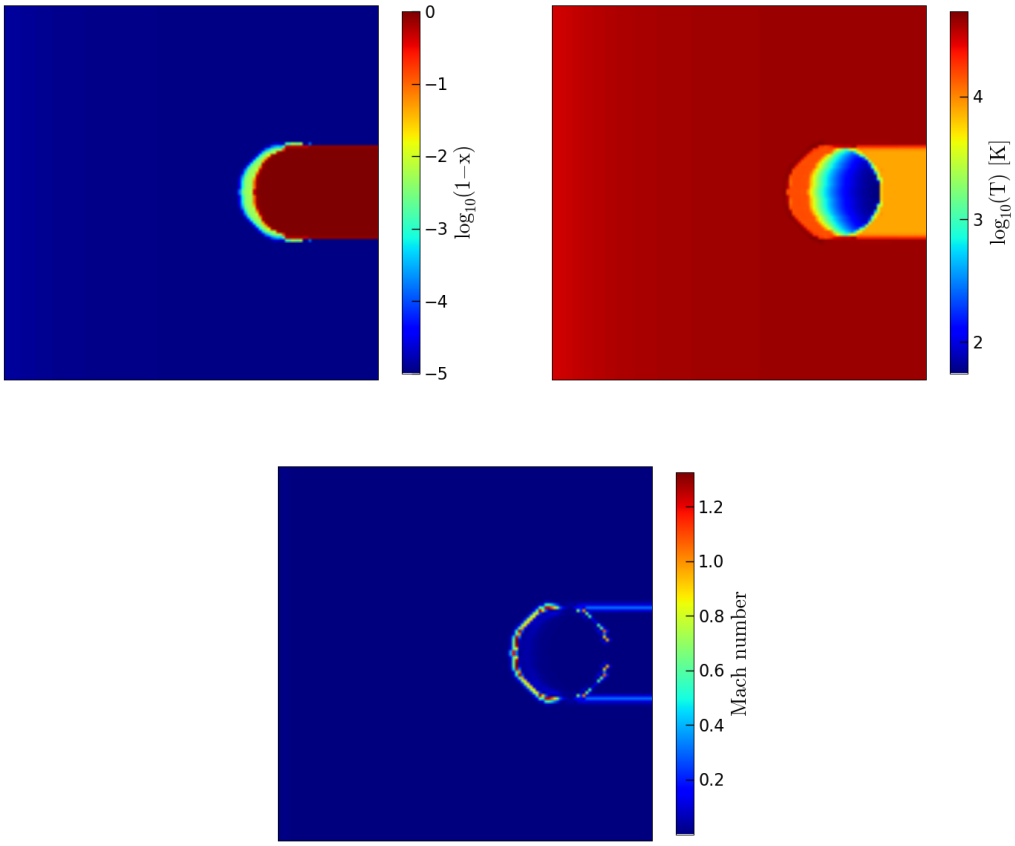

Figure 2.18: Images of neutral fraction $1 - x$ (top left), temperature (top right) and Mach number (bottom) at $\mathrm{t} = 1\,\mathrm{Myr}$, cut through the centre of the clump in FLASH-HC in Test 7. In the early stages of evolution, the results are almost identical to the ones obtained in Test 3 (see Paper 1). The non-zero Mach numbers show that the photo-ionization flow appears almost immediately.

irregularities in the left-hand side edge of the shadow, at the centre of the clump, in the HI values (top left panel of Fig. 2.19). Irregularities are also seen in the outward velocity of the photoevaporative wind (bottom left panel of Fig. 2.19). The source of these irregularities is almost certainly a bug in the ray tracing scheme. We repeated Test 7 with AMR at higher level of refinement (maximum cell resolution corresponding to 256^3 uniform grid), presented in Fig. 2.20. It is clear that increasing of resolution only amplifies the problem, suggesting it is not rooted in any density fluctuations introduced by limited

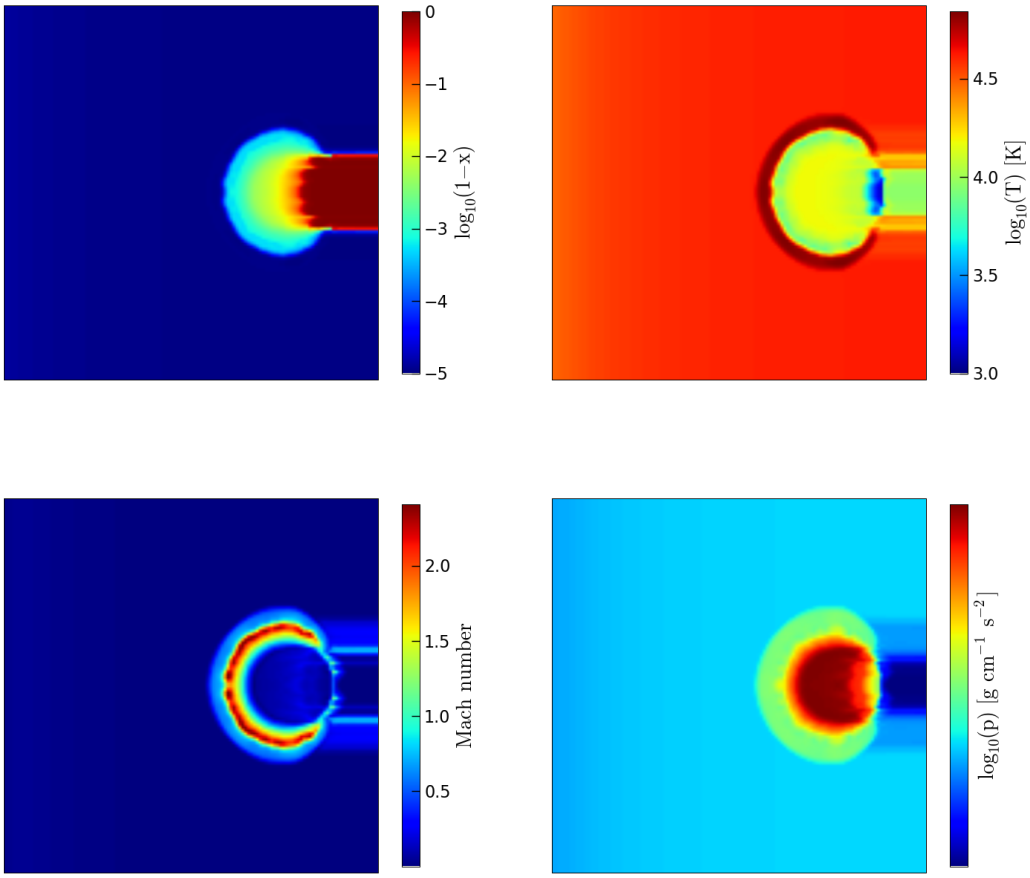

Figure 2.19: Slices through the FLASH-HC results at $t = 10\,\mathrm{Myr}$. Plotted variables are, top to bottom left to right, neutral fraction, temperature, Mach number and pressure. The neutral fraction shows uneven features in the source-side edge of the shadow. The irregularity is also seen in in other variables but to a lesser extent. The root of these irregularities is probably numerical and may be related to our plane-parallel front implementation. See Fig. 2.20 and text for further discussion.

grid resolution. Fig. 2.20 shows that the "fingers" of more highly ionized gas that penetrate the shadow edge trace the edges of blocks. Placing sources out of the box is not a very thoroughly tested functionality and it is possible that some subtle bug leads to these unexpected features. The original HC implementation performed Test 3 in Paper 1 and these irregularities were not found. Still, in that work, the source was not placed outside of the computational box, but the box was extended to place the source far from

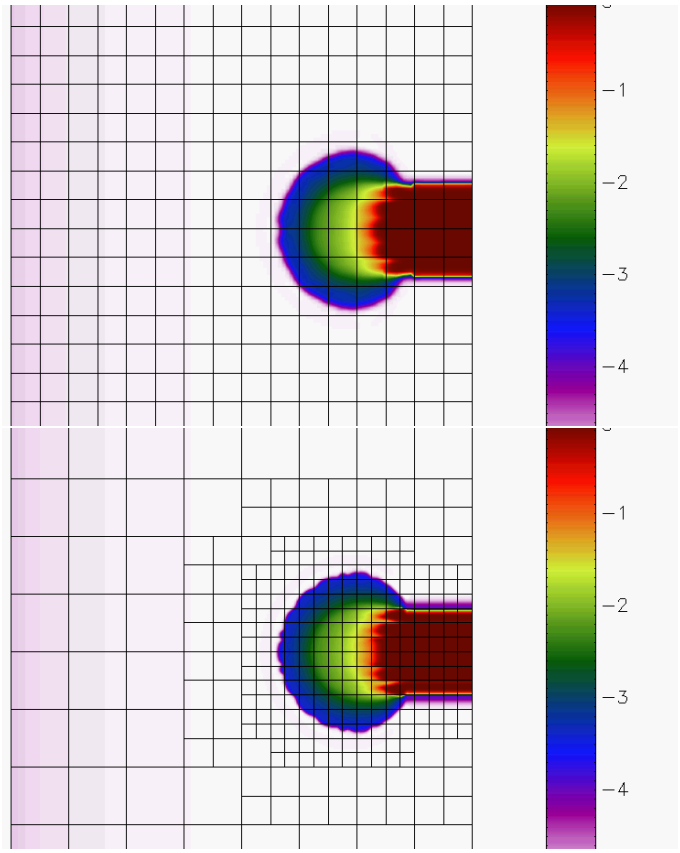

Figure 2.20: The effect of grid resolution on irregular protrusions in the left-hand side edge of the shadow in the neutral fraction results of FLASH-HC at t = 10 Myr. The top panel is the default, 128^3 grid result with block boundaries overplotted for reference. The bottom plot is the same problem with higher resolution in the region of interest (AMR grid). The features in the shadow edge only increase in intensity once the resolution is increased, meaning the density field resolution is not the cause. Features seem to follow the block edges in both panels, potentially pointing to a subtle bug in the ray tracing procedure as a possible cause.

the clump. This lead to many unnecessary computations, as blocks must be placed in those unused regions. The number of extra blocks makes placing a source at our desired distance impossible, so the front coming into the regions interest was not plane-parallel in the same terms we presented in the test description. Finally, as we will show in the rest of this section, the effect of this irregularity does not significantly affect the evolution

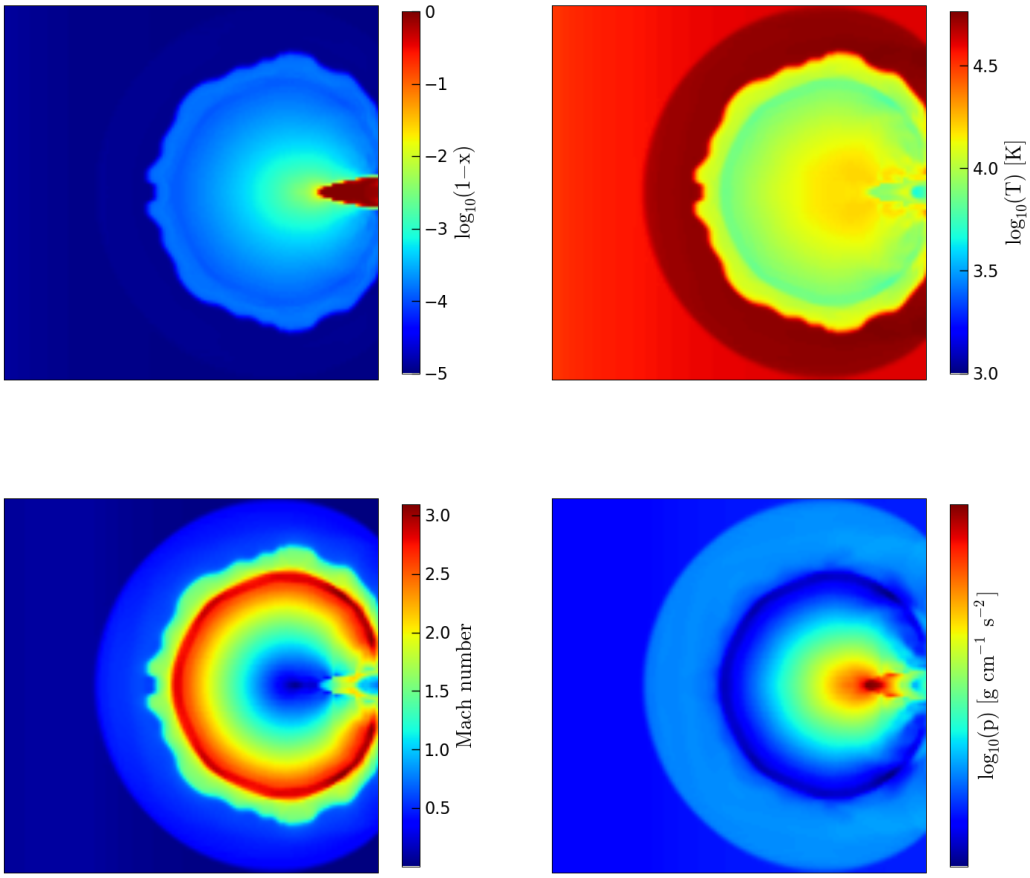

Figure 2.21: Slices through the FLASH-HC results at $t = 50\,\mathrm{Myr}$. Plotted variables are, top to bottom left to right, neutral fraction, temperature, Mach number and pressure.

of the system, nor produce results very different from other codes.

Figure 2.21 shows the state of the system in FLASH-HC at t = 50 Myr. By that time, the photoevaporation is well advanced, with only a small core of neutral material left in the clump. The shape of the core is distorted due to pressure forces from the surrounding hot, ionized gas. The temperature and pressure (top and bottom right panels in Fig. 2.21) clearly show a supersonic shock front sweeping the previously photoevaporated material. The inner region of the shock cools adiabatically, while the material in front of the front is very hot, $\approx$ 60,000 K in FLASH-HC, within the range of 40,000-70,000 K found in other codes.

We compare the morphology of our result with other codes in Fig. 2.22, where we

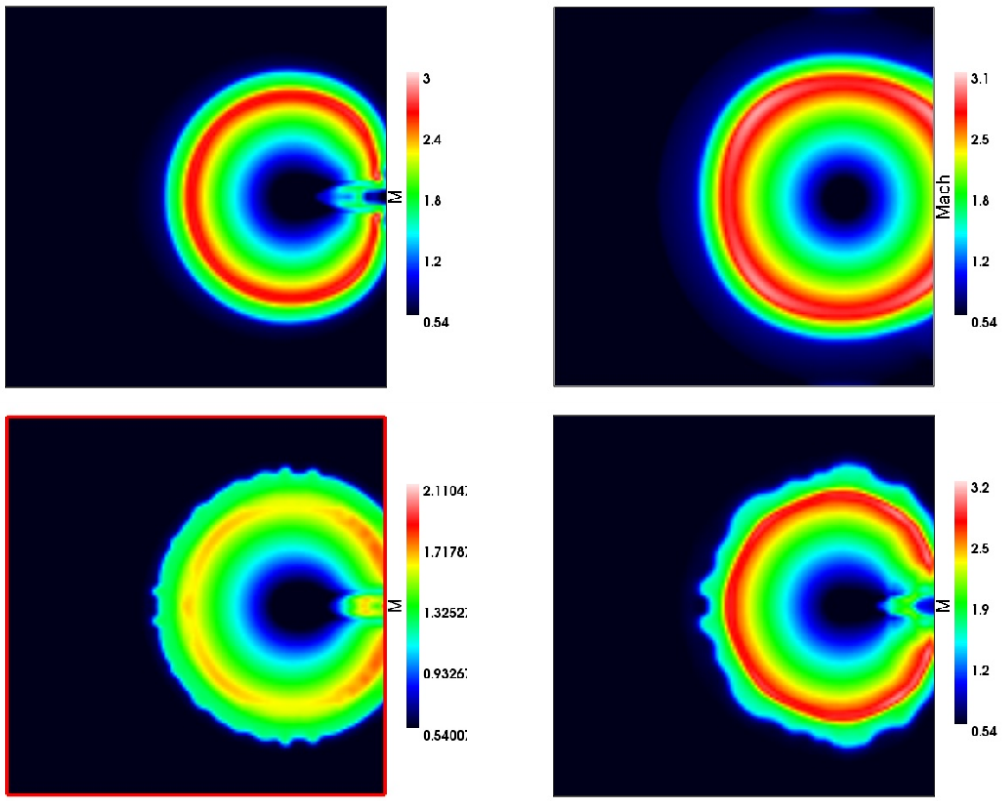

Figure 2.22: Mach number images cut through $z = 3.3\,\mathrm{kpc}$ at t = 50 Myr for (from left to right and top to bottom) Capreole+C^2-ray, ZEUS-MP, Coral and FLASH-HC. The morphology found by FLASH-HC, while somewhat irregular, generally agrees with other codes.

show the Mach numbers at t= 50 Myr. The results we obtained with FLASH-HC, even though they show some flow irregularities, are not far off from the rest of the codes. For example, Coral finds a significantly weaker shock, probably due to the particular hydrodynamic solver employed (based on van Leer flux splitting, see Mellema et al. (1998); Shapiro et al. (2004) for more details). ZEUS-MP produces an almost spherical supersonic wind, a feature not reproduced by other codes.

Finally, the morphological issues we presented do not affect the general evolution, as seen in spherically-averaged profiles in Fig. 2.23. At early times all codes agree very well on the position of the I-front and its profile (first two panels of Fig. 2.23). At later time, the hydrodynamic evolution introduces some differences, yet the scatter remains small. Temperature plots in Fig. 2.23 show that all the codes agree in the shape and

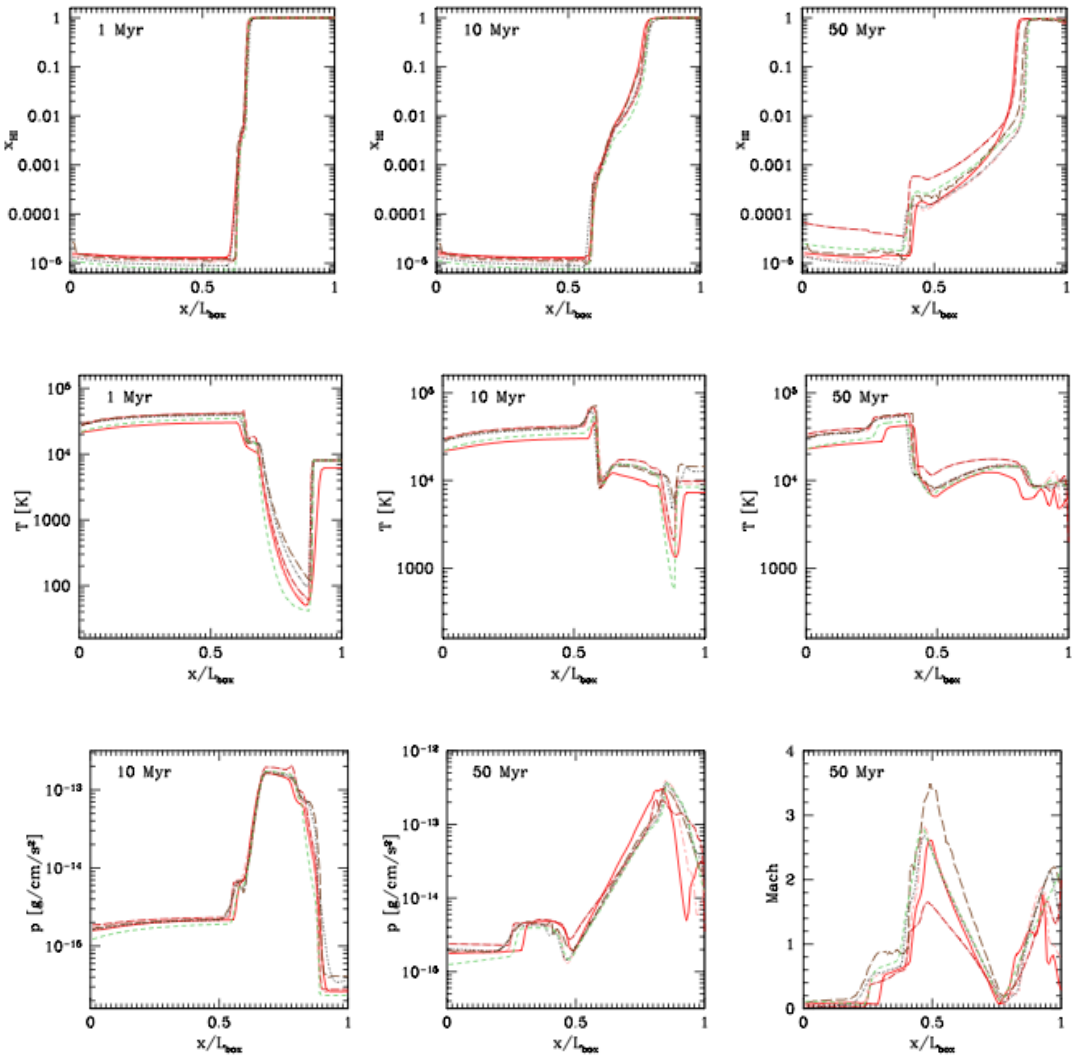

Figure 2.23: Evolution of neutral fraction (top row), temperature (middle row), pressure and Mach number (bottom row) Test 7 results of all codes. FLASH-HC results are shown as pink dashed line. See text and Paper 2 for discussion.

position of the major flow features, with the distinction being the absolute values. The largest variance is found in the temperature of the neutral gas in the clump at t=10 Myr, which is again explained with the different treatment of high energy photons. This fine agreement is also seen in pressure and Mach number plots, not presented here.

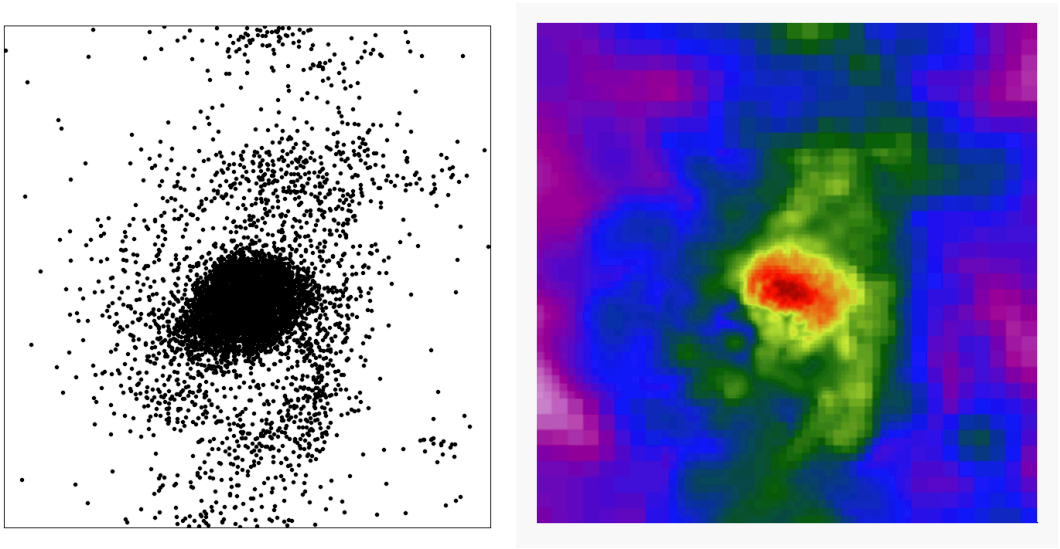

Figure 2.24: A single galaxy from the GIMIC simulation: the position of SPH particles (left) and the density field on a FLASH grid obtained with the converter module (right).

2.6.4 Conclusion:

We presented the results of 3 radiative hydrodynamic tests we performed with FLASH-HC scheme for the RT Code Comparison project (Iliev et al., 2009). While some problems in the HC scheme were uncovered during the tests, the results we obtain are very close to the other methods. We also note that the improvement in computation speed we introduced, made possible by the inclusion of the photon-conserving chemistry solver, does not affect the final result. We have shown that our improvements allow for the time step of the whole calculation to be chosen by the hydrodynamics criteria, significantly improving the performance of the whole scheme. See Chapter 6 for more discussion and future uses of this scheme.

2.7 Gadget-to-Flash converter

To facilitate future projects, we have written a new FLASH module that converts output of Gadget (Springel et al., 2001; Springel, 2005a) SPH simulations to the FLASH AMR grid. Gadget is an SPH hydrodynamics method (Monaghan, 1992), therefore the density field is represented by discrete particles and the mass of each is assumed to be smoothed

in a sphere of radius h, called the smoothing radius. Here we briefly describe the algorithm for mapping the SPH fluid variables from particles to the FLASH AMR grid.

In the first step, the AMR block hierarchy is created. First, all the blocks are refined to the minimum level. Next, we mark blocks for further refinement if a particle found within it has $h < f_p l_{\rm block}/2$, where $l_{\rm block}$ is the physical size of the block and f_p is the fraction of the block given as an input parameter. For example, when $f_p = 1/n_x$, the block is marked for refinement if the smoothing diameter of a particle, $2h$, found in the block is smaller than one cell. Once all the particles have been tested or all blocks marked for refinement, the next level of refinement is added. The procedure is repeated until all the blocks are fully refined in respect to $f_{\rm p}$ or the maximum level of refinement is reached. We point out that at no stage must all SPH particles be kept in memory: all operations involving particles can be performed on a chunk of SPH data, with the chunk size being chosen by the user.

Once the grid is constructed, the mass from each particle is assigned to grid cells that are intersected by the particle smoothing volume. We use a normalised 3D Gaussian instead of the default spline kernel of Gadget-2 (Monaghan & Lattanzio, 1985; Springel, 2005a) to represent the mass distribution within the smoothing volume. The standard deviation along every coordinate is chosen so that it provides a good match for the shape of the spline kernel used in Gadget-2. The Gaussian is truncated after 3σ and normalised so the mass of smoothed particles is conserved. The Gaussian kernel is employed because it can be easily integrated over in Cartesian coordinates to obtain the exact amount of mass to be assigned to every cell. It is also computationally efficient, since the error functions, $\rm erf(x)$, that result from the integration of Gaussian functions are implemented as intrinsic functions in FORTRAN. All the other fluid variables are assigned to cells as the mass weighted mean from the contributing particles.

As an example, Fig. 2.24 shows the density field of a single galaxy extracted from the GIMIC simulation (Crain et al., 2009) converted to a FLASH grid. We intend to use this FLASH module in our future projects, as well as make it available to the whole FLASH community.

Chapter 3

Radiative transfer with SimpleX

3.1 Introduction

In this chapter we will discuss the radiative transfer scheme, SimpleX, we employ for simulating reionization. Section 3.2 presents the SimpleX method in some detail. The following sections discuss the tests we performed as well as improvements we introduced to make the method more suited for our purposes. We discuss the sampling strategies for computational grid creation in Section 3.3. The representation of the density field on a computational grid is discussed in Section 3.4. We also present the results of a cosmological Strömgren sphere test that lead to the correct way of calculating recombinations in a highly optically thick cells in Section 3.5. Finally, the parallelization of the SimpleX algorithm we implemented is briefly presented in Section 3.6.

The epoch of reionization is a global transition involving most baryons in the Universe. In order to obtain a converged statistical representation of this process from a computer model, a simulation box of order of 100 Mpc/h side length must be used (e.g. Barkana & Loeb, 2001; Iliev et al., 2006a). The intensities of the density field in such a box span several orders of magnitude: from the highest density virialized halos, through overdense filamentary structures to under-dense voids. The sources of ionizing radiation are found in the highest density regions, where the recombination of ionized gas is most effective as well. It is then imperative for the numerical RT method to correctly represent the smallest spatial scales in a cosmological density field. This means that some kind of adaptive computational grid, in the same vein as the one discussed in Chapter 2, should be used to provide the necessary detail while reducing computational costs.

The number of dark matter halos that can host ionizing radiation sources in such cos-

mological volumes is counted in the millions. Therefore, any RT method employed for reionization must be capable of efficiently handling such a high number of sources. This instantly disqualifies most characteristics-based techniques (i.e. Hybrid Characteristics, Chapter 2), whose computational costs scale linearly with the source count. Note that there are exceptions, like the C^2-ray code (Mellema et al., 2006).

The SimpleX scheme discussed in this chapter uses a spatially adaptive unstructured grid to represent the density field and is explicitly independent of the number of sources. It is also very computationally efficient, allowing for quick production of ionization field realisations with modest computational resources. This makes it ideal for our purposes.

3.2 The SimpleX method

3.2.1 Photon packet representation of the radiation field

On the smallest scale, the movement of a single photon through a medium can be described as a random walk, fully specified by the interaction cross-section, α. The interaction coefficient is a combination of two processes: $\alpha = \alpha^{\rm abs} + \alpha^{\rm scat}$, where $\alpha^{\rm abs}$ is the cross-section for absorption and $\alpha^{\rm scat}$ is the scattering cross-section, where the direction of scattering is given by a normalised probability distribution function, $f(\vec{n}, \vec{n'})$ that gives the probability for a photon incoming from direction $\vec{n}$ to be scattered in the direction $\vec{n'}$. In the large number of photons limit, the microscopic random walk transform into the macroscopic radiative transform equation (Chandrasekhar, 1943).

The small scale picture motivates the SimpleX approach to computationally solving the radiative transfer equation, Eq. (1.16), as well as many Monte Carlo methods (e.g. Maselli et al., 2003; Semelin et al., 2007; Altay et al., 2008; Cashwell & Evrett, 1959, for a general discussion). Tracing single photons through a medium is impractical. Instead, photons are grouped in *photon packets*, discrete amounts of radiation energy emitted from a source during a time step. In the case of ionizing radiation:

$$N_\gamma = \frac{\Delta t}{N_{\rm packets}} \int_{\nu_{\rm th}}^{\infty} \frac{L_\nu}{h\nu} {\rm d}\nu, \tag{3.1}$$

where N_γ is the number of photons in the packet, the integral over luminosity, L_ν, gives the total number of ionizing photons emitted per unit time (since it is integrated from the Lyman limit frequency, $h\nu_{\rm th} = 13.59\,{\rm eV}$, $\nu_{\rm th} = 3.29 \times 10^{15}{\rm Hz}$), Δt is the simulation time

step and $N_{\rm packets}$ is the number of photon packets representing the radiation emitted by the source during Δt. Photon packets are transported along paths defined by the required spatial sampling (i.e. radiation energy must reach all the computational cells that it is supposed to) and the interaction with matter (e.g. photon packets may change paths due to scattering). Along each path, a photon packet may interact with the traversed density field and lose energy due to absorption or gain energy due to emission.

3.2.2 The grid

Voronoi tessellation

The photon packet representation of the radiation field is independent of the computational grid used to represent the density field or the exact choice of photon paths. Most commonly employed grids for this purpose in Monte Carlo methods are Cartesian grids, either regular (e.g. Maselli et al., 2003) or adaptive (e.g. Semelin et al., 2007; Altay et al., 2008). SimpleX uses an unstructured mesh, the Voronoi tessellation, to represent the density field.

Broadly defined, a tessellation is any set of geometrical figures that covers ("tessellates") space without any overlaps or gaps (see Okabe et al., 2000, for more details on the Voronoi diagram and tessellations in general). Let $\Phi = x_i$ be a finite set of points in the d-dimensional space R^d. The Voronoi tessellation of that space is:

$$V(\Phi) = C_i, \tag{3.2}$$

where C_i is a Voronoi cell with nucleus x_i. The cells are defined as:

$$C_i = \left\{ y \in R^d : \| x_i - y \| \leq \| x_j - y \|, \forall i \neq j \right\}. \tag{3.3}$$

Simply put, Voronoi cell C_i is a set of all points in R^d closer to x_i than to any other point in Φ. An example Voronoi diagram is shown in the left panel of Fig. 3.1. In R^3 space, Voronoi cells are polyhedrons. Their *facets* are polygons, bounded by lines called *Voronoi edges*. The intersections of Voronoi edges give a set of points called *Voronoi vertices*. A set of Voronoi vertices and edges fully defines the Voronoi diagram corresponding to the set of points Φ.

At a first glance, using a Voronoi diagram as a computational grid may seem unnecessarily complicated. Indeed, in comparison to regular Cartesian meshes, the conceptualisation and construction of the Voronoi diagram is much more challenging. Why, then,

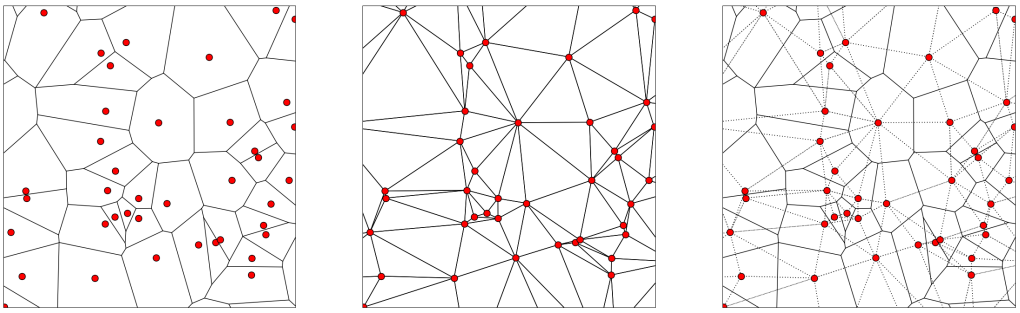

Figure 3.1: *Left:* An example Voronoi diagram in 2 dimensions. The red circles are the Voronoi cell nuclei. *Middle:* Delaunay triangulation of the same set of points. The Voronoi nuclei are tips of the Delaunay triangles. *Right:* Combination of the previous two plots, showing the relation of the Voronoi diagram and the Delaunay triangulation. The Delaunay lines connect all Voronoi cells that share an edge.

should one chose to use it over the regular meshes? First and foremost, it allows for a more precise estimate of the density field. Voronoi based density estimation techniques have been shown to be much more precise than the regular grid based ones (e.g. Zaninetti, 1991; Pelupessy et al., 2003; Melnyk et al., 2007; van de Weygaert, 2007). This is due to the fact that the Voronoi grid can be constructed directly on the points defining the density field, whether they are galaxies obtained form observational surveys or the particle positions in an N-body simulation, and it naturally adapts to the spatial scales of the data. When used as a computational grid, it avoids common mesh tangling that plagues other unstructured grids (e.g. Gnedin, 1995). This is why the Voronoi diagram was recently used by Springel (2009) in a hydrodynamics scheme that bridges the gap between the Lagrangian and Eulerian approaches. Still, for radiative transfer, the most important benefit of using the Voronoi diagram is its dual graph, the Delaunay triangulation.

Delaunay triangulation

Graph theory (e.g. Harary, 1994) tells us that every planar graph G, i.e. a graph whose edges intersect only at the vertices, has a dual planar graph G^*, which has a vertex in every region of G and its edges connect neighbouring regions in G. The term "dual"

references the "symmetric" property of the graphs: if G^* is a dual of G, then G is a dual of G^*. The dual of the Voronoi diagram is the Delaunay triangulation. Given the Voronoi diagram, the Delaunay triangulation is easily obtained: simply connect nuclei of all Voronoi cells that share an edge with straight lines. The resulting diagram is a *triangulation* of space, since its unit figure is a *simplex*, a d-dimensional generalisation of a triangle (e.g. it is simply a triangle in 2D space). The connection with the Voronoi diagram is even deeper, as it can be shown that the circumcenters of all simplices are Voronoi vertices (Okabe et al., 2000). The triangulation example is shown in the middle panel of Fig. 3.1.

From the RT perspective, the most important property of the Delaunay triangulation is that it connects every Voronoi cell with its neighbours. Thus, the Delaunay lines provide natural paths for the propagation of photon packets. This approach removes the need for costly ray tracing operations, but it also introduces some diffusion, since the Delaunay line directions are isotropic only on average. Furthermore, the triangulation is computed only once thus defining all the possible paths before the RT calculation starts. Thanks to this property, the scheme does not have to move only one photon packet at a time as it is commonly done in Monte Carlo RT methods. Instead, *all the photon packets can be moved at the same time*. In practise, this means that the loop over individual photon packets is replaced by a loop over Voronoi cells. In every step, a cell receives packets and redistributes them to all neighbours after accounting for the interaction with the medium. It is this algorithm that is the root of the major strength of the SimpleX method: the independence of the computational cost from the number of radiation sources. The number of operations for the radiation transport in a single step is of order $O(N)$ at the most, where N is the number of cells, in a case where all cells send and receive photon packets.

SimpleX uses the QHull library (Barber et al., 1996) to construct the Voronoi/Delaunay diagram. This is one of the fastest available libraries designed for this purpose. It also includes provisions for handling poorly behaving point sets (e.g. overlapping points). Before we continue with the overview of the interaction of radiation with matter in the SimpleX scheme, we will take a brief look at the specific grid choice that motivated the original development of SimpleX.

Grid based on the photon mean free path

Ritzerveld et al. (2003) advocated a specific setup of the Voronoi/Delaunay grid for use in radiative transfer. They argued that the points used as Voronoi nuclei should be chosen in such a way that the Delaunay lengths correlate, on average, with the mean free path of photons in the medium. This means that the probability for interaction with matter will be the same along every line, i.e. it needs to be computed only once.

In the simplest case, a uniform density field, this desired relation can be obtained by sampling the mass distribution with a Poisson point process (see e.g. Stoyan et al., 1996, for more details about point processes). For general density fields, a correlated point process must be used for the same purpose, where the probability of placing a point is weighted by the local density at its potential position. The local mean free path of photons (e.g. Chandrasekhar, 1950) at point $\vec{r}$ is:

$$\lambda(\vec{r}) = \frac{1}{\sigma n(\vec{r})}, \tag{3.4}$$

where σ is the interaction cross-section and $n(\vec{r})$ is the local number density. On the other hand, the average local Delaunay line length can be obtained from stochastic geometry (see Okabe et al., 2000):

$$\lambda_{\mathrm{D}}(\vec{r}) = \frac{\zeta}{n_{\mathrm{D}}(\vec{r})^{1/d}}, \tag{3.5}$$

where $n_D(\vec{r})$ is the local grid point density, d is the number of spatial dimensions and ζ is a constant that depends on d, $\zeta_{2\mathrm{D}} = 32/9\pi \approx 1.132$ and $\zeta_{3\mathrm{D}} = 1715/2304(3/4\pi)^{1/3}\Gamma(1/3) \approx 1.237$. The originally stated requirement is that the two lengths defined in Eqs. (3.4) and (3.5) scale linearly:

$$\lambda_{\mathrm{D}} = Q\lambda, \tag{3.6}$$

where Q is a *global* constant, i.e. it applies to all Delaunay lines. Replacing Eqs. (3.4) and (3.5) into Eq. (3.6), we see that the linear scaling is reproduced if the density of Voronoi grid nuclei is:

$$n_D(\vec{r}) \propto n^d(\vec{r}). \tag{3.7}$$

For example, in a 3D density field, the probability of putting a grid point at a given position is weighted by the cube of the local density at that point.

Following Eq. (3.6), we see that on the grids constructed with such a scaling the optical depths will, on average, be the *same* along all Delaunay length, $\tau = \lambda^{-1}\lambda_D = Q$.

While elegant, this idea is almost unattainable in practise. For one, when modeling ionizing radiation transport, this scaling holds only when the grid is constructed. The reason is clear: the mean free path depends only on the neutral hydrogen density, $n_{\rm HI}$, so as soon as the gas ionization state changes, the mean free path scaling is broken. This can be avoided by changing the grid every time the ionization state is updated (see Paardekooper et al., 2009, for a possible implementation and discussion). The update requires either a complete reconstruction of the grid or the use of more sophisticated triangulation techniques that allow for the incremental change of the triangulation. In any case, the computational costs would be much higher. More importantly, the steep sampling weights of Eq. (3.7) result in many more points placed in the high density regions which, in turn, exacerbates the numerical diffusion. Simply put, photon packets are not getting where they should. We will come back to this issue in Section 3.3, where we will discuss and test the sampling techniques.

3.2.3 Evolving the ionization field with SimpleX

In this section we discuss how the SimpleX schemes obtains the mean ionizing intensity and solves the chemistry equations to obtain the evolution of the ionization field on the defined computational grid. The problem in which we employ the scheme, cosmic reionization of hydrogen, allows us to make a few simplifying equations. First, radiation is assumed to be monochromatic, with all photons at the Lyman limit frequency, $\nu_{\rm th}$. This approximation is reasonable when sources have stellar spectra (e.g. Loeb, 2006), as they do in our work. Second, we assume that radiation only interacts with hydrogen gas. Again, this approximation is justified by the source spectrum being strongly peaked at the Lyman limit, meaning only hydrogen can be fully ionized before the number density of QSOs increases enough to produce significant numbers of high energy photons (e.g. Madau, 1999; Barkana & Loeb, 2001). The heating rate, Eq. (2.20), for Lyman limit photons only is zero, leading to the same temperature in the neutral and ionized gas. This is clearly wrong, which motivates our final approximation: we set the temperature of the ionized gas to 10^4 K, a typical temperature in HII regions after thermal equilibrium is reached (e.g. Osterbrock, 1989, see also Fig. 2.7). The assumed constant temperature removes the need for solving the energy equation, Eq. (2.19), reducing the system of chemistry equations to a single rate equation.

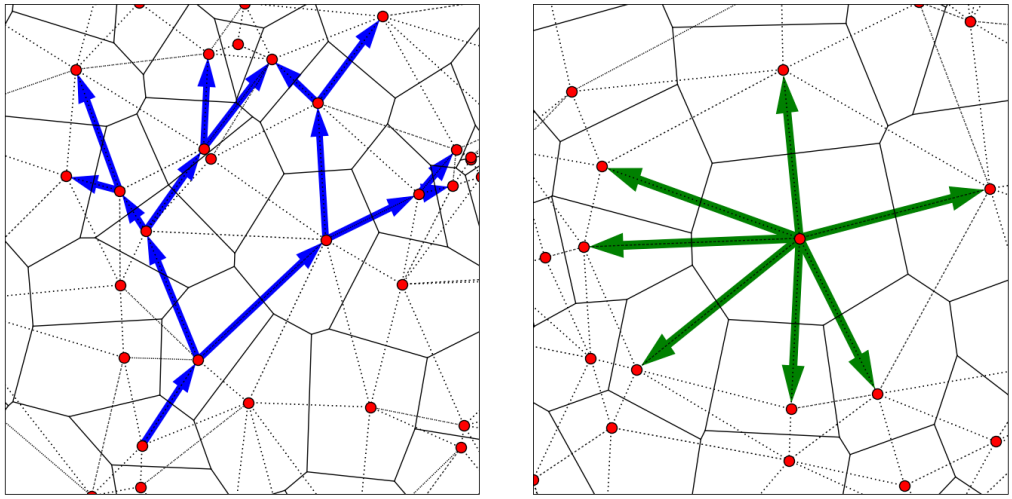

Figure 3.2: *Left:* Propagation of radiation in the SimpleX scheme. At every grid node, the incoming photon packet (decreased by the number of photons remaining in the cell) is split among the most straightforward outgoing lines. The splitting of the photon packet is done to improve angular sampling. *Right:* Source cell emitting new photon packets. The total number of photons in a source cell is split among all the Delaunay lines leading from it. It is easy to implement anisotropic radiation, i.e. sent photon packets can have different sizes. Diffuse radiation, e.g. recombination radiation, can easily be added by making every cell a source.

The ionization rate is obtained by transporting photon packets along Delaunay lines and modeling the absorption of radiation energy along each line. When a photon packet is moved from one cell to the next, the number of photons deposited in the cell is:

$$N_{\rm kept} = N_{\rm in}(1 - e^{-\tau}), \tag{3.8}$$

where $N_{\rm kept}$ is the number of photons that interacts with neutral atoms in the cell, $N_{\rm in}$ is the number of photons in the packet entering before interaction in the cell and $\tau = \sigma_{\rm th} n_{\rm HI} l_D$ is the optical depth along the Delaunay line of length l_D the photon packet travels along. In addition, $n_{\rm HI}$ is the number density of neutral hydrogen in the cell[1]

[1]Even though a Delaunay line passes through two Voronoi cells, we simplyfy the calculation by assuming that the density along the line is the same as in the origin cell. We tested this assumption and found it not to affect our results, as neighbouring cells typically have similar densities.

and $\sigma_{\rm th} = 6.3 \times 10^{-18}{\rm cm}^2$ is the ionization cross-section at the Lyman limit frequency $\nu_{\rm th}$ (e.g. Spitzer, 1978; Osterbrock, 1989). The number of photons in the packet that do not interact with the cell is:

$$N_{\rm out} = N_{\rm in} e^{-\tau}, \tag{3.9}$$

so $N_{\rm out}$ photons leave the cell on the opposite side. Instead of propagating further a single packet, it is instead split in three (two in 2D) new packets propagated along the three most straightforward paths leading away from the initial incoming direction. The splitting is done to improve sampling of radiation directions. An example of a single photon packet propagating over a 2D grid is shown in the left panel of Fig. 3.2. The photon packets are created in source cells, cells containing a source of radiation. The source luminosity L is used in Eq. (3.1) to compute the number of photons initially assigned to each packet, where the number of packets $N_{\rm packets}$ is equal to the number of Delaunay lines leading from the cell. Emission of new packets is shown in the right panel of Fig. 3.2. The same procedure is used to model diffuse radiation, as every cell can be treated as a source at no extra computational cost.

If N_H is the number of hydrogen atoms in the cell, then the ionized fraction x defines the number of ionized atoms, $N_{\rm HII} = xN_{\rm H}$, and neutral atoms, $N_{\rm HI} = N_{\rm H} - N_{\rm HII} = (1-x)N_H$. With the assumptions made above, the evolution of the ionization state of the cell is given by the ionized hydrogen rate, Eq. (2.18), which we rewrite in terms of the number of atoms instead of densities:

$$\frac{{\rm d}N_{\rm HII}}{{\rm d}t} = \Gamma N_{\rm HI} - \alpha_B(10^4\,{\rm K})\frac{N_{\rm HII}^2}{V}, \tag{3.10}$$

where V is the cell volume and we used $N_{\rm HII} = N_{\rm e}$ for a hydrogen-only medium. Note that we ignore the collisional ionization rate, since it is not a significant contributor at $T = 10^4\,{\rm K}$. We also ignore the diffuse recombination radiation and use case-B recombination rate throughout the rest of this work (see Section 2.3.3 for the definition of case-B recombination rate).

We assume that all the photons that interact with the cell within the time step go towards ionizations, therefore the change in the number of ionized atoms due to ionizations is:

$$\Delta N_{\rm HII}^{\rm ion} = \sum_{i=1}^{N_{\rm packets}} N_{\rm kept,i}, \tag{3.11}$$

i.e. it is the sum of interacting photons, given by Eq. (3.8), from all photon packets that cross the cell in a time step. In case that $\Delta N_{\rm HII}^{\rm ion} > N_{\rm HI}$, the number of extra photons kept is returned to outgoing photon packets, i.e. there can be no more ionizations than there are neutral atoms in the cell. The photo-ionization rate directly follows from $\Delta N_{\rm HII}^{\rm ion}$:

$$\Gamma = \frac{\Delta N_{\rm HII}^{\rm ion}}{N_{\rm HI}\Delta t}, \tag{3.12}$$

where Δt is the time step. The rate equation (3.10) is integrated explicitly:

$$N_{\rm HII}(t+\Delta t) - N_{\rm HII}(t) = \Delta N_{\rm HII}^{\rm ion}(t) - \alpha_B(10^4\,{\rm K})\frac{N_{\rm HII}^2(t)}{V}\Delta t, \tag{3.13}$$

where (t) marks the value of the variable at the beginning and $(t+\Delta t)$ at the end of the time step.

The time step is intimately connected to moving photon packets on the grid, the grid *sweeps*. A single sweep is a loop over the grid in which all photon packets are moved from one cell to the next. The time step is then: $\Delta t = T_{\rm sim}/N_{\rm sweeps}$, where $T_{\rm sim}$ is the total simulation time of a single RT run and $N_{\rm sweeps}$ is the number of sweeps. The minimum $N_{\rm sweeps}$ is given by the requirement that the photon packets be moved enough times to cross the simulation box. In practise, a much larger $N_{\rm sweeps}$ is needed to reach a converged result. Convergence is achieved when Δt is a fraction of the shortest recombination and ionization times on the grid. For example, through some trial and error, we found $N_{\rm sweeps} = 1500$ to be an appropriate number for reionization runs we perform in this thesis (see Chapter 5). For example, in the 128^3 RT runs on a $50\ {\rm Mpc}/h$ boxes used in Chapter 5, when $z \approx 15$ (at early redshifts the recombinations are more effective since density scales as $n \propto a^{-3}$), $t_{\rm rec} \approx 1{\rm Myr}$ in cells with overdensity $\Delta \approx 100$. The time between N-body snapshots at that redshift is $T_{\rm sim} \approx 15{\rm Myr}$. Assuming $N_{\rm sweeps} = 1500$, the time step in that RT run is then $\Delta t \approx 0.01{\rm Myr}$, therefore two orders of magnitude shorter than the recombination time in the highest density regions. At the same redshifts the cosmological I-fronts move slowly due to the high densities and small fluxes from the rare sources that drive them (see the I-front jump condition, Eq. (2.23), that can be used to estimate the I-front speed), meaning that only a few cells are ionized in $T_{\rm sim}$. At low redshifts, near the end of overlap, the I-fronts driven by many more sources may move through the voids at much lower density, therefore many more cells can be crossed in $T_{\rm sim}$. Even then, the number of cells crossed in $T_{\rm sim}$ is much lower

than our chosen $N_{\rm sweeps}$, which guarantees that no cell is fully ionized in a single Δt and the I-front propagation is governed by ionization time instead of the speed of light.

The use of Eq. (3.8) guarantees spatial photon conservation as it takes the optical depth of the cell into account, as we discussed in Section 2.5.1. The temporal photon conservation is also a natural consequence of our choice of Δt, discussed in the previous paragraph, since the ionization state x is updated every integration step of Eq. (3.13). Note also that the temporal photon conservation is not a large issue for very optically thick cells such as the ones typically found in reionization simulations (e.g. even at $z = 5$, the lowest redshift in our simulations, practically all cells in a 50 Mpc/h box have $\tau \gtrsim 100$), as we discussed in Section 2.5.2.

The number of sweeps is motivated by both physical time scales in the rate equation and the spatial transport of photon packets. In order to separate the two issues, we introduced a slightly different sweeping procedure. A single loop over sweeps has been separated into two loops. The outer one is a time stepping loop, during which new photon packets are emitted and recombinations are estimated. The embeded loop is used to propagate the existing photon packets, which are moved along the grid as long as either they have photons to deposit in the cells or they move as far as the speed of light allows[2]. The step of the outer loop is fully specified by physical time scales, e.g. the recombination time in the highest density cell, while the step of the inner one is computed internally to satisfy the described conditions. This removes $N_{\rm sweeps}$ as an input parameter. However, we found this more accurate procedure not to be necessary on the scale of our reionization simulations (tens of megaparsecs) because the speed of light limits the packets to crossing no more than a single cell per time step anyway for the typical $N_{\rm sweeps}$ and $T_{\rm sim}$ at all redshifts.

3.2.4 Comparison of SimpleX to other Cosmological RT methods

Before we proceed to the tests and improvements of the method, we will briefly discuss the reason why we chose to use SimpleX for simulating the evolution of the ionization fields during reionization in the first place, instead of one of many techniques devel-

[2]To estimate the maximum number of cells a photon packet is allowed to move, we use the global average Delaunay line length. Due to that, some photon packets may move faster or slower than the speed of light

oped for this purpose. Our prime concern was speed. The major issue in that regard is, as we already mentioned, the number of sources that is found in simulation boxes. The characteristics-based schemes (see Chapter 2 for more detailed examples) must loop over individual sources, thus having a computational cost that scales linearly with their number. Still, there are some optimised characteristics methods that alleviate this issue. For example, in Chapter 2 we already discussed the C^2-ray method (Mellema et al., 2006) that allows the use of very long integration time steps by properly taking photon conservation along a ray into account.

Another popular method is the adaptive ray-tracing of Abel & Wandelt (2002b) (used by e.g. Sokasian et al., 2001; McQuinn et al., 2007; Zahn et al., 2007; Trac & Cen, 2007). This technique is in essence a photon packet method. The speed of the scheme comes from the fact that the rays are followed only as long as there are photons in the packet initially cast along the ray. The eponymous adaptivity comes from the fact that rays can be split to improve the spatial sampling (same as we do in the SimpleX scheme) or merged where the high precision is not necessary. Even with all the improvements, the number of operations in these schemes still depends on the number of sources, though the scaling is not always linear. Furthermore, all the groups use uniform meshes to represent the density field, losing resolution in the high density regions.

The Monte Carlo codes (e.g. Maselli et al., 2003; Semelin et al., 2007; Altay et al., 2008) also scale with the number of sources. This scaling is only broken once the number of sources is high enough for one to claim that the radiation field is sufficiently sampled with some given number of photon packets emitted from all sources.

A more apt comparison for SimpleX should be methods that do not depend on the number of sources. TRAPHIC, a radiative transfer method implemented in Gadget-2 (Pawlik & Schaye, 2008), is to a large degree inspired by SimpleX. The main difference is that TRAPHIC does not use the Voronoi/Delaunay grid: photon packets are transported directly on SPH particles in Gadget-2. The packets are not sent along direct connecting lines from one particle to the next, but are instead emitted isotropically in equal spherical angle cones. This scheme does not suffer from the numerical diffusion that the SimpleX transport along Delaunay lines may introduce, but is more computationally expensive since the travel paths need to be recomputed during the calculation.

Another class of methods that we considered are the moment methods (e.g. Gnedin

& Abel, 2001; Petkova & Springel, 2009). While they are explicitly independent of the number of sources, the use of a single Eddington tensor in both optically thick and thin regimes necessarily leads to errors. The precision is much better if one computes the Eddington tensor for every cell explicitly (as in Finlator et al., 2009b), but this entails the use of ray-tracing methods, thus severely reducing the performance speed of the scheme.

In terms of computational speed, SimpleX compares very well to all of the previously mentioned codes. For example, while Iliev et al. (2006a) quote tens of thousands of CPU hours needed to perform a single reionization realisation, a run of comparable or better spatial resolution performed with SimpleX takes only a few hundred CPU hours. Comparable speeds can be obtained with approximate RT solvers designed to quickly produce reionization maps (e.g. Alvarez et al., 2009; Thomas et al., 2009), but their semi-analytic approach does not accurately represent the density field, which results in approximate morphologies of HII regions during reionization and, more importantly, the less precise calculation of recombinations. We concluded that SimpleX provides us with a very good compromise between speed and precision compared to other available methods.

3.3 Grid sampling

In Section 3.2.2 we argued that the main strength of the Voronoi/Delaunay graph pair in RT applications is that it provides connections between neighbouring cells, along which the radiation can travel. This is true irrespective of the actual choice of Voronoi grid point positions. Still, this choice can have important consequences because it determines the detail in which the density field data is resolved by the grid. In this section, we will discuss our choice of grid point sampling and compare it to some other possible choices. We will test the grid sampling in two cases: the Test 4 of the RT Code Comparison project (Iliev et al., 2006b) and on a number of test cosmological density fields we obtained from N-body runs.

The adaptivity of the Voronoi diagram comes from the complete freedom of choosing the position of grid points. We already gave an example of the grid choice in Section 3.2.2: using Eq. (3.7) the original SimpleX method created a grid that correspond to the mean free path in the medium. There we claimed that the high gradient in the number of

grid cells produced by Eq. (3.7) may result in unphysical propagation of radiation. This problem has been observed in the Test 4 of the Code Comparison project (Iliev et al., 2006b) and will be discussed in more detail in the following section.

Let us introduce here our preferred grid sampling technique. For our reionization simulations (see Chapter 5), we will obtain the density fields using N-body simulations, for which the outputs are just positions of equal mass particles. The easiest way to choose a Voronoi grid that corresponds to such density field data is simply to use the particles themselves as Voronoi cell nuclei. If a lower RT grid resolution is necessary[3], we choose a random subset of the particles. This sampling approach results in the following grid sampling relation:

$$n_{\rm D} \propto n_{\rm H}. \tag{3.14}$$

This is the same scaling the N-body technique has, i.e. the local density is represented by the particles found in that volume. We prefer this technique for two reasons. Firstly, when using Eq. (3.14) the spatial scales resolved in the N-body data are naturally mirrored by the RT grid. Secondly, we do not need to estimate the density field before constructing the RT grid. In the following section, we will show that this scaling is, in fact, a much better choice than Eq. (3.7).

3.3.1 Test 4 of the RT Code Comparison project

This test involves the propagation of ionization fronts from multiple sources through a static cosmological density field. The initial conditions are provided on a 128^3 uniform grid computed from a time slice of a cosmological N-body and gasdynamic simulation performed using the PM+TVD code (Ryu et al., 1993). The simulation box used is a cubic region with side length of 0.5 Mpc/h populated by dark matter and hydrogen gas. The 16 most massive halos identified in the box are used as sources of ionizing radiation. Ionizing luminosities were assigned to the halos by assuming a linear mass-luminosity relation:

$$\dot{N}_\gamma = f_\gamma \frac{M\Omega_{\rm b}}{\Omega_0 m_p t_s}, \tag{3.15}$$

[3]This will be almost always the case in the reionization simulations we present in Chapter 5. High resolution N-body runs are needed to produce the dark matter merger trees that are used to model the sources of reionization, but no RT method can match that resolution due to RT being a much more computationally expensive problem then gravity.

where f_γ is an efficiency parameter, M is the total halo mass, Ω_b and Ω_0 are baryon and dark matter density parameters, m_p is the proton mass and t_s is the lifetime of the source, chosen to be 3 Myr. The efficiency parameter, f_γ, defines how many photons are produced per hydrogen atom in the halo during its lifetime. The codes taking part in the test were expected to follow the time evolution of the ionization and temperature fields for 0.4 Myr and provide outputs at 5 chosen times.

Figure 3.3 compares the results we obtained with our version of SimpleX against the SimpleX used in the original Code Comparison project and another code, C^2-ray (Mellema et al. 2006). Perfect agreement between SimpleX and other codes cannot be expected, since all the other codes ran the problem with multi-frequency transport and self-consistent ionization gas heating, while SimpleX assumed a temperature of 10^4 K in ionized regions and approximated the hardness of the spectrum through the use of a single frequency bin and a spectrum-averaged ionization cross-section:

$$\langle\sigma\rangle = \frac{\int_{\nu_{\rm th}}^{\infty} \sigma_\nu I_\nu {\rm d}\nu}{\int_{\nu_{\rm th}}^{\infty} I_\nu {\rm d}\nu}, \tag{3.16}$$

where σ_ν was defined in Eq. (2.15) and I_ν is the intensity of the radiation source (assumed to be a blackbody with effective temperature $T_{\rm eff} = 10^5\,{\rm K}$, so the intensity is given by the Planck law: $I_\nu = (2h\nu^3/c^2)(\exp({\rm h}\nu/{\rm kT_{eff}}) - 1)^{-1}$). Regardless of these differences, general shape of the HII regions should be similar. Radiation originates in dark matter halos located in the peaks of the density field. Once the I-front breaks out of the high density filaments, it starts expanding much faster in the lower density voids (recall that the I-front speed, Eq. (2.23), depends on the local density). This results in I-fronts having a typical "butterfly" shape, as their propagation speed stays much lower in the filament than in the voids. In Fig. 3.3, all codes demonstrate this I-front behaviour, but it is very clear that the front speeds in voids are lower in the original SimpleX (blue line) than in C^2-ray (red line). This is a direct consequence of undersampling the low density regions through the use of Eq. (3.7). If there is an insufficient number of vertices in a region, there are not enough potential routes for radiation packets to reach it. Since there are many more connections in the high density regions, the radiation will preferentially stay there, somewhat bypassing the lower density ones and leading to the delay of the I-fronts. Also, the more jagged shape of the fronts in the voids is due to the violation of radiation isotropy because of the lack of potential radiation travel routes.

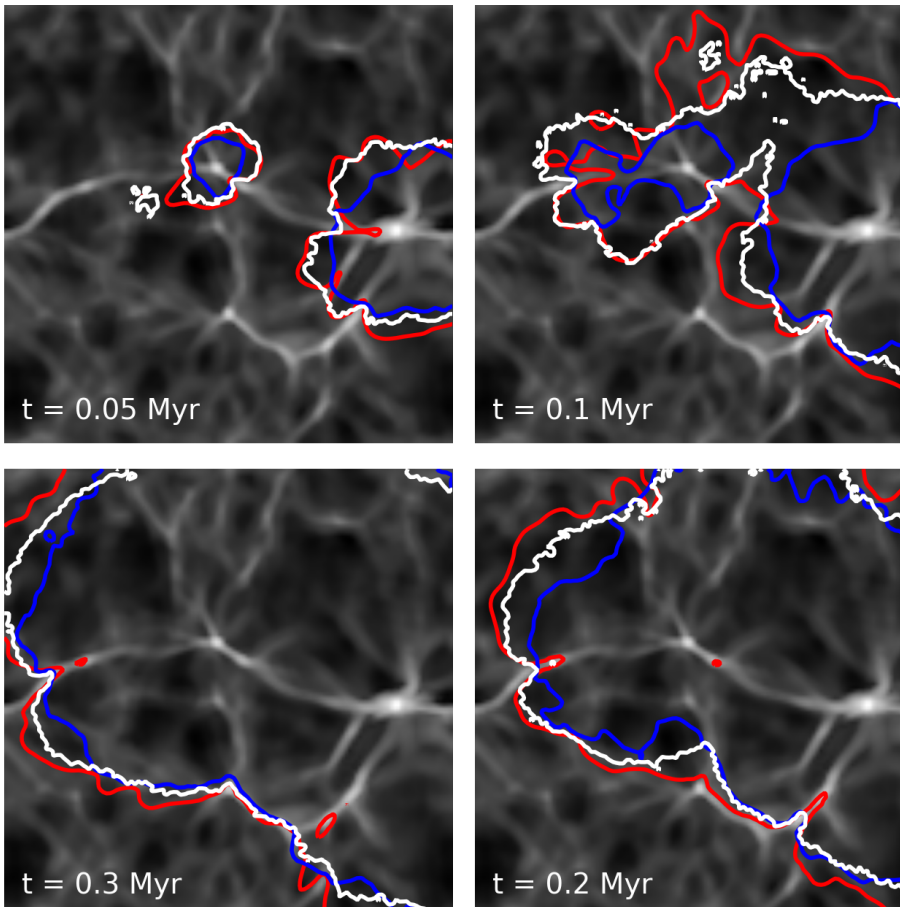

Figure 3.3: Slices through the middle of the Test 4 simulation box (see text for test description) at 4 output times: 0.05, 0.1, 0.2 and 0.3 Myr (clockwise from top left). The grey-scale image in the background is the density field, $\log_{10} n_{\rm H}$. Contours represent the $x = 0.7$ ionized fraction boundary from three different runs: C^2-ray code (red line), SimpleX with Eq. (3.7) sampling function (blue line) and SimpleX with direct point sampling, Eq. (3.14) (white line). The blue lines under-sample the low density regions resulting in delayed propagation of I-fronts in voids. The more spatially uniform SimpleX sampling (white lines) results in much better agreement with C^2-ray and the other codes that took part in the Code Comparison project.

The white lines in Fig. 3.3 show the results obtained with our current version of SimpleX. Before discussing the results we will present the sampling technique we employ and how it differs from the technique used by original SimpleX. The main difference is

in the way the density field is supplied at input. The original SimpleX scheme uses the uniformly gridded density values. Our version, on the other hand, uses the direct output of N-body simulations - positions of equal mass particles. From this data, we reconstruct the density field on the Voronoi grid in several steps. First, we choose a desired size subset of all particles to be Voronoi nuclei, where the probability of choosing any particle is weighted by some function of density, $W = f(\rho)$. Our preferred sampling technique, described earlier in this section, Eq. (3.14), is given by the weighting function: $W = 1$, i.e. a uniform (unweighted) selection of particles. Secondly, using an octree-based neighbour search, we assign the mass of all the points we did not select to their closest neighbours among the ones that we did select. Remembering the definition of the Voronoi cell, Eq. (3.3), it is clear that this step corresponds to the nearest grid point interpolation on the Voronoi grid. Finally, the density of every grid cell is computed after the full triangulation is set up and it is simply equal to the total mass assigned to the Voronoi nucleus divided by the volume of the corresponding Voronoi cell. This procedure is equivalent to the density reconstruction scheme called Voronoi Tessellation Field Estimator (Schaap, 2007), which was found to be a much more precise technique for density field reconstruction than the commonly employed SPH (Monaghan, 1992) or grid-based (Hockney & Eastwood, 1981) schemes.

The original SimpleX procedure for setting up the density field on the Voronoi grid is somewhat different. First, given a uniform grid of density values, Voronoi nuclei are created through a correlated point process. A random number generator is used to generate point coordinates. The probability for a point to be kept as a Voronoi nucleus is weighted by the density of the uniform grid cell in which it falls, using Eq. (3.7). This procedure is repeated until the required number of points is kept. After the Voronoi diagram has been generated, the value of density of every Voronoi cell is obtained by linear interpolation from the density values of the neighbouring uniform grid cells. Contrary to our method, this procedure operates with densities and not masses, thus it is not explicitly mass-conserving (i.e. the interpolation of densities from one grid to another does not guarantee that the total mass on both grids is going to be the same).

For Test 4, the original SimpleX has used the default grid assignment method described in the previous paragraph. Since the data for the test is supplied on a uniform grid, we were not able to employ our default method. To bridge the gap, we generated

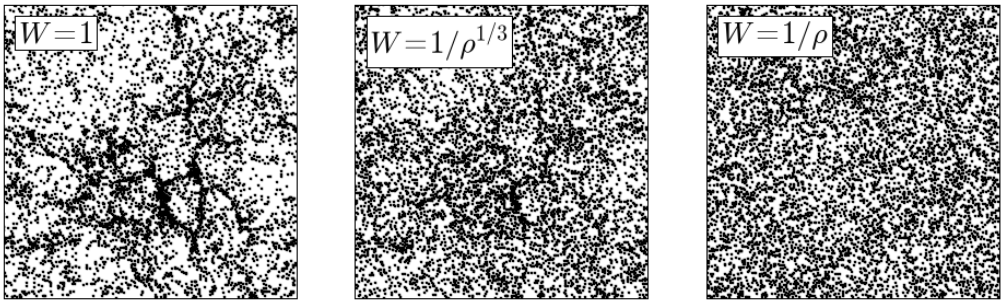

Figure 3.4: Different sets of Voronoi nuclei representing the same N-body run. Weighting functions are given in top-left corners of each panel. See text for discussion.

the point data from the supplied grid. In every uniform grid cell, we generate a number of points proportional to the density in that cell. We require that there be a minimum of one point per cell. We also assign mass to every point equal to the mass in the parent grid cell divided by the number of points spawned in that cell. Using these points, we proceed with our default technique: we choose the subset of points to be the Voronoi nuclei, allocate left-over mass to nearest nuclei and compute the density once the Voronoi cell volumes are computed.

Coming back to Fig. 3.3, we see that our sampling scheme (white line) produces a much better agreement with other codes (represented by C^2-ray, red line) than the original SimpleX (blue line). This is entirely due to more Voronoi nuclei being placed in the voids. We can conclude that a simple change in the sampling procedure is all it takes to fix the perceived inadequacies of SimpleX radiative transfer in cosmological density fields. The same conclusion was reached by Paardekooper et al. (2009), who modified the original, grid-based sampling procedure with sampling functions similar to the ones we used here.

To make sure that our preferred sampling procedure places a sufficient number of points in the low density regions, we compare the result presented in Fig. 3.3 with the results obtained from Voronoi grids created by two other weighting functions chosen to sample the low density regions in even more detail: $W = 1/\rho^{1/3}$ and $W = 1/\rho$. The difference between Voronoi nuclei sets created with these weighting functions is illustrated in Fig. 3.4. This figure shows the range of possible samplings our weighting functions

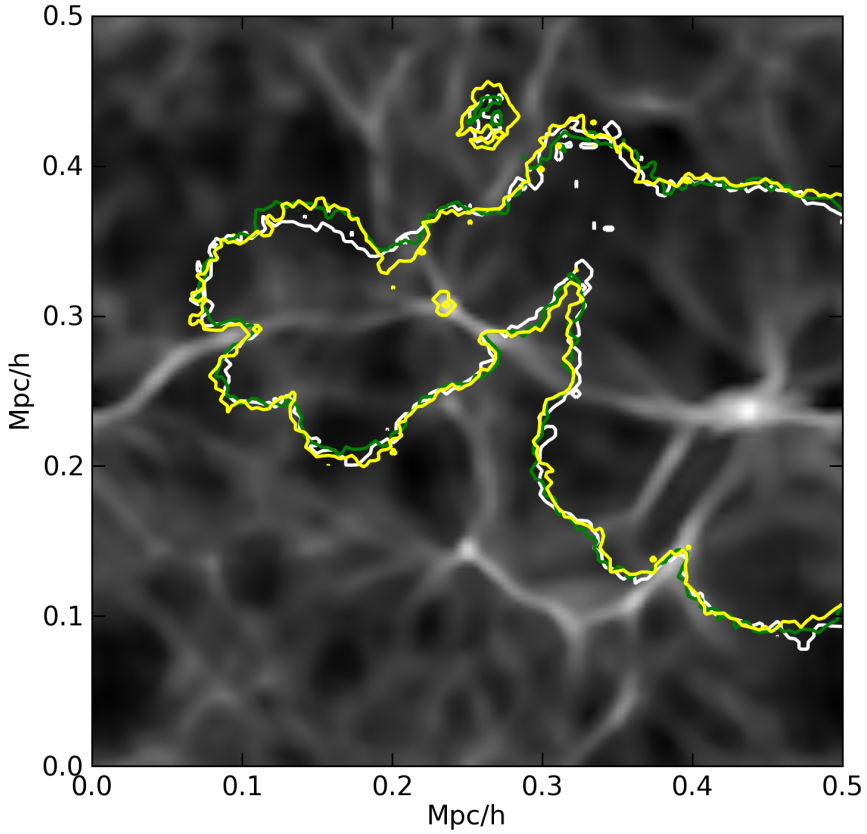

Figure 3.5: The effect of different sampling on the Test 4 results. Shown is a slice through the middle of the density field (grey scale) with contours representing the $x = 0.7$ ionized fraction boundary at time t = 0.1 Myr for runs with different samplings: W = 1 (white line), W = $1/\rho$ (yellow line), W=$1/\rho^{1/3}$ (green line) with the same grid resolution, 64^3 cells. Using more spatially uniform sampling than $W = 1$ does not produce any significant changes in I-front position, suggesting that $W = 1$ adequately resolves the low density regions.

cover: while $W = 1$ is proportional to the local density, the $W = 1/\rho$ is inversely proportional to the density which results in a constant number of vertices per unit volume.

We repeated the Test 4 runs with these two grids and we present the results in Fig. 3.5. We see that further "smoothing" of the Voronoi nuclei spatial distribution does not result in any significant change in the position of I-fronts at any time. This leads us to conclude

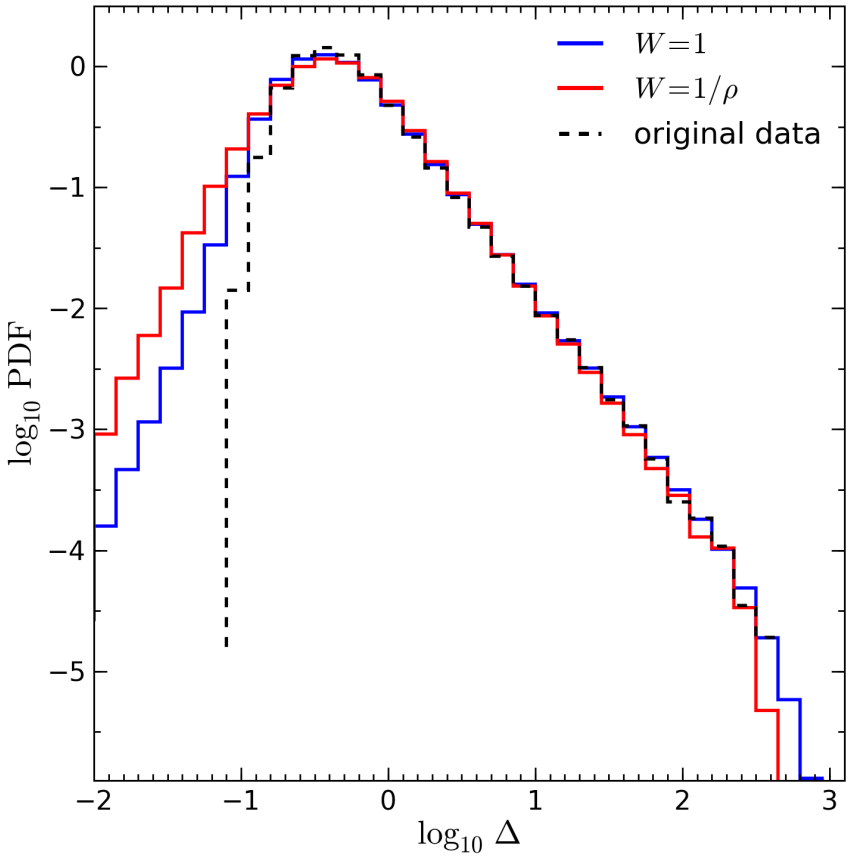

Figure 3.6: Comparison of volume-weighted density PDFs of differently sampled realisations of the Test 4 density field. The PDFs of fields sampled with $W = 1$ and $W = 1/\rho$ are shown as blue and red histograms, respectively. The PDF of the original density field supplied for the Code Comparison project is given for reference (dashed black line). See text for details about the sampling procedures. Even though the use of different point sampling can affect the highest or lowest density regions, it should not make a significant difference in the Test 4 case because the original data is sufficiently resolved (as seen in Fig. 3.5).

that our preferred sampling procedure, which reproduces the Eq. (3.14) relation, already sufficiently resolves the low density regions in Test 4.

The agreement between differently sampled Voronoi grids presented in Fig. 3.5 is not enough to put the Test 4 to rest, as it might be possible that all those Voronoi grids

represent the original density field in the same, yet incorrect way. To show that this is not the case, we present the volume-weighted probability density functions (hereafter, PDF) of the cell densities in the original uniformly gridded data and two Voronoi grids (Fig. 3.6). At a first glance, the Voronoi grids resolve both higher and lower densities than the original data provides. This is an unavoidable consequence of using the Voronoi grid. Even the $W = 1/\rho$ grid that was, as we discussed previously, constructed to be as uniform as possible, still does not have equal volume cells. This results in densities that differ from the values supplied on the uniform grid. On the $W = 1$ grid, the number of vertices is proportional to the local density, so the cell volumes exhibit the same variance as the density field. Even though the number of grid points is the same for Voronoi grids (128^3 in both), the spatial resolution is not the same and this leads to the difference in the cell density PDFs. The spatial sampling also causes the differences between the two Voronoi grids. Still, the agreement in PDFs is excellent in the density range that all grids resolve.

The largest discrepancies between the original data and the Voronoi grids are in the low density regions. These differences should not affect the obtained I-front positions in a Test 4 run. For one, the low density regions have a smaller fraction of total gas to be ionized, thus the loss of ionizing photons in those regions will not be as important as in the high density regions. The same is true for the number of recombinations in a cell. Finally, the low density voids are the final regions to be ionized in a cosmological volume[4], so even if some discrepancies do arise due to the density field differences they will be relevant only in the final overlap stages (Gnedin, 2000a), when reionization is practically complete. Taking all this into account, Fig. 3.6 leads us to conclude that the density field is sufficiently well represented by the Voronoi grids we used in our Test 4 runs.

3.3.2 Sampling effects on N-body data runs

Now that we have shown that SimpleX with our sampling procedure can obtain satisfactory results in the Test 4 of the Code Comparison project, we want to make sure that the

[4]We find that reionization proceeds completely "inside-out", so that the highest density regions are the first to be ionized and low density voids are the last. Note that this is not a universally agreed-upon scenario (e.g. Finlator et al., 2009a). This issue is discussed in more detail in Chapter 5

Simulation	$L_{\rm box}$ [Mpc/h]	$N_{\rm DM}$	$m_{\rm DM}$ [$10^5\,{\rm M}_\odot$/h]
L12.5N128	12.5	128^3	646.2
L20N512	20	512^3	41.35
L10N512	10	512^3	5.17
L10N1024	10	1024^3	0.65
L20N1024	20	1024^3	5.17

Table 3.1: Parameters of the N-body runs used in Chapter 3. See text for the description of the labelling scheme. The stated values are as follows: size of the computational box, $L_{\rm box}$; number of dark matter particles, $N_{\rm DM}$; mass of dark matter particles, $m_{\rm DM}$. The Cosmological parameters [Ω_0, Ω_Λ, $\Omega_{\rm b}$, σ_8, n_s, h] used for all runs are [0.25, 0.75, 0.045, 0.9, 1, 0.73]. The runs marked by bold names are used in the single snapshot tests of grid sampling effect on RT results (Section 3.3.2)

same procedure is appropriate for use in the context of the reionization runs we intend to perform. As already mentioned in the previous section, we intend to compute the density field on the Voronoi grid directly from the N-body simulation particles. In order to test the effects of sampling, but also of grid resolution, we perform runs on $z = 5$ snapshots from two N-body runs, L12.5N128 and L20N512. The labelling of the runs is as follows. Letter "L" means "length" and the following numerical value gives the side length of the simulation box in Mpc/h. The letter "N" marks the number of particles used in the run and the numerical value is the cubic root of that number, i.e. if a simulation uses 512^3 particles, it is marked as N512. All the N-body runs in this section have been performed using the "lean" version of Gadget-2 (Springel, 2005a). More details about the simulations are given in Table 3.1.

The L12.5N128 box is chosen mainly for quick test runs, since a representative Voronoi grid can be constructed quickly. The L20N512 box is chosen because it has approximately

the same particle mass as our target simulation for the large scale reionization runs, the Millennium-II simulation (Boylan-Kolchin et al., 2009). The typical RT resolutions we employ with SimpleX are 64^3 and 128^3 RT cells, so only a sub-section of L20N512 particles will be used to construct the RT Voronoi grid, as discussed earlier. For the tests in this section, we use the $z = 5$ snapshots, because the density contrasts are highest at that time and all inadequacies due to sampling or grid resolutions should be most pronounced. The friends-of-friends algorithm (Davis et al., 1985) was employed to identify dark matter halos. We use 16 and 37 of the most massive halos found at $z = 5$ in the L12.5N128 and L20N512 runs, respectively, as the sources of ionizing radiation, the numbers chosen to make these tests comparable to Test 4 presented in the previous section. We used Eq. (3.15) to assign ionizing luminosities to the chosen halos. The efficiency parameter, f_γ, was not the same in all runs, but varied with the sole intention of producing a significant change in the ionized fraction within the time chosen for each specific run. In simple words, we change the luminosity to get the whole box ionized in an arbitrarily chosen time. For these runs we were not concerned with reproducing the correct physical conditions as we only wanted to study the numerical effects stemming from the properties of the Voronoi grid.

We first look at the effect of the choice of Voronoi nuclei positions on the radiative transfer results. The left panel of Fig. 3.7 shows the results of RT runs on Voronoi grids created with three sampling functions introduced in the previous section (Fig. 3.5) in the L12.5N128 box. All RT grids have 64^3 Voronoi nuclei, i.e. every eight particle was used as a cell nucleus. All sources have the efficiency parameter of $f_\gamma = 25000$ and the total simulation time was 1 Myr[5]. With these parameters, the total number of ionizing photons produced is ≈ 2 times larger than the total number of atoms in the box. The I-fronts move much faster on grids that have less grid cells in high and more in the low density regions. The change of Voronoi grid resolution, i.e. the increase of the total number of grid cells, shows a similar effect as seen in the right panel of Fig. 3.7. The lower resolution runs produce faster I-fronts than the higher resolution ones.

These result seems to contradict the conclusion of the previous section where we

[5]The running time is much shorter than the recombination time at the average density at $z = 5$, but the recombinations are still an important factor, as evidenced by the fact that the box is not fully ionized even if more than enough photons have been emitted.

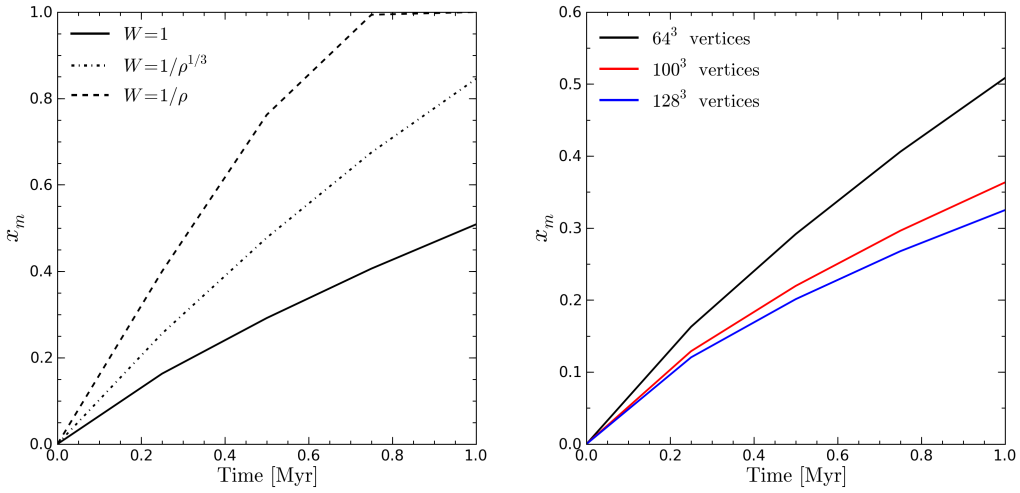

Figure 3.7: The effect of sampling on the L12.5N128 test run results (see text for run setups). *Left:* The effect of Voronoi nuclei placement on the evolution of the mean mass-weighted ionized fraction, $x_{\rm m}$. Line styles represent different sampling function weights: W = 1 (solid line), W = $1/\rho$ (dashed line), W=$1/\rho^{1/3}$ (dot-dashed line). All grid have 64^3 cells. More spatially uniform grids produce significantly faster I-fronts. *Right:* The effect of the number of grid points on the evolution of $x_{\rm m}$. Line colors represent different Voronoi grid resolutions: 64^3 (black line), 100^3 (red line) and 128^3 (blue line). All points sampled with the same weighting function ($W = 1$). The decrease of RT grid resolution results in the same effect as the spatial "smoothing" of cell numbers per unit volume shown in the left panel.

found that all the differently sampled grids we are testing produce the same RT result, as shown in Figure 3.5. The reason for this discrepancy is found in a much smaller importance of recombinations in Test 4, compared to the tests discussed here. For example, the difference in the total recombination rate on 64^3 RT grids sampled with $W = 1$ and $W = 1/\rho$ from the Test 4 density field data is about 30% in favour of the former. On the other hand, for the same grids sampled from L12.5N128, $W = 1$ has an ≈ 10 times larger recombination rate than $W = 1/\rho$. In the case of Test 4 density data, the importance of recombinations is probably smaller because the details of the density field have been smoothed out on the Cartesian grid on which it is supplied. Note also that the density

field for Test 4 was obtained from a hydro simulation, meaning that the Jeans smoothing may have also played a role.

Figure 3.7 showed that both the resolution and sampling of the Voronoi RT grid can have a large effect on the final result. The main difference between the runs shown in both panels of Fig. 3.7 is the spatial size of cells in regions of different densities. As we pointed out earlier, a grid chosen with a $W = 1/\rho$ weighting function will have approximately equal-sized cells in the whole volume, thus having a much lower spatial resolution in high density regions than the $W = 1$ grid, which adapts to the density field and has approximately constant cell mass. Equally, a larger number of RT grid cells (right panel of Fig. 3.7) results in higher spatial resolution in all regions. Note that we emphasise the high density regions: these are the regions where the bulk of recombinations are produced, as the rate depends on the square of the local density (see the recombination term in Eq. (3.10)). This means that insufficient grid resolution in high density regions can result in a significantly underestimated total number of recombinations.

To confirm our suspicion that the RT grids are not resolving sufficiently well the recombination rate on the provided N-body density fields, leading to the large discrepancies seen in Fig. 3.7, we repeat the RT runs on Voronoi grids created with different samplings, but this time without including recombinations in the calculation. We ran on two grids that showed the biggest difference in Figure 3.7, $W = 1$ and $W = 1/\rho$. We decreased the efficiency parameter of sources to $f_\gamma = 10000$ in order to avoid having the box completely ionized too quickly. We show the results of those runs in Figure 3.8. Once the recombinations are neglected, different grid samplings give virtually indistinguishable results. Figure 3.8 confirms that it is not the position of grid cells that introduces the differences in the RT results. Instead, they are the result of the unsatisfactory representation of the underlying density field. More precisely, the lack of spatial resolution smooths the density field thus decreasing the number of recombinations produced on the grid. The following section discusses how to include the unresolved details from the N-body data to the RT grid.

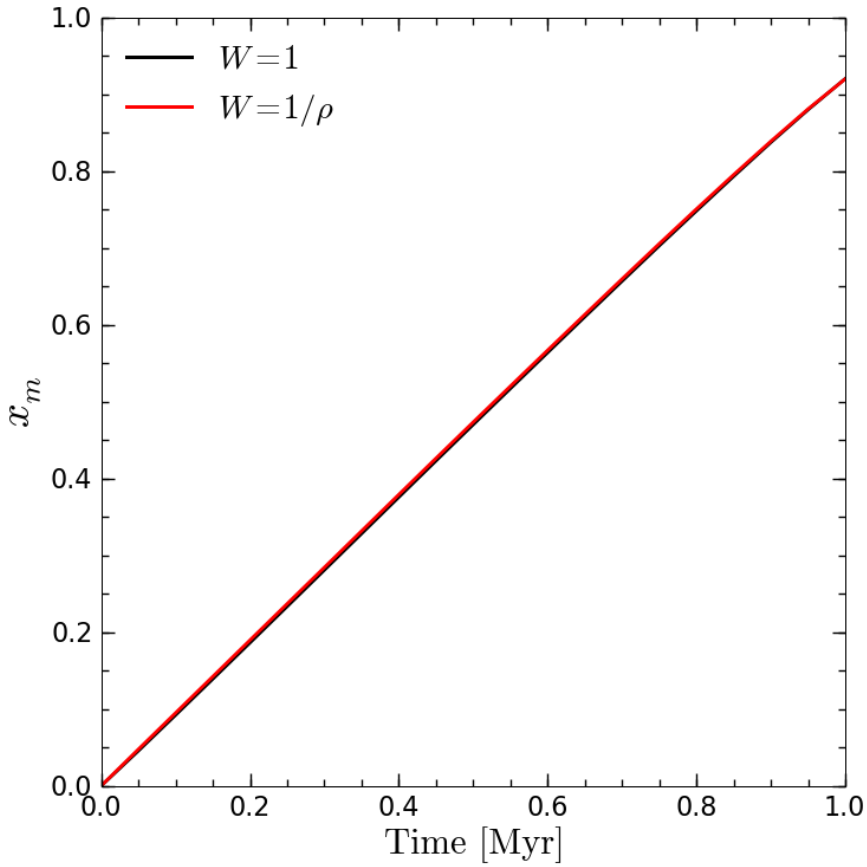

Figure 3.8: The evolution of the mean ionized fraction in the L12.5N128 test run without recombinations. The number of RT cells is 64^3 in both runs. Two lines represent two different sampling function weights: $W = 1$ (black solid line) and $W = 1/\rho$ (red solid line). This figure clearly demonstrates that the position of RT grid cells is not the determining factor in the evolution of I-fronts.

3.4 Representing the density field

As we suggested in the previous section, the representation of the density field on the computational grid can have a severe influence on the simulated propagation of I-fronts. In this section, we describe a simple technique to include all the available density field data from N-body simulations to the RT grids, not only for SimpleX but for any RT scheme in general. First, in Section 3.4.1, we discuss the backdrop for reionization, the intergalactic medium (hereafter, IGM) and its representation using N-body simulations.

Next, we introduce the tool used to represent the sub-cell density field, the clumping factor, in Section 3.4.2. We discuss its usage by other groups, how we compute it from N-body data and, most importantly, how it can be used to improve convergence of RT results. In Section 3.4.3, we revisit the single snapshot tests performed in Section 3.3.2 and show how they are improved through the use of the clumping factor.

3.4.1 IGM from N-body simulations

The largest part of the baryons found in the Universe fills the space that separates the galaxies, creating the *intergalactic medium* (IGM; see Madau, 2000; Meiksin, 2007, for recent reviews of the subject). It was the IGM that was reionized during the eponymous epoch that is the main subject of this thesis. It is then clear that we must understand the density distribution of the IGM if we are to simulate the process of its ionization.

Most of the information about the intergalactic gas comes from studying the *Lyman alpha forest*, a pattern of absorption lines associated with structures in the IGM, that is observed in the spectra of high redshift quasars (see e.g. Rauch, 1998; Bechtold, 2003, for detailed reviews of the subject). Computer models (e.g. Cen et al., 1994; Hernquist et al., 1996; Theuns et al., 1998) successfully connect the Lyman alpha forest features and the large scale IGM density field with the paradigm of hierarchical structure formation.

In this picture, on the largest, linear scales, it is safe to assume that the baryons closely follow the distribution of dark matter as on those scales gravitational dynamics dominate over the thermodynamics of the gas (e.g. Zhang et al., 2004). This assumption allows us to use the N-body simulations (i.e. dark matter only) to estimate the IGM density field, as is commonly done in large scale reionization numerical work (e.g. Ciardi et al., 2000; Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007). Still, one must be aware that the N-body simulations do not provide an accurate representation for the gas on smaller scales. For one, as gravitational collapse proceeds, the gas pressure forces become more important and decouples it from dark matter as Jeans length scales are approached (a characteristic length at which the gravitational force is opposed by the pressure force, see e.g. Peebles (1993), mentioned in Chapter 1). The Jeans smoothing is also increased by the photo-ionization heating of the gas during reionization (e.g. Pawlik et al., 2009). The hydrodynamical effects are usually ignored in current large-scale simulations (e.g. Iliev et al., 2007; McQuinn et al., 2007; Shin et al., 2008).

The gas that has collapsed into dark matter halos makes up the rest of the cosmic baryons, but it is not straightforward to include it in simulations. For one, gas in halos with $T_{\rm vir} \gtrsim 10^4 K$ is able to efficiently cool through atomic line transitions which eventually leads to the formation of galaxies. In that case, the assumption that baryons trace the dark matter fails completely. Furthermore, the virialized structures take up only a small fraction of the typical computational cell volume, meaning that most of the ionizing photons passing through the cell will not interact with them. This is, of course, not true for the ionizing photons produced in the halos themselves. Since it is impossible to properly model the breakout of photons from their sources in large scale simulations needed for reionization study, this effect is included with a parameter called the *escape fraction* ($f_{\rm esc}$ for short), the fraction of photons produced by sources within a halo that escapes into the IGM. The escape fraction can be obtained from both higher resolution simulations (e.g. Razoumov & Sommer-Larsen, 2007; Gnedin, 2008; Wise & Cen, 2009) and observations (e.g. Leitherer et al., 1995; Bland-Hawthorn & Maloney, 1999; Heckman et al., 2001; Bergvall et al., 2006; Shapley et al., 2006), but its exact value during reionization remains largely unknown. These arguments suggest that gas in halos should not be considered part of the density field in a reionization simulation. Therefore, we neglect the N-body particles associated with dark matter halos when constructing the density fields for the reionization RT runs in Chapter 5. We single out those particles by choosing a threshold overdensity a particle associated with the IGM is allowed to have, as we discuss in more details in the following sections. We will, however, demonstrate the role of halo particles on the propagation of I-fronts in the following discussion.

Finally, the necessary N-body resolution is given by the need to resolve the halos that can host sources of ionizing radiation (generally taken to be halos with $T_{\rm vir} \gtrsim 10^4$ K). Smaller so-called "minihalos", where the gas is not capable of atomic line cooling, can have a large effect on both the time evolution of the ionized fraction (e.g. Ciardi et al., 2006) and the shape of the HII bubbles (e.g. Furlanetto et al., 2004; McQuinn et al., 2007). The issue of minihalos is complicated due to the fact that the gas can be photoevaporated from their shallow potential wells so more complex modeling is required to gauge their true importance (Shapiro et al., 2004; Iliev et al., 2005b). Minihalos may or may not be the main source of the Lyman limit systems (LLS for short) observed in the Lyman alpha forest, absorption features that owe their name to the fact that they are optically

thick to Lyman limit radiation, $\nu_{\rm th}$. They correspond to column densities of $10^{17.2}{\rm cm}^2 \lesssim N_{\rm HI} \lesssim 10^{20}{\rm cm}^2$. These still poorly understood objects (Tytler, 1982; Sargent et al., 1989; Prochaska, 1999), and to a lesser extent the even more optically thick damped Lyman-α objects (reviewed by Wolfe et al., 2005), determine the mean free path of ionizing photons in the post-reionization Universe (Haardt & Madau, 1996; Miralda-Escude, 2003) and may also be important during the EoR itself. All these features must be included in the simulation if one wants to claim to have a proper representation of the density field.

3.4.2 Resolving the recombinations with clumping factors

As we discussed in Section 3.4.1, an accurate representation of the IGM for reionization simulations is not a trivial problem. Here, we focus on a single aspect - the recombination rate, which depends on the gas distribution (the self-shielded structures such as LLS or minihalos do not affect the ionizing photon budget in the same way). In all our reionization simulations, to be presented in Chapter 5, we obtain the IGM from N-body simulations by assuming that the gas is a perfect tracer of the dark matter. Still, as pointed out in Section 3.3.2, due to the computational cost constraints RT grids may not always match the spatial and mass resolution provided by the N-body runs. This results in a "smoothing" of the density field by its representation on the lower resolution RT grid. In this section we focus on how the effect of this smoothing on the recombination rate can be alleviated.

What is the clumping factor?

The recombination rate (number of recombinations per unit time) in a single computational cell (or any finite volume for that matter) in a hydrogen-only medium is (right hand side term in Eq. (3.10)):

$$\dot{N}_{\rm rec} = \alpha n_{\rm HII}^2 V, \tag{3.17}$$

where V is the volume of the cell, $n_{\rm HII}$ is the ionized number density of hydrogen in the cell and α is the recombination rate coefficient (could be either case A or B). What one must always remember is that the values of cell variables are actually meant to represent a *continuous field*. The discrete cell values are averages of that field over a finite volume

of the cell, V:

$$n_{\rm HII} = \langle n_{\rm HII} \rangle = \frac{\int_V n_{\rm HII}(\vec{r}) {\rm d}^3 r}{V}, \tag{3.18}$$

where $n_{\rm HII}(\vec{r})$ is the value of the continuous ionized number density field at position $\vec{r}$ and the $\langle\rangle$ brackets denote an average over the volume V. We rewrite Eq. (3.17) to explicitly show what is computed on a discrete grid cell:

$$\dot{N}_{\rm rec}^{\rm dis} = \alpha \langle n_{\rm HII} \rangle^2 V. \tag{3.19}$$

A discrete grid representation is adequate for evaluating physical quantities that depend on the continuous density field *linearly*. For example, to get the number of ionizations per unit time in the volume V:

$$\dot{N}_{\rm ion} = \Gamma \int_V n_{\rm HI}(\vec{r}) {\rm d}^3 r = \Gamma \langle n_{\rm HI} \rangle V. \tag{3.20}$$

In other words, using the mean density in the cell, $n_{\rm HI}$ is equivalent to integrating the continuous density field to get the number of ionizations in volume V. This result implies that the number of ionizations computed on a discrete grid does not depend on the grid construction, i.e. the choice of V for each cell, as shown in Fig. 3.8.

The recombinations, on the other hand, depend on the *square* of the density. Given a density field, the correct recombination rate in volume V, obtained by integrating the field, is:

$$\dot{N}_{\rm rec}^{\rm cont} = \alpha \int_V n_{\rm HII}^2(\vec{r}) {\rm d}^3 r = \alpha \langle n_{\rm HII}^2 \rangle V. \tag{3.21}$$

A comparison of Eq. (3.21) with (3.19) shows that using the mean value of the density in the volume V does not produce the correct number of recombinations in that volume (to be more precise, it does so only when $V \to 0$). The same comparison also reveals that $\dot{N}_{\rm rec}^{\rm cont} = C \dot{N}_{\rm rec}^{\rm dis}$ where:

$$C = \frac{\langle n_{\rm HII}^2 \rangle}{\langle n_{\rm HII} \rangle^2}. \tag{3.22}$$

The linear factor C is the *clumping factor*, computed over the volume V.

We have shown that, in principle, one should be able to compute the correct number of recombinations on a discrete grid irrespective of the grid construction, if the clumping factor in every cell is known. Note that the number of recombinations obtained will be only as correct as the continuous density field data. This density field is provided by other numerical simulations (the dark matter-only N-body runs in our case), which

are themselves limited in resolution and included physics, as mentioned in Section 3.4.1. Therefore, we should not talk about the "correct", but instead "converged" recombinations - our goal is to compute the same number of recombinations for a given N-body density field, irrespective of the RT grid used to represent it. In the next section, we show how we compute the clumping factor from the N-body simulations.

Computing the clumping factor from the N-body data

To compute the clumping factor using Eq. (3.22), we must first estimate the continuous density field from the N-body outputs, the positions of equal mass particles. There are a number of techniques one can employ for this purpose. Very well known examples are the grid techniques, where the mass of every N-body particle is redistributed among regularly spaced grid cells according to some weighting function and the density obtained by dividing the assigned mass with the volume of the cell (Hockney & Eastwood, 1988). While grid techniques are easy to use, they are not appropriate for density fields with a large dynamical range. The cell size must match the smallest scales resolved by the N-body run which results in too many cells placed in lower density regions, many of them remaining empty (we mentioned this issue as an argument for AMR grids, see Chapter 2).

A more efficient method for cosmological density fields are smoothed particles hydrodynamics techniques (see Monaghan, 1992, for review and further references). This is a particle-based scheme for solving the hydrodynamics equations and the density field at an arbitrary point can be recovered from the set of particles representing the fluid. This is accomplished by smoothing the mass of every particle within some 3D compact kernel, usually a spline function, but many variations exist. The density at the point of interest is computed as the sum of contributions from all the particles in its smoothing region:

$$\rho(\vec{r_i}) = \sum_{j=1}^{N} m_j W(\vec{r_i} - \vec{r_j}; h_i). \tag{3.23}$$

Here N is the total number of particles in the smoothing region, W is the smoothing function and h is the smoothing length. The contributing particles j can be either all particles that are within h_j of particle i (the "scatter" approach) or all particles j that are within h_i of the particle i (the "gather" approach). We use a "gather" SPH interpolation

scheme, as in Gadget-2 (Springel, 2005a), to calculate the density field at with every N-body particle in our simulations[6].

In Section 3.3 we mentioned that our sampling procedure incorporates the density estimation technique (VTFE from Schaap, 2007). In principle, it would be more consistent to estimate the density field for clumping factor calculations with the Voronoi grid. The main reason we employ the SPH technique instead is that, in our current implementation, the creation of the tessellation demands a very large amount of computer memory. Construction of the Voronoi grid with grid resolution matching the number of particles in even the intermediate N-body runs (e.g. 512^3 particles) is beyond the machines we had available for this project. The tessellation can be created in chunks to use less memory and we intend to use such a technique in future work. For all the runs presented in this thesis, we computed the clumping factor with the SPH densities. We have compared the SPH densities to Voronoi diagram-computed densities on smaller N-body runs. The density results and the corresponding clumping factors differ no more than by a factor of a few. We conclude that the use of SPH densities to compute the clumping factors is adequate.

We evaluate the integrals over the volume in Eq. (3.22) as the sum over the values at N-body particles found in the volume V, $\int_V \rho(\vec{r})\mathrm{d}^3 r \approx \sum_{i=1}^{N_{\rm part}} \rho_i V_i$. Equation (3.22) thus becomes (note that in hydrogen-only medium $\rho = m_H n_H$ and m_H terms cancel out):

$$C = \frac{\frac{\sum_{i=1}^{N_{\rm part}} \rho_i^2 V_i}{\sum_{i=1}^{N_{\rm part}} V_i}}{\left(\frac{\sum_{i=1}^{N_{\rm part}} \rho_i V_i}{\sum_{i=1}^{N_{\rm cell}} V_i}\right)^2}. \tag{3.24}$$

Here, $N_{\rm part}$ is the number of particles found in volume V, ρ_i are SPH densities at N-body particle positions and V_i is the SPH volume associated with particles. The volume V_i can be estimated in two ways: $V_i = m_i/\rho_i$ or $V_i = 4\pi/3h_i^3$, where m_i is the mass of the particle and h_i is the SPH smoothing length at the particle's position. We chose to use the former definition because it allows us to write Eq. (3.24) in a more concise way, as we will show shortly. The actual choice does not affect our results as both definitions yield similar volume estimates. For the N-body simulations, all the particles have the same

[6]SPH densities are automatically calculated in Gadget-2. However, for the N-body runs we used the "lean" version of Gadget-2, which solves the gravity-only problem and does not compute the densities. This is why we had to compute them in post-processing, using an SPH density calculator based on the original Gadget-2 method, also written by Volker Springel.

mass, $m_i = m$. Note that this need not be the case in an SPH hydrodynamics calculation. Taking the same mass into account and including the chosen definition of the particle volume we can rewrite Eq. (3.24) as:

$$C = \frac{\frac{\sum_{i=1}^{N_{\rm part}} \rho_i^2 \frac{m}{\rho_i}}{\sum_{i=1}^{N_{\rm part}} \frac{m}{\rho_i}}}{\left(\frac{\sum_{i=1}^{N_{\rm part}} \rho_i \frac{m}{\rho_i}}{\sum_{i=1}^{N_{\rm part}} \frac{m}{\rho_i}}\right)^2} = \frac{1}{N_{\rm part}^2} \sum_{i=1}^{N_{\rm part}} \rho_i \sum_{i=1}^{N_{\rm part}} \frac{1}{\rho_i}. \tag{3.25}$$

Clearly, we can evaluate the clumping factor in any N-body volume by knowing only the SPH density of all particles in that volume. Note that, in general, the sum of SPH volumes of particles found in a volume is not equal to that volume: $\sum_{i=1}^{N_{\rm part}} V_i \neq V$. We can re-scale the particle volumes so that their sum equals the volume, $\sum_{i=1}^{N_{\rm cell}} AV_i = V$, where A is a constant scaling factor. If we use the re-scaled volumes in Eq. (3.24) all the scaling terms A will cancel out and the value of the clumping factor does not change. We conclude that the exact volume representation is not a factor in the clumping factor calculation - it is only the relative spatial distribution of matter that enters in the clumping. The same argument applies when calculating the clumping in any volume. We will use Eq. (3.25) to compute all the clumping factors in the following discussion.

Clumping factor in reionization simulations

In Section 3.4.1, we discussed what an accurate model of IGM density field for reionization simulations would entail. A simulation that would include all these components in a $\approx 100\,\mathrm{Mpc/h}$ side computational box is probably decades away[7]. This is why large-scale reionization simulations tend to either completely ignore the gas distribution beyond the resolution of their RT grids (Ciardi et al., 2000) or to include it through a clumping factor representative of the sub-cell matter distribution.

This sub-cell clumping factor is usually obtained from simulations with higher mass resolution and smaller box size (e.g. Kohler & Gnedin, 2005; Iliev et al., 2007; Pawlik et al., 2009). What is missing in this approach is the clumping factor's dependence on the local density - cells with different densities are assumed to have the same sub-cell

[7]For example, if an N-body simulation is to resolve the pre-reheating mean density Jeans mass at redshift 10 ($M_J \approx 10^4 M_\odot/h$) in a $L_{\rm box} = 100\,\mathrm{Mpc/h}$ box would require $\approx 10,000^3$ particles. Simulations have been run to date with $\approx 4000^3$ particles (e.g. Habib et al., 2009), about 10 times less than required. The hydrodynamical simulations that reach these mass and spatial scales are nowhere in sight.

gas distribution. McQuinn et al. (2007) tested the effects of an overdensity-dependent clumping factor on HII region morphology during reionization. Still, they used only a simple analytical estimate without any comparison to simulations.

We will argue that the clumping factor is not only a function of overdensity, but also of the spatial window in which it is computed. The latter effect follows from the existence of the smallest scales on which the matter is "clumped". In the N-body simulations we consider here, this limit is set by the mass resolution. While this is an artificially imposed scale, a physical one does exist in the IGM, the Jeans scale, where the pressure forces smooth out the density field (e.g. Bryan et al., 1999; Schaye et al., 2000). This means that the conclusions we reach here on the N-body data will apply for representing the IGM density fields in general.

Figure 3.9 shows the value of the clumping factor as a function of IGM threshold overdensity, $\Delta_{\rm thresh}$, computed from several N-body simulations (see Table 3.1 for simulation parameters). The overdensity is defined as $\Delta(\vec{r}) = \rho(\vec{r})/\bar{\rho}$, where $\bar{\rho}$ is the mean density of the Universe. The vertical dashed line marks our default IGM overdensity value, $\Delta_{\rm thresh} = 100$, above which we chose to exclude particles (see discussion why in Section 3.4.1). The chosen value is between the mean overdensity of a spherical top-hat halo (≈ 178 in the Einstein-de Sitter universe, Coles & Lucchin, 2002) and the overdensity at the virial radius of an isothermal halo (≈ 60, Lacey & Cole, 1994). We tested a range of values between 50 and 200 and found that our conclusions are unaffected by the exact choice.

While the clumping factor values in Fig. 3.9 are converged with respect to the box size (blue solid and red dashed lines), the same is not the case with respect to particle mass. This is not surprising, since the dark matter-only simulations lack the gas pressure forces that provide the limit to gas clumping, as we discussed previously. Indeed, if we compare Fig. 3.9 with the same tests performed on SPH hydrodynamical simulations by Pawlik et al. (2009) (their Fig. 4), we see that even though their particle mass is larger than ours, their runs do reach convergence for the same choice of $\Delta_{\rm thresh}$.

Considering the discussion so far, the N-body simulations we use are not an accurate representation of the IGM. This is why we ask the reader to keep in mind that the clumping factors we discuss in the following section are not meant to represent the correct density field beyond the scales resolved in N-body runs, as in other works. Our

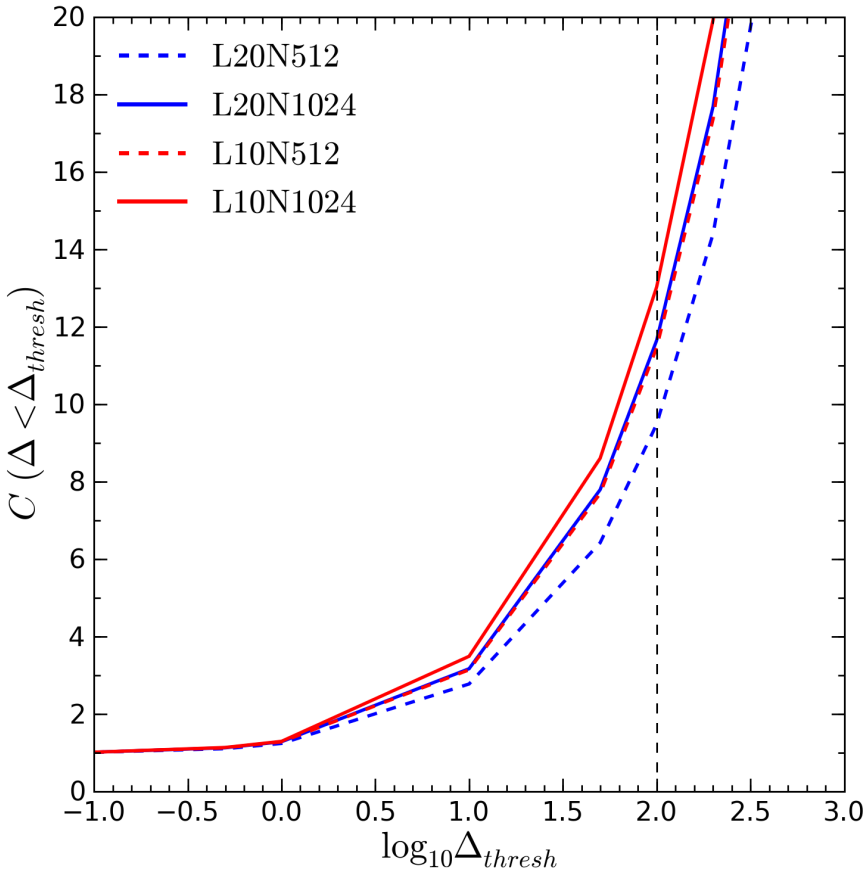

Figure 3.9: Global clumping factor as a function of IGM overdensity threshold. Lines represent the results of Eq. (3.25) using all particles from different N-body simulations. Red and blue lines represent runs in 10 and 20 Mpc/h boxes, respectively. The solid lines represent runs with 1024^3 particles, while the dashed lines are runs with 512^3 particles. See Table 3.1 for particle masses. The vertical dashed black line marks $\Delta_{\rm thresh} = 100$, the value we choose for our default overdensity threshold. The clumping factor results are converged with respect to the box size, but there is no convergence with respect to particle mass. We cannot claim that any of these simulations is an appropriate representation of the IGM.

clumping factor represents the resolved matter clumping, only what is given by the N-body simulation. If we could construct an "ideal" RT grid, with one Voronoi cell for each N-body particle, computing this clumping factor would be pointless. As this is not

the case, we will show that the clumping factor computed from the *available N-body data* can significantly improve the convergence of the RT results with respect to grid spatial resolution and construction. This idea is comparable to the method used by Trac & Cen (2007) who compute the ionizations on a coarse RT grid, but compute the recombinations on individual N-body particles. All other groups ignore this issue.

Global versus local clumping factors

An almost unanimous approach to modeling sub-grid density distribution for reionization simulations is to compute a single clumping factor (but evolving with redshift, e.g. Iliev et al., 2007) from a higher resolution run and assume that the same clumping factor describes the density distribution in any other volume in the Universe e.g. the computational cells of a larger simulation box. This assumption implies that the clumping factor over some volume V is the accurate representation of the density distribution in a *sub-volume* of the same volume, V_i, $V = \sum_i V_i$.

Figure 3.10 shows that this is in fact incorrect. We computed a single redshift-dependent clumping factor from all the particles with $\Delta < 100$ in the L20N1024 simulation, using Eq. (3.25). From this point on, we will name such clumping factors *global*, $C_{\rm global}$, to signify they were computed from all available density field data. So, $C_{\rm global} = \langle {n_{\rm HII}}^2 \rangle / \langle n_{\rm HII} \rangle^2$ where the brackets signify the mean over the whole simulation volume, introduced in Section 3.4.2. The obtained value of $C_{\rm global}$ for L20N1024 is shown with a solid line in Fig. 3.10. We compare $C_{\rm global}$ with the fit provided by Iliev et al. (2007) (dotted line): while their run is very different from the one used here (they computed C in a smaller box with higher mass resolution and assumed cosmology parameters from the WMAP 3-year best fit, Spergel et al., 2007), the clumping factor results are approximately in the same range.

As a next step, we subdivided the L20N1024 computational volume in 64^3 uniform square cells and computed a clumping factor in each, again using Eq. (3.25). We call these the *local* clumping factor, $C_{\rm local}$, to point out that they are computed using sub-volumes. The local clumping in a cell with volume V_i is $C_{\rm local,i} = \langle {n_{\rm HII}}^2 \rangle_i / \langle n_{\rm HII} \rangle_i^2$, where the $\langle \rangle_i$[8] indicates volume averaged over the cell i. The dashed line in Fig. 3.10 is the

[8]Note that while the global $\langle n_{\rm H} \rangle$ corresponds to the mean density in the Universe, $\langle n_{\rm H} \rangle_i$ does not, since it is an average over a region with $\Delta \neq 1$ in general.

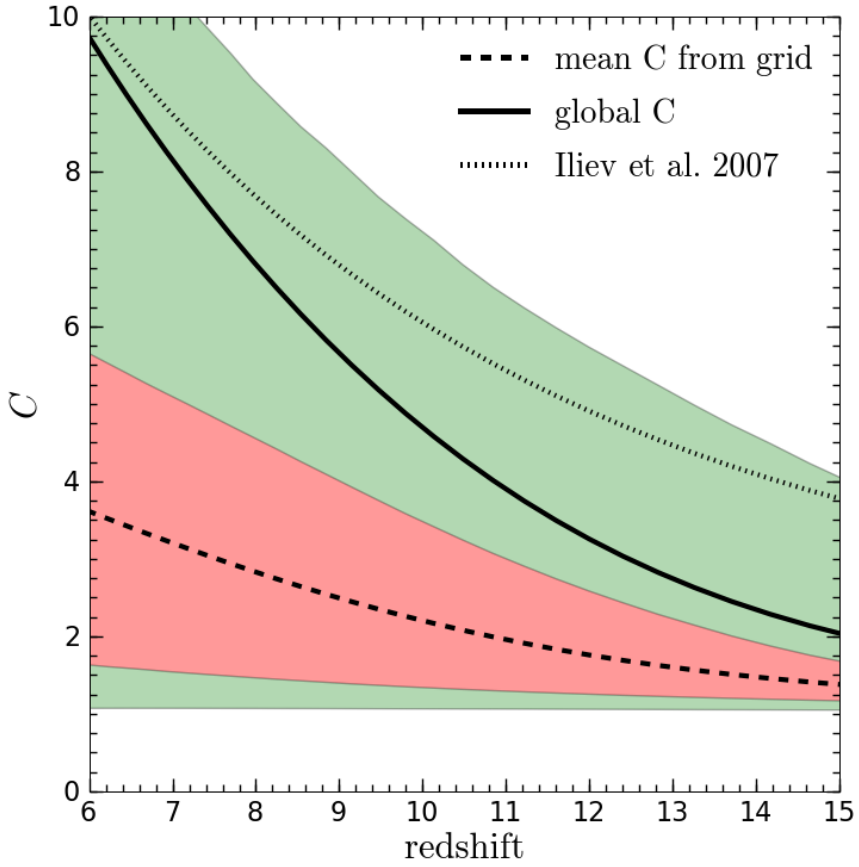

Figure 3.10: Evolution of the clumping factor with redshift in the L20N1024 simulation. The solid black line gives the clumping factor computed using all the particles in the box, $C_{\rm global}$. The red and green areas respectively show the 50% and 99% ranges of $C_{\rm local}$, in this case computed on 64^3 uniform cell subdivision of the original volume. The dashed line gives the mean $C_{\rm local}$ as a function of redshift. All the clumping factors are computed assuming $\Delta_{\rm thresh} = 100$. The global clumping factor is a poor representation of clumping on smaller scales, severely overestimating the actual values in most cells.

mean of the $C_{\rm local}$ distribution and the red and green shaded areas show the 50% and 99% range of values, respectively. It is clear that $C_{\rm global}$ is much higher than what one finds in most regions in the simulation box. Therefore, a single $C_{\rm global}$ is not only an inappropriate representation of matter on all scales, it will necessarily lead to overestimated

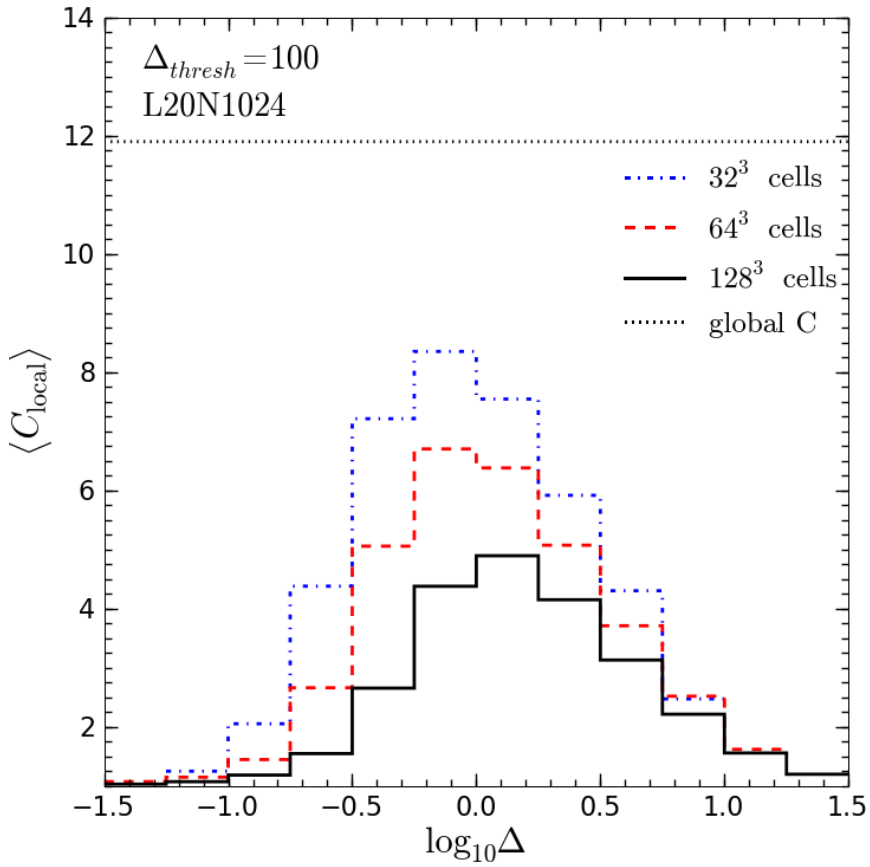

Figure 3.11: The mean local clumping factor as a function of overdensity in L20N1024 at z=5. The clumping factors are computed assuming $\Delta_{\rm thresh} = 100$. Histograms show the results on three subdivisions of the L20N1024 box, uniform Cartesian grids with 32^3 (blue dot-dashed), 64^3 (red dashed) and 128^3 (black solid line) cells. For reference, we show the global clumping factor (see text for definition) with the same threshold overdensity (black dotted line). See text for discussion.

recombination rates if used as the representation of matter distribution in the majority of the volume.

Figure 3.11 shows the volume-weighted mean $C_{\rm local}$ as a function of overdensity in the cell. First, $C_{\rm local}$ decreases with increasing grid resolution. This is not surprising since the N-body simulation mass resolution sets a limit on the matter clumping, as mentioned earlier. For the same reason, increasing RT resolution leads to more and more converged RT results. Second, it is clear that $C_{\rm global}$ (dotted horizontal line) overestimates the mean

clumping on all scales. This will necessarily lead to overestimated recombination rates as we will show shortly. Interestingly, the clumping is highest in moderately overdense regions. This is true irrespective of whether or not the halos have been removed from the density field. This kind of overdensity dependence of the clumping factor has not been previously considered in the literature, as the usual choice is a monotonically increasing or decreasing function (e.g. McQuinn et al., 2007). On higher resolution grids, the peak of the distribution shifts towards higher overdensities, simply because these grids resolve a larger overdensity range, but the shape of the distribution remains largely the same. More importantly, the same shape is obtained when using an adaptive Voronoi grid instead of the uniform Cartesian ones used for this discussion.

To understand the distribution of clumping factors, in Fig. 3.12 we show slices through the middle of the L20N1024 density field, with overplotted contours of regions with high clumping factor. In the left panel, all particles have been included in the clumping estimate ($\Delta_{\rm thresh} \rightarrow \infty$). The highest density regions correlate with high $C_{\rm local}$ regions. Once the halos are removed (by setting $\Delta_{\rm thresh} = 100$; right panel) the clumping decreases significantly everywhere, but the least so in the regions between voids and filaments, consistent with the peaks of $\langle C_{\rm local} \rangle$ in Fig. 3.11. This would imply that the large scale density gradients contribute significantly to the clumping factor, often overlooked in favour of the smaller scale density fluctuations. We defer a rigorous study of the contribution of different scales in the density field to total gas clumping to future work.

In the end, the clumping factor is just a tool for obtaining the correct number of recombinations. The total recombination rate in the L20N1024 simulation box, assumed to be fully ionized, $x = 1$, is:

$$\dot{N}_{\rm rec,total} = \alpha \int_V n_{\rm H}(\vec{r}) {\rm d}^3 r = \alpha C_{\rm global} \langle n_{\rm H} \rangle^2 V, \tag{3.26}$$

where we remind the reader of definition of the clumping factor as presented in Section 3.4.2.

The question is whether $\dot{N}_{\rm rec,total}$ can be obtained from the discrete values in subvolumes of V, the computational cells V_i, i.e. what is the error in the recombination rate introduced by representing the continuous density field with discrete volume cells. The

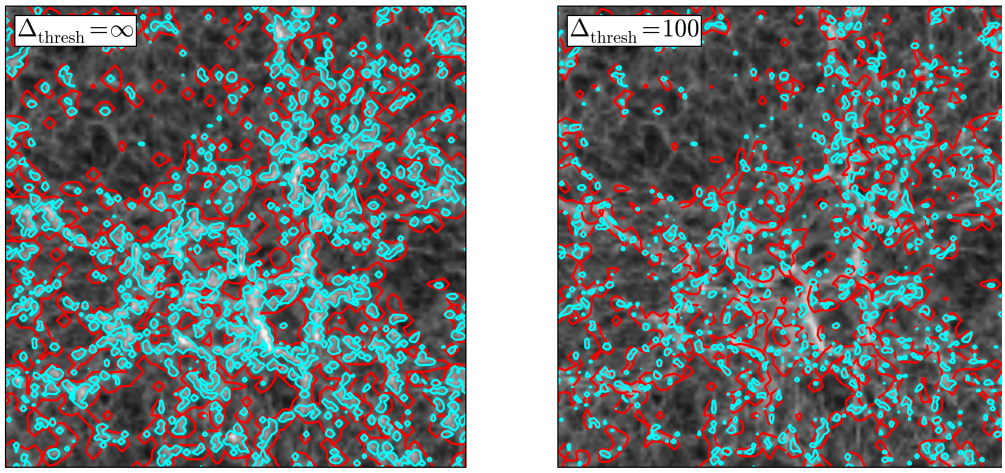

Figure 3.12: Slices through the middle of the L20N1024 density field (z=5) showing high clumping regions. All contour lines represent regions with clumping factor $C_{\rm local} \geq 5$, computed on a 64^3 and 128^3 uniform grids (red and cyan lines, respectively). The left-hand side plot shows the regions where the clumping is computed using all the N-body particles, while the right-hand side plot shows the regions where particles with $\Delta > \Delta_{\rm thresh} = 100$ were not included. Both plots demonstrate that high clumping regions trace high density regions, irrespective of the grid resolution. If $\Delta > \Delta_{\rm thresh}$ particles are excluded, the highest clumping regions are *not* in the highest density regions, pointing to the important contribution of the large scale gas distribution to the clumping factor.

total recombination rate computed as a sum of rates in each cell is:

$$\dot{N}_{\rm rec,sum} = \sum_{i}^{N_{\rm cell}} \alpha C_i \langle n_{\rm H} \rangle_i^2 V_i, \tag{3.27}$$

where C_i is the estimate of the clumping in cell i.

Figure 3.13 shows the ratio of Eqs. (3.27) and (3.26) as a function of grid resolution and clumping factor employed. The bar colors correspond to various choices of C_i: yellow bars assume there is no matter clumping in the cell, $C_i = 1$, blue bars use the locally computed clumping factor in each cell, $C_i = C_{\rm local,i}$, and the red bars assume the same globally computed clumping factor in all cells, $C_i = C_{\rm global}$. The dashed line gives the

target value, where $\dot{N}_{\rm rec,sum} = \dot{N}_{\rm rec,total}$. Ignoring subgrid matter clumping (yellow bars) underestimates the recombination rate. This is entirely due to the smoothing of the density field, as evidenced by the increase of the recombination rate with increasing grid resolution. Using the local clumping factor (blue bars) largely solves this issue. While there is still some dependence on the grid resolution, it is much smaller than when C_i is ignored, e.g. while the 64^3 grid with $C_i = 1$ produces 3 times less recombinations per unit time, the same resolution grid with $C_i = C_{\rm local,i}$ produces only about 20% more. The 256^3 grid with $C_{\rm local}$ gets to within a few percent of the desired result. Finally, the most important insight is that the use of a single clumping factor $C_{\rm global}$ (red bars) drastically overestimates the recombination rate. Furthermore, it does nothing to boost convergence between results from different grids as it only linearly amplifies the differences that exist when $C_i = 1$ (orange bars). The locally computed C does the opposite and works towards converging different grid results to the "correct" solution. In the next section, we will use the locally computed clumping to fix the issues stemming from grid sampling presented in Section 3.3.

3.4.3 Repeating the N-body test runs with the local clumping factor

We have demonstrated how the use of locally computed clumping factors can produce converged recombination rates computed on density data grids. The volume over which the clumping is computed can have any shape, so there is nothing preventing us from computing it on Voronoi cells. We identify the N-body particles belonging to each Voronoi cell using the same octree-based neighbour search procedure used to assign masses to cells, mentioned in Section 3.3 and use Eq. (3.25) to compute local clumping factors, then repeat the test runs presented in Section 3.3.2.

First we repeat the runs presented in the left panel of Fig. 3.7, where we showed the effect of grid cell sampling of the evolution of the ionized fraction in the L12.5N128 box. The results are shown in Fig. 3.14. The inclusion of the local clumping factor in the calculation (lines marked with circles) slows down the I-fronts and brings the evolution significantly closer to the desired result (blue line). The same effect is observed in Fig. 3.15, where we repeat the runs with different numbers of grid cells originally presented in the right panel of Fig. 3.7.

While using local clumping factors produces a noticeable improvement (bringing the

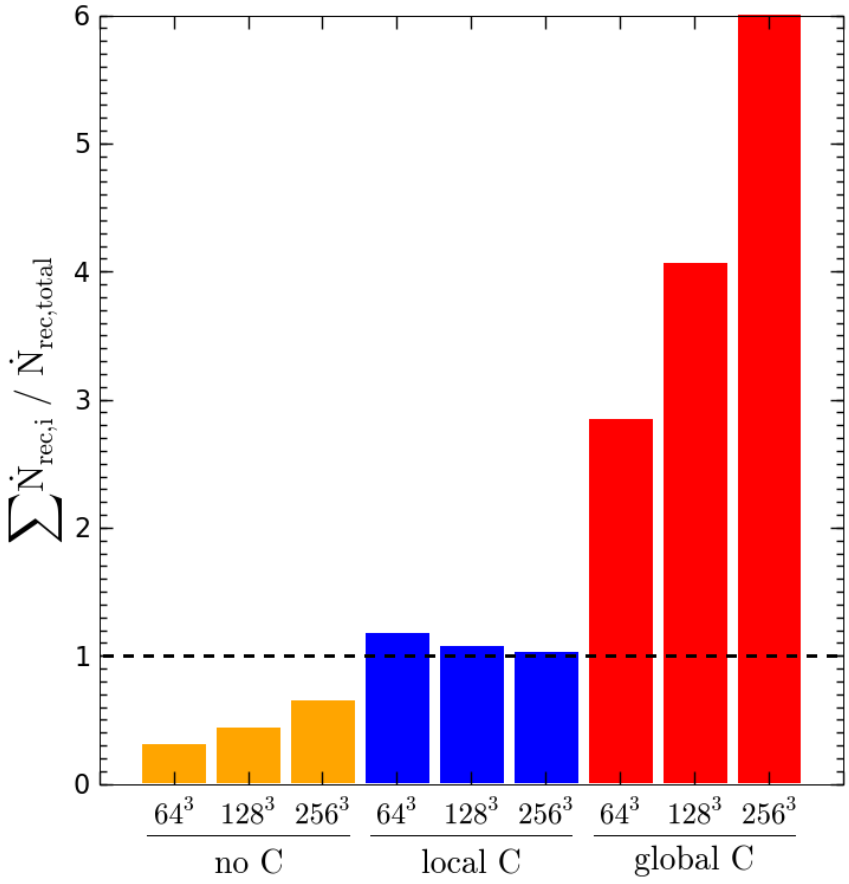

Figure 3.13: Total number of recombinations computed in the L20N1024 simulation with respect to grid resolution and clumping factors employed. The bars represent ratios of recombination rate integrated over the grid cells representing the density field and the total value defined by Eq. (3.26). The use of $C_{\rm local}$ converges the results from different resolution grids towards the desired value, while the use of a single $C_{\rm global}$ only amplifies the issues due to grid resolution. See text for further discussion.

low resolution I-front positions about 50% closer to the desired result), the local clumping factor used in the L12.5N128 runs does not produce perfect convergence with the desired result (blue line in both figures). This is probably due to the small number of particles in this N-body run. In all the runs presented in Fig. 3.14 and 3.15 a Voronoi cell has, on average, only 8 particles. This results in relatively low values of the clumping factors, decreasing their effect on the I-front propagation.

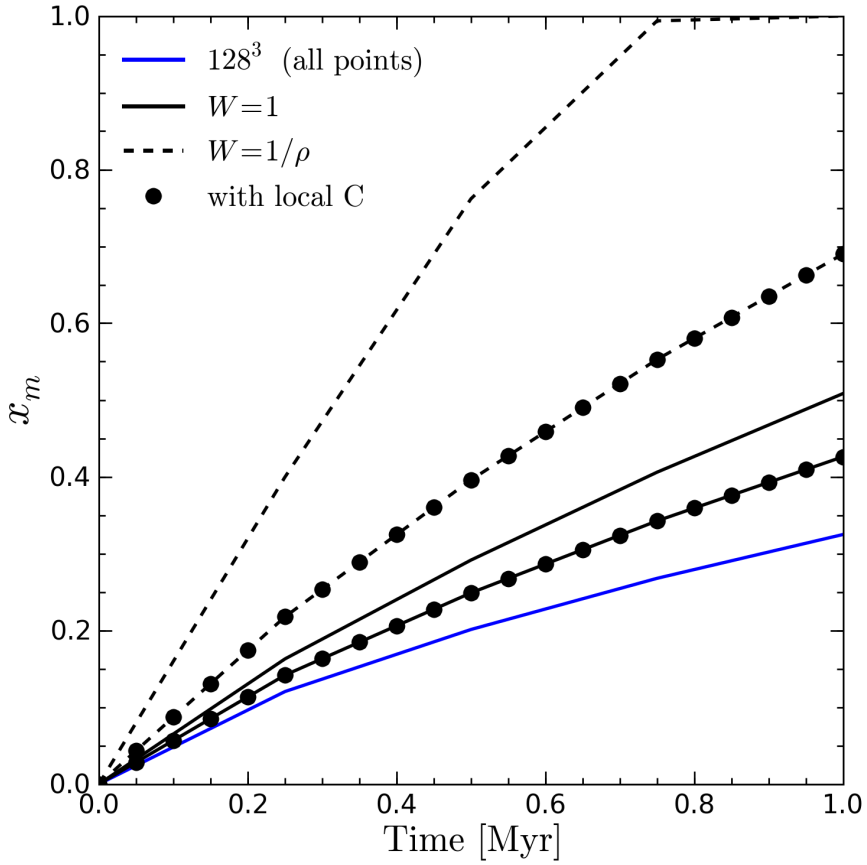

Figure 3.14: The evolution of the mass-weighted ionized fraction in the L12.5N128 test runs with different sampling functions, with and without the local clumping factor. Runs using a local clumping factor are marked with circles. While different sampling functions make a large difference for the 64^3 resolution run (black lines, solid for $W = 1$, dashed for $W = 1/\rho$), inclusion of the local clumping factor converges the results towards the run with all N-body points used as Voronoi nuclei (128^3 resolution, blue solid line).

This is not the regime we are primarily interested in. In the reionization runs we perform in Chapter 5, the RT cells represent many more N-body particles. To test the effect of the local clumping factors in such runs we do single snapshot runs on the L20N512 box. This simulation box is used in Chapter 5 to study the numerical effects in full reionization runs. We again chose the final output snapshot at z=5 as the underlying density field. The 37 most massive halos identified in the box at that redshift are used as radia-

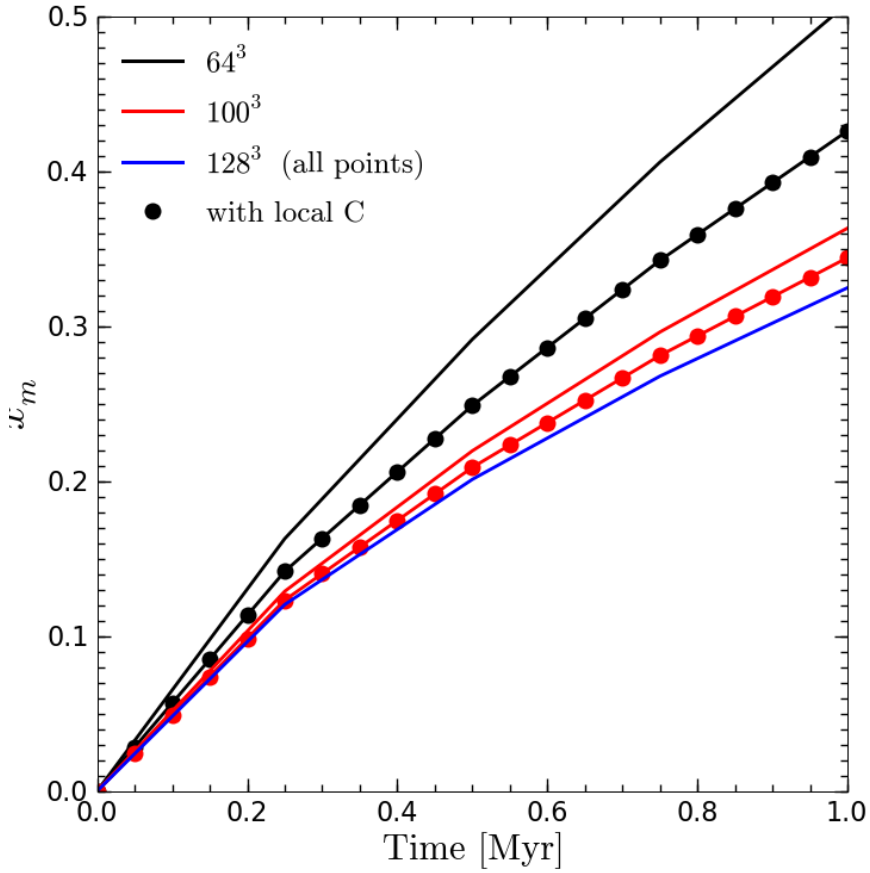

Figure 3.15: Same as Fig. 3.14 , but now with different RT grid resolutions (all grids sampled assuming $W = 1$). The local clumping factor brings lower resolution run results (black lines for 64^3 and red lines for 100^3) towards the highest resolution one (128^3, blue solid line). Once again, the runs that include the local clumping factor are marked with circles.

tion sources, again assigning them luminosities with Eq. (3.15). The efficiency parameter was $f_\gamma = 1000$. It is lower than in the previous runs because we chose both a larger number and more massive halos as sources than in the L12.5N128 box. The system is again evolved for 1 Myr and the ratio of ionizing photons produces to the total number of atoms in the box is ≈ 1.5. We present only results from grids created with the $W = 1$ weighting function: since we are interested in the detailed representation of the density field, using a grid that explicitly adapts to it is a logical choice.

Figure 3.16 shows the effect of the local clumping factor on runs in the L20N512

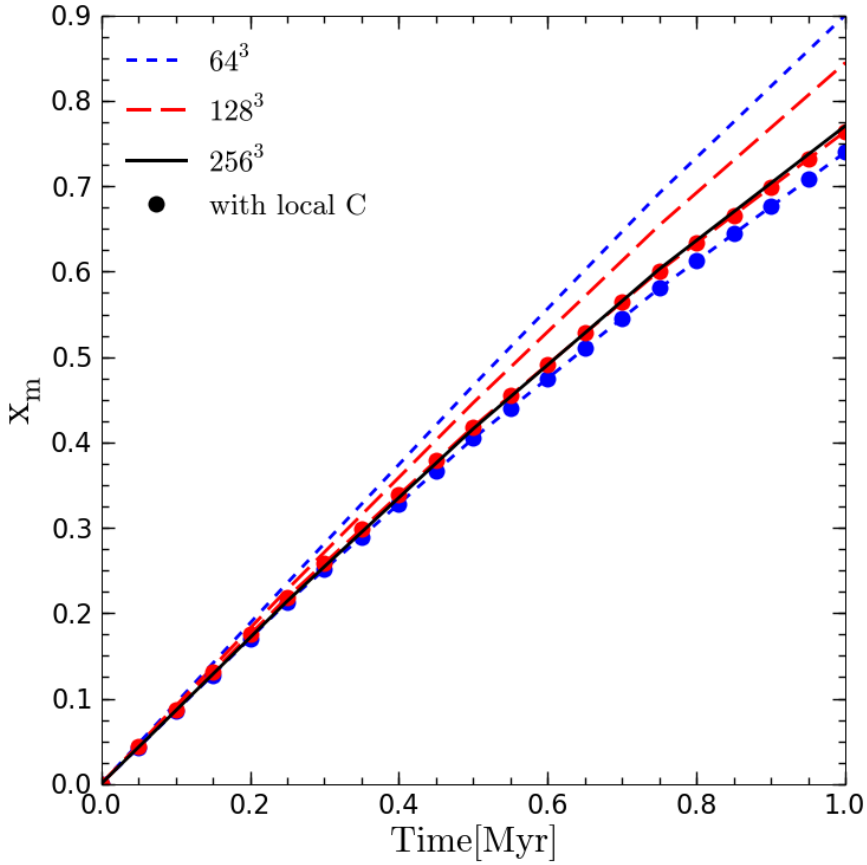

Figure 3.16: The effects of the local clumping factor on the evolution of the mean mass-weighted ionized fraction in the L20N512 box. The runs that use the local clumping factor are marked with circles. All runs use the $W = 1$ sampling function with different numbers of cells: 256^3 (black solid line), 128^3 (red long dashed lines) and 64^3 (blue short dashed lines). We do not show the result of the 256^3 run with C, because it gives identical results to the runs without C, meaning the grid effectively resolves all the available density field data. As in Figure 3.15, the inclusion of the local clumping factor helps convergence of the different resolution results. See text for further discussion.

box. On these grids there are many more N-body particles per Voronoi cell: on average 512 and 64 for 64^3 and 128^3 grid, respectively. The effect of the inclusion of the local clumping factor is as expected from previous results. Without the local clumping factors, the I-fronts in the 64^3 resolution run are about 20% faster at final time than in the 256^3

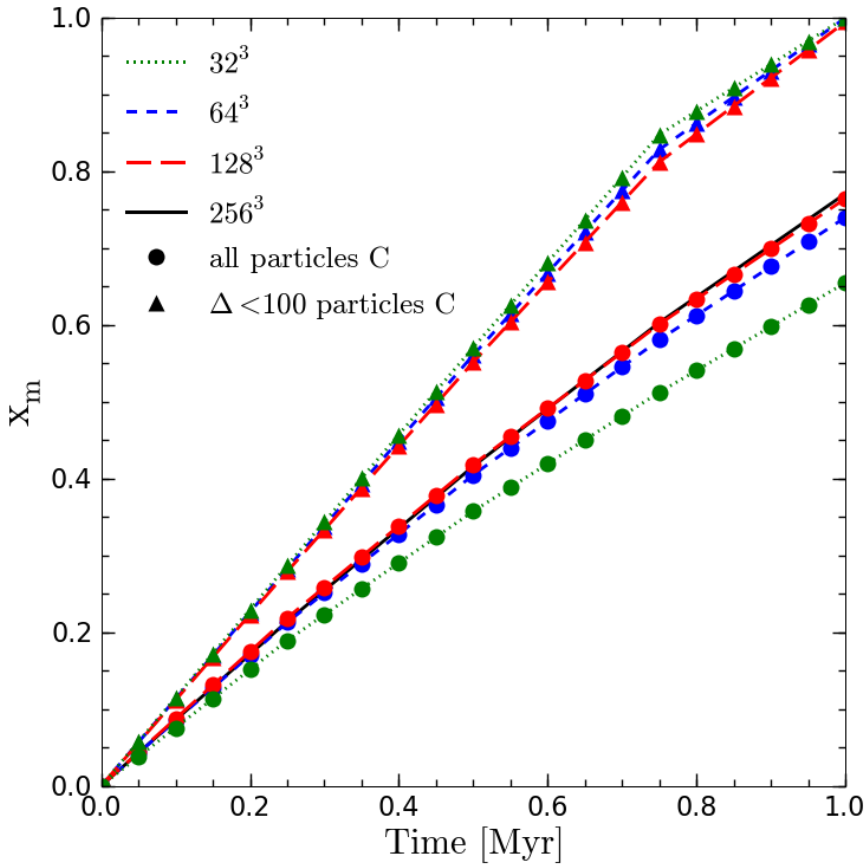

Figure 3.17: The effects of excluding particles with overdensities higher than 100 on the evolution of the ionized fraction in the L20N512 test run. Line styles are the same as in Figure 3.16, with the addition of a 32^3 run (green dotted line). All runs used the local clumping factor. The triangle symbol marks the runs that do not include high overdensity particles in both the density field and the local clumping factor. The local clumping factor applied on the density field without high overdensity particles gives a higher level of convergence of the global ionization state.

run. With the clumping factors, the difference between runs is much smaller, with 64^3 run being just a few percent slower than the other two runs, that agree almost perfectly.

The test runs presented in Figs. 3.14 and 3.15 have not excluded the halo particles (i.e. $\Delta_{\rm thresh} = \infty$). Removing them will decrease the very high recombination rate otherwise found in source cells. Figure 3.17 shows the evolution of the mean ionized fraction

with and without the halo particles in the density field. When the halo particles are excluded, the convergence of RT results improves even further (lines marked with triangle symbols). Note that a further decrease in the RT grid resolution in the case when all particles are included starts breaking the convergence in evolution of $x_{\rm m}$ (green dotted line with circle symbols). We must remember that the clumping factor by itself is not a replacement for a properly resolved RT grid, but only a tool for representing the available density data that is not resolved in practise. To get the best results, one should still use the highest RT grid resolution possible[9] The results presented in Figures 3.16 and 3.17 show that the local clumping factor helps in obtaining converged RT results in typical reionization run conditions. We show this explicitly in Section 5.3.1.

3.5 Cosmological Strömgren sphere test

In this section we present the results of the cosmological Strömgren sphere test, performed here with SimpleX for the first time. The test entails the evolution of a spherically symmetric HII region in a uniform cosmologically expanding medium. This is a test relevant for all cosmological radiative transfer applications, such as reionization simulations.

In Section 2.4, we introduced the basic equations for the I-front expansion in physical (non co-moving) coordinates. To obtain the evolution of the radius of a spherically symmetric I-front, R_I, we substitute the flux at front interface, Eq (2.24) into the I-front jump condition, Eq. (2.23):

$$n_{\rm H}\frac{{\rm d}R_I}{{\rm d}t} = \frac{1}{4\pi R_I^2}\left(\dot{N}_\gamma - \frac{4\pi}{3}\alpha_B n_H^2 R_I^3\right), \tag{3.28}$$

where all the variables have been introduced in 2.4. Equating the terms in the bracket, defines the Strömgren radius R_S, Eq. (2.27), where the I-front speed becomes zero.

The HII region expansion given by Eq. (3.28) is modified for a density field in the expanding Universe. Let us rewrite Eq. (3.28) in co-moving coordinates:

$$n_{H,c}\frac{{\rm d}R_{I,c}}{{\rm d}t} = \frac{1}{4\pi R_{I,c}^2}\left(\dot{N}_\gamma - \frac{4\pi}{3}\alpha_B n_{H,c}^2 R_{I,c}^3 a^{-3}\right). \tag{3.29}$$

[9]At the moment, the highest number of RT cells used in a large-scale reionization simulation is 512^3 (McQuinn et al., 2009), while the typical number is $\approx 256^3$ (e.g. Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007). All these works use different N-body box sizes and mass resolutions, so a comparison of the detail with which the density field is resolved is not straightforward.

Here, we have included the scale factor a, introduced in Chapter 1, which relates the physical and co-moving coordinates: $R_I = aR_{I,c}$ for the HII region radius and $n_{\rm H} = n_{\rm H,c}a^{-3}$ for the hydrogen number density. Equation (3.29) is simplified by the fact that the peculiar velocity of the I-front dominates over the cosmic expansion speed, $a\frac{{\rm d}R_{I,c}}{{\rm d}t} \gg R_{I,c}\frac{{\rm d}a}{{\rm d}t}$, which is true for all typical sources (Shapiro & Giroux, 1987). Comparing Eqs. (3.28) and (3.29) we see that the one difference is the a^{-3} term in the recombination rate in Eq. (3.29), i.e. that the number of recombinations decreases with decreasing redshift. This suggests that it is possible for an HII region never to reach its Strömgren radius.

We compare the SimpleX results against the work of Shapiro & Giroux (1987), who have shown that Eq. (3.29) has an analytic solution in the matter-dominated Einstein-de Sitter universe:

$$y(t) = \lambda e^{\lambda\frac{t_i}{t}}\left[\frac{t}{t_i}{\rm E}_2(\lambda\frac{t_i}{t}) - {\rm E}_2(\lambda)\right], \tag{3.30}$$

where $y \equiv [R_{I,c}(t)/R_{S,c}]^3$, i.e. the co-moving position of the front in the units of the co-moving Strömgren radius, $R_{S,c} = \left[3\dot{N}_\gamma/4\pi\alpha_B n_{H,c}^2\right]^{1/3}$, $\lambda = t_i/t_{\rm rec,i} = t_i\alpha_B n_{H,c}$ is the ratio of the recombination time and the age of the Universe at the epoch of source turn-on, and $E_2(x) \equiv \int_1^\infty \frac{e^{-xt}}{t^2}{\rm d}t$ is the second order exponential integral. The scale factor is normalised to the epoch of source turn-on, $a_i = 1$.

The setup for the SimpleX run was as follows. The source was always turned on at redshift $z_{\rm on} = 20$, but we varied the density of the medium to correspond to various overdensities of the mean hydrogen density in the Universe. Since reionization in most of our runs starts at about $z \approx 20$, we were more interested to see how SimpleX models the transport through regions of different overdensities at that epoch. The effect of increasing the density is equivalent to turning the source on at an earlier redshifts, since the density enters the solution only through the λ term. We assumed a matter-dominated Universe, $\Omega_{\rm m} = 1$, with baryonic component of $\Omega_b = 0.045$, Hubble parameter $h = 0.73$ and finished the runs at $z = 0$. The source luminosity was set to be $\dot{N}_\gamma = 10^{52}\,{\rm photons\,s^{-1}h^{-2}}$. Simulation box has the side length of 10 Mpc/h, chosen so that the HII region never leaves the box for the density used. We use a 64^3 Voronoi grid with randomly positioned cell nuclei and impose a constant density in the cell[10]. During the run, we modify the Delaunay line lengths and cell densities to correspond

[10]Note that this is different from our grid creation procedure described in Section 3.3, where the grid is created on equal mass N-body particles.

to $a(t)$ at every time step (sweep). We also tried evolving the density in various coarser steps, equally spaced in $\log_{10}a$ as in our reionization runs (see Chapter 5), and found only negligible differences from the fine step density updates. We compute the position of the I-front by assuming all the ionized atoms are located in a sphere centred on the source, $R_I = \sqrt[3]{3V_I/4\pi}$, where $V_I = (N_{\rm HII}/N_{\rm H})V = xV$ is the volume of the ionized region, V is the total volume of the computation box while $N_{\rm H}$ and $N_{\rm HII}$ are respectively the total number of hydrogen atoms and the number of hydrogen ions in the box. This was done because the unstructured Voronoi grid is unable to produce perfectly spherical HII regions, especially at lower resolutions. Using the actual number of ionized atoms to compute the front position allows for a comparison that is independent of the actual grid.

Figure 3.18 compares the results from the SimpleX runs at two overdensities, $\Delta = 1$ and $\Delta = 10$, with the analytic results of Shapiro & Giroux (1987). The behaviour hinted at by Eq. (3.29) is clearly observed here. As the density decreases due to the expansion of the Universe, the recombination rate reduces, leading to an increase of the Strömgren radius. This increase of R_S is large enough to out-race the propagation of the I-front at our chosen source turn-on redshift[11] and $R_{I,c} < R_{S,c}$ at all times. Therefore, the cosmological I-fronts are expected to remain very fast, R-type (see Section 2.4), throughout the lifetime of the source. This result provides the rationale for ignoring the effects of the radiation driven gas flow in large-scale reionization simulations, which is the assumption we adopt in our work (Chapter 5).

The original results we obtained with the SimpleX method are presented with the green X symbols. Even at moderate overdensities, $\Delta = 10$, the result does not agree with the analytic solution. Even worse, the SimpleX method seems to produce unphysical results: I-fronts that reach beyond their Strömgren radius (points above the y-axis value of one). The root of this problem is, once again, an inadequate RT computational grid. As the density is uniform, the issue is not the representation of the density field we discussed in Section 3.4. Instead, it is purely a result of inadequate spatial resolution.

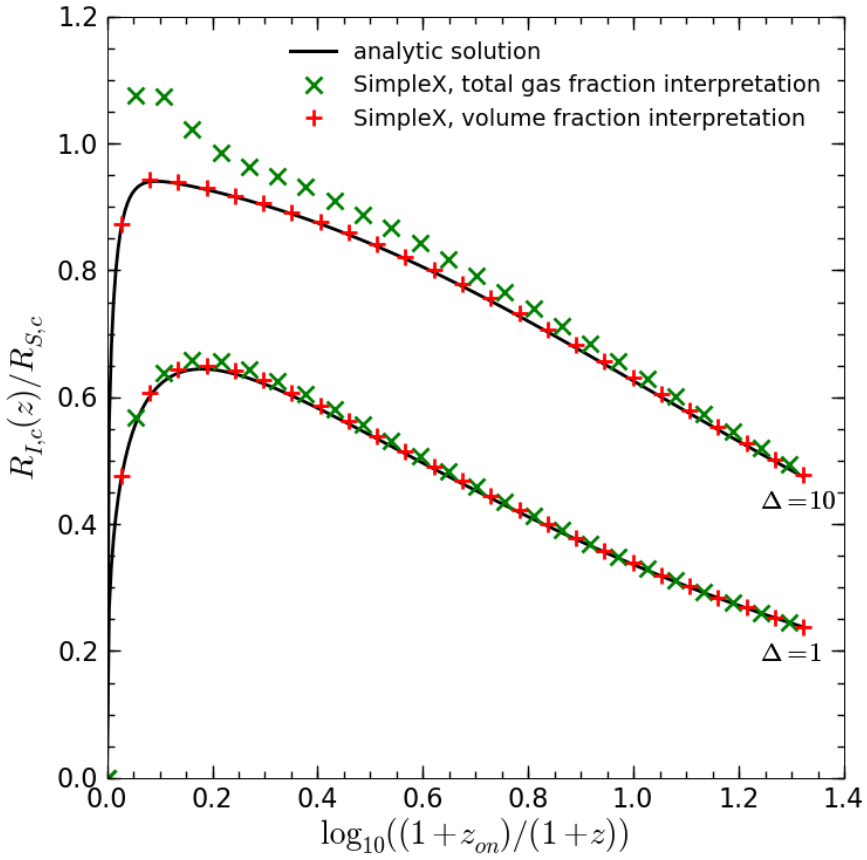

Figure 3.18: The evolution of the co-moving I-front position in an expanding density field. The I-front positions are given in units of the co-moving Strömgren radius, $R_{S,c}$ computed at source turn-on redshift, $z_{\rm on} = 20$. Bottom set of curves represents runs at the mean density, $\Delta = 1$, while the top ones represent moderately overdense regions, $\Delta = 10$. The black solid lines give the analytic solution (see text), while the two sets of symbols represent SimpleX runs: green X symbols give the result of the original SimpleX implementation while the plus symbols represent the runs with the new, volume interpretation of the ionized fraction. See text for further discussion.

Volume interpretation of the ionized fraction

As we pointed out in Section 3.4.2, the values in computational cells are averages of continuous field values over the cell volumes. The recombination rate in a cell can be

[11]For higher turn-on redshifts (or overdensities), the I-front can reach the Strömgren sphere in the early evolution. See Shapiro & Giroux (1987) for more details.

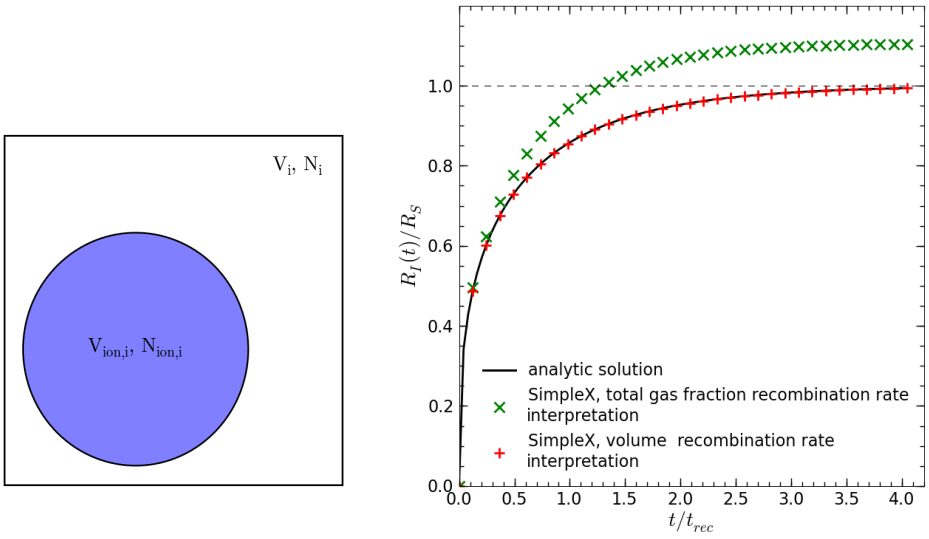

Figure 3.19: Diagram showing the actual distribution of ionized gas in the case when a whole HII region is contained within one cell. In that case, we know that all of the ions, $N_{\rm ion,i}$ are located in a sphere around the source, $V_{\rm ion,i}$. The total volume and number of atoms in the cell are V_i and N_i, respectively.

written as:

$$\dot{N}_{\rm rec} = \alpha \langle x \rangle^2 {n_{\rm H}}^2 V. \tag{3.31}$$

This is the same as Eq. (3.17), but since the density field is uniform in each cell, the recombination rate depends only on the ionized fraction x. The usual assumption (and the one used by the original SimpleX implementation) is that the ionized fraction is constant throughout the cell. This is accurate for optically thin cells, but becomes increasingly incorrect for optically thick cells. To understand why, first recall the relation between the optical depth of a cell and the mean free path: $\tau = l/\lambda$, where l is a spatial dimension of the cell (e.g. the Delaunay line in the SimpleX scheme) and $\lambda = (\sigma n_{\rm HI})^{-1}$ is the mean free path. Secondly, as we discussed in Section 2.4, the I-front thickness is of order of a few mean free paths, e.g. for an I-front driven by Lyman limit photons only ($h\nu_{\rm th} = 13.59\,{\rm eV}$) has the thickness of $\approx 18\lambda$ (Ritzerveld, 2007, see also Test 1 results in Fig 2.6). This means that an I-front crossing the grid will be fully contained in a cell with $\tau \gtrsim 18$. For even higher optical depths, $\tau \gtrsim 100$, the I-front thickness will only correspond to a small fraction of the cell's spatial dimension, so one can assume that it represents an infinitely

sharp transition between neutral and ionized regions, see Eq. (2.23).

The average value $\langle x \rangle$ used in Eq. (3.31) cannot distinguish between these two extreme cases. For example, $\langle x \rangle = 0.5$, either when the whole cell has $x = 0.5$ (uniformly ionized cell) or when $0.5V$ of the cell is fully ionized ($x = 1$) and the rest is neutral, which occurs when a sharp front crosses half of an optically thick cell. However, the recombination rate should not be the same in those two cases. The left panel of Fig. 3.19 shows an example of the latter case: a very optically thick cell holding a radiation source around which a spherical HII region is expanding. In that case, the ionized fraction should be interpreted as the *fraction of the cell volume V that is completely ionized*, $V_{\rm ion} = \langle x \rangle V$. As only ionized gas can recombine, the recombination rate is:

$$\dot{N}_{\rm rec}^{\rm vol} = \alpha {n_{\rm H}}^2 V_{\rm ion} = \alpha \langle x \rangle {n_{\rm H}}^2 V, \tag{3.32}$$

where we assume that $x = 1$ in $V_{\rm ion}$.

Comparing Eqs. (3.31) and (3.32), we see that the recombination rate in a cell depends on how we *interpret* the mean ionized fraction in that cell. The right panel of Fig. 3.19 shows the consequences of two interpretations we are discussing for a highly optically thick cell presented in the left panel. When $\langle x \rangle$ is interpreted as the fraction of the volume to be ionized and Eq. (3.32) used to compute the recombination rate (red pluses), the exact analytic solution, given in Eq. (2.25) (black solid line), is reproduced. This is no surprise since using the volume interpretation of $\langle x \rangle$ in a cell is equivalent to solving the I-front jump condition, Eq. (3.28). On the other hand, assuming that $\langle x \rangle$ represent a uniformly ionized cell, with $x(\vec{r}) = \langle x \rangle$ throughout the cell as is assumed in Eq. (3.31), the obtained I-front speed (or the number of neutral atoms ionized per unit time) is incorrect and leads the front to surpass its Strömgren radius (green X symbols). Comparison of Eqs. (3.31) and (3.32) shows that the difference in the recombination rate evaluated with two interpretations is exactly $\langle x \rangle$ i.e. the assumption of uniformly distributed x leads to $\langle x \rangle$ times fewer recombinations produced per unit time in very optically thick cells.

This is exactly the issue that leads to the wrong results in Fig 3.18 (green X symbols): due to limited spatial resolution and very high density in cells, the optical thickness is very high ($\tau \gtrsim 100$) meaning that x is to be interpreted as a fraction of volume that is fully ionized. The original SimpleX rate equation solver, Eq. (3.13), does not discriminate between the scenarios we described here. Instead, it always uses Eq. (3.31) to evaluate the recombination rate which is an underestimate in the conditions of this test. A simple

switch to using Eq. (3.32) produces excellent agreement between SimpleX results and the analytic solution of Eq. (3.30) (red pluses).

Simply using Eq. (3.32) instead of Eq. (3.31) in all possible problems is also incorrect as it will overestimate the recombination rate in optically thin cells. We therefore use each when it is appropriate. We use the volume interpretation if the optical thickness in the cell is $\tau > 10$, i.e. if the I-front is resolved by less than 2 cells. Since the cells always transition to the optically thin regime when they are fully ionized (see Fig. 2.12), the classic interpretation of the ionized fraction will always be used. The volume interpretation is therefore necessary only for highly optically thick cell in the process of being ionized. This is still a very important contribution, since highly optically thick cells remain optically thick for almost all values of x, as demonstrated by the $\Delta\tau = 100$ result in the left panel of Fig. 2.12.

Finally, this issue is not specific to SimpleX. Indeed, any method that uses Eq. (3.31) exclusively will underestimate the recombination rate in the very optically thick cells. The solution we discussed in this section is therefore complementary to the photon conservation idea of Abel et al. (1999): where their fix guarantees the correct number of *ionizations* in the optically thick cell, ours does the same for *recombinations*.

3.6 Parallelization

Dividing up an RT problem in order to solve it in parallel is different from hydrodynamics or gravity problems. The "area of influence" of radiation is much larger than that of gas flow, due to the vastly different characteristic speeds. This non-locality of influence is similar to that of gravity, but with an important distinction: for gravity one density element act on another at some distance without any interference from all other elements. This is not true for radiation, as the energy arriving from one element to the other depends on all the matter between the two. The subdivision of an RT problem for optimal parallel performance is complicated by the fact that the number and spatial distribution of active computational cells changes in unpredictable ways. As an example, consider the evolution of an HII region around an isotropically radiating source in an uniform density field. We can chose to divide the volume so that the number of cells affected by radiation energy is equal on each processor used, by splitting the volume around the ra-

diation source. The number of cells affected by radiation increases with the expansion of the HII region, but when the subdivision is symmetric this increase will be equal on all processors. Any departure from these "ideal" conditions will break the load balancing e.g. if the density field is not uniform the I-front will expand faster in some directions than others. Indeed, one needs to know the solution of the problem to plan a proper division in advance, which in a sense negates the whole point of modeling it.

The original SimpleX could not be run in parallel. In fact, this was used as an example of the method's efficiency: interesting RT problems can be modelled on a single desktop machine. While this is certainly impressive, we wanted to get the highest possible speed out of the scheme, which necessarily meant using a multi-processor machine. To avoid the issues with domain decomposition discussed in the previous paragraph, we chose to use shared memory parallelism paradigm and used the OpenMP programming directives. In a shared memory implementation, the whole computational domain is visible to all processors. The loops to be executed by multiple threads are marked by simple compiler directives. Since the cell indices are not correlated with the spatial distribution of cells, simply executing equal-sized chunks of a loop over cells on each thread provides, on average, good load balancing. This is not strictly true when the mean ionized fraction of the whole volume is low, but in that case the total operation count is small anyway.

In order to use OpenMP to parellelize the scheme, we had to make a small change to the way the photon packets are moved on the Delaunay grid. We already discussed the sweeps in Section 3.2.3. In the original, serial implementation, photon packets are moved at every cell visitation. In some compositions of the grid, this can result in photon packets that are moved more than once per sweep. For example, say that a photon packet is moved from cell with index i to cell with index j. If $j > i$ the photon packet will be moved again once the cell j is visited later on in the sweep. This scheme can lead to undefined behaviour in a multi-threaded implementation (i.e. two threads simultaneously accessing the same memory register). Furthermore, the final result depends on the order in which the grid is swept. To avoid this problem, we split photons transported along a Delaunay line into the incoming and outgoing photons. When a photon packet is moved from cell i to cell j, it is written in the outgoing section of cell i. Once all the photons have been moved, the outgoing section of i is copied to incoming section of

j, which is then propagated further in the next sweep. This implementation results in photon packets that move only once per sweep and no simultaneous memory access by multiple threads.

The creation of the Delaunay triangulation was also parallelized with OpenMP. For that purpose, the whole domain is divided into equal volume cubes and each of them is triangulated individually. In order for the triangulation to be correct (i.e. identical to the one created serially over the whole volume), every cube needs to have a boundary region populated with grid points from neighbouring cubes. If the density field is very irregular, this implementation will not provide sufficient load balancing. Fortunately, for the high redshifts and the box sizes we employ for simulation reionization, this is not the case. Furthermore, the triangulation part of the computation takes a smaller fraction of the time compared to actual sweeps, thus not requiring further optimisation.

Finally, while not used for any runs presented in this thesis, we have also developed a version of the code for distributed memory machines. This version is "multi-threaded" to take advantage of machines that have multiple cores on a single node. Such a node can act as an independent shared memory machine. We distribute the computational domain in planes nodes. Then, on every node we use the original OpenMP implementation to evolve the radiation field. With this we reduce the number of boundary regions that need to be used to set up the triangulation and to communicate between distributed regions, thus reducing MPI communication cost. Before we proceed with the RT calculation, we test whether the triangulation was properly created (by comparing the neighbours of cells in boundaries with the same cells in the main region on another node) and make a list of all Delaunay lines that connect regions on different nodes. The fact that we distribute larger domains means that the number of particles is better load balanced i.e. distributed volumes have on average the same density. This implementation will be tested in more detail and used in our future work.

3.7 Conclusions

This chapter presented the SimpleX radiative transfer scheme. The main strengths of the scheme include an operation count that is independent from the number of radiation sources (much more efficient when many sources are present as compared to

characteristics-based scheme, such as HC, Chapter 2), an adaptive grid representation of the density field (an important factor in resolving the range of densities in a cosmological density field) and the replacement of the ray-tracing procedure with the construction of the grid itself (Delaunay lines used as paths for photon packets). All these features make SimpleX an attractive method for simulation the reionization process and we justified our choice by comparing its strengths and weaknesses with other codes used for the same purpose in Section 3.2.4.

All of the tests and improvements we introduced in this chapter were motivated by our use of SimpleX in large-scale reionization simulations. First, in Section 3.3.1 we showed that the perceived inaccuracy of the SimpleX method in the Test 4 of Iliev et al. (2006b) code comparison is completely due to the point sampling scheme used in the original implementation. We argued that the original sampling strategy, based on Eq. (3.7) was motivated by scaling arguments unattainable in practise and showed that our preferred sampling procedure, based on matching the scales present in the density field, results in much better agreement with other codes in the Test 4 problem. We elaborated on the point sampling process with further tests in Section 3.3.2. The runs we performed on single snapshots of N-body data, mirroring a step of a procedure we will use in Chapter 5 to model reionization, showed that the RT results may depend on the point selection and grid resolution. However, the cause of these differences is not native to the SimpleX scheme, but a consequence of representing the detailed density field on coarser discrete computational grids.

We focused on this issue in Section 3.4. There, we argued that the root of differences seen in Section 3.3.2 is the "smoothing" of the N-body density field by the computational grid used to represent it. We introduced the clumping factor as a numerical tool to combat this issue. We showed that using a clumping factor evaluated locally in each computational cell can indeed result in much better convergence of RT results, a claim demonstrated by repeating tests from Section 3.3.2 in Section 3.4.3. Equally important, we showed that the common use of a single clumping factor as a representation of the sub-cell density field, as is commonly done in most other reionization modeling work, is inadequate and may lead to severe overestimates of the recombination rate, as demonstrated in Fig. 3.13. Indeed, taking into account the definition of the clumping factor we introduced at the beginning of the same section, a single value of C can be considered

an accurate representation of the density field only in the volume in which it was computed. We conclude that an accurate representation of gas clumping in a computational cell must take into account both the cell's local density and its spatial size, as shown in Fig. 3.11.

In Section 3.5 we tested the SimpleX scheme on the problem of an HII region propagating through a density field that expands with the Universe, for which the analytic solution was provided by Shapiro & Giroux (1987). The test results we obtained pointed out the issue of correctly computing the recombination rate in very optically thick cells. There, we introduced a simple fix to the recombination rate evaluation, based on reinterpretation of the ionized fraction in the cell as a fraction of the cell volume that is fully ionized, and showed that it completely fixes the issue. We pointed out that the recombination rate of optically thick cells is the other side of the coin presented in the work of Abel et al. (1999), who only fix the ionization rate with what has been called spatial photon conservation in Chapter 2.

Some of the work presented here, namely the clumping factor as a tool for RT convergence, will be elaborated upon further in Chapter 5. For now, we conclude that the improved SimpleX scheme is a fast and accurate method for simulating RT during the reionization epoch. =

Chapter 4

GALFORM: modelling the galaxies that reionize the Universe

4.1 Introduction

On the list of potential sources of photons that reionize the Universe, the most plausible dominant sources are the galaxies (see the discussion in Chapter 1). However, the observations of galaxies at redshifts of EoR are still scarce. We must therefore turn to theory to obtain estimates of the production of ionizing photons by galaxies during reionization. For that purpose, we employ the Durham semi-analytic galaxy formation and evolution model, GALFORM (Cole et al., 2000). The current theory of galaxy formation rely heavily on the ΛCDM cosmological paradigm and the hierarchical buildup of structure discussed in Chapter 1. The key elements of the theory were put forward by Fred Hoyle some fifty years ago. Hoyle (1951) suggested that the source of galactic rotation are the tidal torques that affect the collapsing gas, later expended upon by others (e.g. Peebles, 1969; Doroshkevich, 1970; White, 1984; Catelan & Theuns, 1996). Hoyle (1953) also introduced the cooling time argument to explain observed galaxy masses. This idea was elaborated on by Rees & Ostriker (1977) and Silk (1977). White & Rees (1978) combined the cooling arguments with the theory of Press & Schechter (1974) laying the foundation of the modern galaxy formation theory.

The semi-analytic models of galaxy formation, one of which is GALFORM, aim to treat the relevant physics in a simplified manner which allows the use of analytic or easy numerical solutions. The first semi-analytic galaxy formation model was introduced by White & Frenk (1991), who expanded on the work of White & Rees (1978) by includ-

ing many features present in today's models: gas cooling, star formation, feedback effects and stellar population models (see also Cole, 1991; Lacey & Silk, 1991; Kauffmann et al., 1993a; Cole et al., 1994). The alternative to this approach are the numerical simulations that explicitly solve the relevant gravitational and hydrodynamical equations using some of the numerical techniques developed for this purpose (e.g. Crain et al., 2009; Schaye et al., 2010, to name recent examples). The semi-analytic approach makes several important simplifications, e.g. the assumption of spherical symmetry. Simulation methods cannot include the physics that occurs below their resolution limit and treat such unresolved physical processes phenomenologically, in a way similar to semi-analytic methods. The two approaches are complementary and have been shown to agree fairly well (e.g. Pearce et al., 1999; Benson et al., 2001).

The GALFORM models have been carefully calibrated using a wide range of available observational data that sketches an empirical picture of galaxy formation and evolution. Major breakthroughs in observations of high redshift galaxies have been made in the final years of the 20th century. For example, the Hubble Space Telescope (HST) and its Deep Field project (Williams et al., 1996; Ferguson et al., 2000) provided crucial information for understanding early galaxy formation. The Deep Field data played a part in the first determination of the cosmic star formation history and accompanying metal production history from $z \approx 5$ to the present day (Madau et al., 1996; Ellis, 1997; Madau et al., 1998). On the other hand, the development of the Lyman-break dropout selection technique (Steidel et al., 1996a) provided both ground based telescopes and the HST with an efficient method for identifying high-z star-forming galaxies ($z \gtrsim 3$). Further technological advances lead to the discovery of sub-millimetre galaxies, a population of star-forming galaxies at $z \geq 2$ that were too heavily obscured by dust to be easily found in the optical surveys (Smail et al., 1997; Ivison et al., 1998; Blain et al., 2002, for a review). Yet another route for finding young star-forming galaxies at high redshifts is by detecting their Lyman alpha emission (e.g. Hu et al., 2004; Taniguchi et al., 2005; Kashikawa et al., 2006; Ouchi et al., 2008). Our knowledge of the galaxies in the Local Universe has also been vastly improved by several large galaxy survey programs such as the two-degree Field Galaxy Redshift Survey (2dFGRS, Colles et al., 2001) and Sloan Digital Sky Survey (SDSS, York et al., 2000) leading to among other things the precise determination of the galaxy luminosity functions in various bands (e.g. Kochanek et al.,

2001; Norberg et al., 2002; Nakamura et al., 2003; Pozzetti et al., 2003). GALFORM models have been shown to be able to reproduce a large fraction of this huge pool of data, giving it credence as a model of the galaxy formation process.

Most importantly for the subject of this thesis, in the last year the observations have begun to reach the redshifts relevant to the study of reionization: for example, the Lyman-break technique has been used to observe galaxies as high as $z \sim 10$ (Bunker et al., 2004; Bouwens et al., 2007, 2009a,b; Oesch et al., 2009). This new data allows us to compare (and potentially adjust) GALFORM models to the galaxies near the end of reionization as predicted by the CMB data, potentially providing a first direct glimpse to the sources causing the EoR. Finding a significant agreement with this data would further strengthen the case for using GALFORM galaxies as sources of ionizing photons during reionization.

We will use the GALFORM code to model galaxies during EoR, and in particular to investigate how reionization progresses if GALFORM galaxies are used as sole sources of ionizing photons. Before proceeding to couple GALFORM sources with our chosen RT method (SimpleX, discussed in Chapter 3), we must examine whether it is appropriate for this purpose. GALFORM was already used to study some aspects of the reionization of the Universe, such as the effect of photoionization on galaxy formation (Benson et al., 2001, 2002a,b) or the global history of reionization (Benson et al., 2006). These previous models included some of the main physical processes, such as a detailed treatment of gas cooling in the presence of ionizing radiation, but modelled the radiative transfer effects in a very simplified manner. This is why they mostly focus on simply counting the number of ionizing photons without going into more details about their origin or propagation. The discussion of ionizing emissivity we present in this chapter is an expansion of that previous work, focusing mostly on which galaxies dominate the ionizing emissivity, an important component in full RT simulations in that it may affect the spatial distribution of HII regions (McQuinn et al., 2007).

The organisation of the Chapter is as follows. In Section 4.2, we briefly present the basic ingredients that make up the GALFORM scheme. This discussion is based mostly on the paper by Cole et al. (2000), which originally presented GALFORM. The scheme was later updated and used by a number of authors (e.g. Benson et al., 2003; Baugh et al., 2005; Bower et al., 2006; Lacey et al., 2008). In Section 4.3, we discuss the most important

model parameters in terms of the production of ionizing photons. Section 4.4 presents a comparison of predictions of GALFORM models (varying some of the same parameters as in Section 4.3) with the recently observed Lyman-break galaxy data at $z \sim 6 - 10$. Finally, in Section 4.5, we discuss the ionizing emissivities predicted by GALFORM runs performed on N-body data with limited mass and spatial resolution, as will be the case in our full reionization runs in Chapter 5.

4.2 The GALFORM semi-analytic method

In this section we briefly describe the key components of the GALFORM semi-analytic model of galaxy formation. The discussion borrows heavily from Cole et al. (2000), to which we refer the reader interested in a more detailed presentation of the scheme. We present the generation of dark matter halo merger trees (Section 4.2.1), the assumed structure of said halos (Section 4.2.2), gas cooling (Section 4.2.3), star formation (Section 4.2.4), stellar population modelling (Section 4.2.5), galaxy mergers (Section 4.2.6), photo-ionization feedback parameters (Section 4.2.7) and the default Durham GALFORM parameters (Section 4.2.8).

4.2.1 Dark matter halo merger trees

In the current paradigm, the galaxies form at the centres of dark matter halos. The dark matter provides the gravitational scaffolding for the forming galaxy. The evolution of galaxies will depend on the merger histories of the halos containing them, so it is imperative to have an accurate description of how every halo evolves through mass accretion and hierarchical merging.

The halo merger histories (also called merger trees, due to their specific visual representation) for use in the GALFORM model can be obtained in two ways: generated with a Monte-Carlo algorithm or directly extracted from N-body simulations.

Monte-Carlo merger tree generation algorithm

The dark matter halo merger trees can be generated without any numerical calculations of structure formation using a technique based on the extension of the Press-Schechter formalism (Press & Schechter, 1974) as proposed by Bond et al. (1991) and Bower (1991).

The algorithm is described in Cole et al. (2000) and here we present the basic idea behind it.

Lacey & Cole (1993) derived the equation for the average fraction of mass, $f_{12}(M_1, M_2)\mathrm{d}M_1$, found in a halo of mass M_2 at time t_2 that was, at an earlier time t_1, found in halos with masses between M_1 and $M_1 + \mathrm{d}M_1$:

$$f_{12}(M_1, M_2)\mathrm{d}M_1 = \frac{1}{\sqrt{2\pi}} \frac{\delta_{c1} - \delta_{c2}}{(\sigma_1^2 - \sigma_2^2)^{3/2}} \times \exp\left[-\frac{(\delta_{c1} - \delta_{c2})^2}{2(\sigma_1^2 - \sigma_2^2)}\right] \frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\mathrm{d}M_1, \qquad (4.1)$$

where σ_1 and σ_2 are the linear theory root mean squares (rms, abbreviated) of density fluctuations in spheres of mass M_1 and M_2, while δ_{c1} and δ_{c2} are the critical linear overdensity thresholds for collapse at times t_1 and t_2. In an Einstein-de Sitter Universe ($\Omega_\mathrm{m} = 1$), the assumed time dependence of the critical density is $\delta_c = 1.686(1 + z)$, while for low-Ω_0, open and flat universes the expressions for δ_c are given by Lacey & Cole (1993) and Eke, Cole & Frenk (1996), respectively.

In the limit of $t_1 \to t_2$, Eq. (4.1) leads to an expression for the average mass fraction of a halo of mass M_2 which was in halos of mass M_1 at a slightly earlier time:

$$\left.\frac{\mathrm{d}f_{12}}{\mathrm{d}t_1}\right|_{t_1=t_2} \mathrm{d}M_1\mathrm{d}t_1 = \frac{1}{\sqrt{2\pi}} \frac{1}{(\sigma_1^2 - \sigma_2^2)^{3/2}} \frac{\mathrm{d}\delta_{c1}}{\mathrm{d}t_1} \frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\mathrm{d}M_1\mathrm{d}t_1. \qquad (4.2)$$

This leads to the number of halos of mass M_1 at the earlier time $\mathrm{d}t_1$ that will become a part of a halo of mass M_2:

$$\frac{\mathrm{d}N}{\mathrm{d}M_1} = \frac{\mathrm{d}f_{12}}{\mathrm{d}t_1} \frac{M_2}{M_1}\mathrm{d}t_1, \qquad (4.3)$$

valid when $M_1 < M_2$. Introducing the parameter, M_res, which is the minimum halo mass to be considered in the merging hierarchy, Eq. (4.3) can be used to compute:

$$P = \int_{M_\mathrm{res}}^{M_2/2} \frac{\mathrm{d}N}{\mathrm{d}M_1}\mathrm{d}M_1, \qquad (4.4)$$

which is the mean number of progenitor halos with masses in the range $M_\mathrm{res} < M_1 < M_2/2$. For mass conservation, it is important to know the fraction of halo mass M_2 that comes from "unresolved" halos, $M_1 < M_\mathrm{res}$. Again using Eq. (4.3), this fraction is:

$$F = \int_0^{M_\mathrm{res}} \frac{\mathrm{d}N}{\mathrm{d}M_1} \frac{M_1}{M_2}\mathrm{d}M_1. \qquad (4.5)$$

Equations (4.4) and (4.5) are the main tools for constructing the merger trees. Given the required mass of the final halo, its merger tree is built with the following procedure. First, note that both Eqs. (4.4) and (4.5) depend on the choice of $\mathrm{d}t_1$, through the use

of Eq. (4.3). A sufficiently small time step $\mathrm{d}t_1$ is used to have $P \ll 1$, meaning that it is unlikely that the halo had multiple mergers during $\mathrm{d}t_1$. Next, a random number R is generated in the interval between 0 and 1. If $R > P$, the halo is assumed to have only one progenitor in the time step i.e. the halo did not go through any mergers. Its mass still increased in that time by accretion, so the mass of the progenitor is set to be $M_p = M_2(1 - F)$, where the fraction F comes from Eq. (4.5). On the other hand, if $R < P$, the halo is assumed to have grown through both a single merger and accretion. The mass of the merging halo, M_1 is drawn randomly from the distribution defined by Eq. (4.3), while the mass of the accreting halo is $M_p = M_2(1 - F) - M_1$. This procedure is recursively performed on all generated halos until the mass resolution limit, $M_{\rm res}$ is reached.

This procedure results in a binary-type tree with very fine time resolution. Before the semi-analytic galaxy formation prescriptions are applied, the binary tree is converted to a new one with a coarser time resolution, where the number of time steps is also an input parameter, $N_{\rm steps}$.

The merger trees are also divided into "branches" that define the lifetime of each halo in the tree. Each "leaf" progenitor (the halo that has no progenitors itself) is followed until it becomes a part of a halo with $f_{\rm form}$ times its original mass. This procedure leads to the definition for the halo formation times: a new halo is "born" at the point where the mergers produce a halo which is $f_{\rm form}$ times more massive than its most massive progenitor. Cole et al. (2000) set $f_{\rm form} = 2$ to agree with the earlier tree generation technique (Cole et al., 1994), but the final results of the galaxy formation model do not depend strongly on the actual choice. In this work, we employ a modified version of the algorithm designed to produce a closer fit to merger trees in the Millennium simulation (Springel et al., 2005b), was described by Parkinson et al. (2008).

Generation of merger trees from N-body simulation outputs

Alternatively, the dark matter halo merger trees employed by GALFORM can be constructed using outputs of N-body simulations using a technique developed by Helly et al. (2003). As a first step, the halo catalogue must be extracted from every N-body simulation snapshot. The method of Helly et al. (2003) uses the friends-of-friends (hereafter FOF) algorithm (Davis et al., 1985) for this purpose, which connects all the particles

with separations smaller than some chosen value, b, usually expressed in the units of mean inter-particle separation. Once the halos are identified in each simulation output, they are connected in a merger tree. For every pair of snapshots, at times t_1 and t_2 where $t_1 < t_2$, halos at time t_1 are marked as progenitors of whichever halo contains the largest fraction of its particles at t_2. Repeating this procedure for all snapshots leads to fully fleshed out merger trees of halos at the final time. Special care is taken to correct the FOF "bridging" issue, where the algorithm artificially connects two separate halos into a single structure.

The mass resolution limit, which is a free parameter in the previously described MC tree construction technique($M_{\rm res}$), is dictated by the minimum number of particles, $N_{\rm min}$, that must be connected by the FOF algorithm for that group to be considered a halo, with the exact choice motivated by the need for the group to be numerically well defined. Our choice throughout this thesis is $N_{min} = 20$. The cosmological parameters and $N_{\rm steps}$, input parameters in the MC tree-generating technique, are now also given by the N-body simulation.

The generation of N-body merger trees that we use in this thesis also employs the SUBFIND code (Springel et al., 2001), which identifies bound dark matter structures within every FOF halo by searching for overdense regions using the local SPH density estimate. The substructures are defined as regions bounded by an isodensity surface that traverses a saddle point of the density field and are gravitationally bound systems. The halos in the merger tree are still identified by the FOF finder, but the SUBFIND output is used to better separate artificially merged halos and to provide positions for individual galaxies within the halo. These updates to the scheme have been previously discussed in Harker et al. (2006).

The main difference between the N-body and MC merger trees is that, in the former case, the halo mass can decrease with time. The number of particles in a halo can decrease both due to tidal striping and the use of the FOF group finder. In terms of the semi-analytic model, this results in negative mass accretion rates leading to unphysical results. To avoid this, Helly et al. (2003) chose to enforce the mass conservation by increasing the mass of affected halos to ensure that the accretion rate is always positive. Regardless of this difference, it was found that the galaxy formation model results do not strongly depend on the exact way the merger trees are obtained. In Section 4.5, we will

compare the ionizing photon emissivities obtained from both merger tree construction techniques.

4.2.2 Structure of dark matter halos

The merger history is not the only information about dark matter halos needed by the galaxy formation model. First, as we already mentioned in the introduction, the galaxy spin is thought to be a consequence of the angular momentum of dark matter halos, itself caused during the halo collapse by tidal torques resulting from matter inhomogeneity (Fall & Efstathiou, 1980; Mo et al., 1998). The amount of angular momentum gained by the halo in this fashion is conventionally described by a dimensionless spin parameter:

$$\lambda_{\rm H} = \frac{J_{\rm H}|E_{\rm H}|^{1/2}}{GM_{\rm H}^{5/2}}, \tag{4.6}$$

where $J_{\rm H}$ is the total angular momentum, $E_{\rm H}$ the total energy (potential and kinetic) and $M_{\rm H}$ the total mass of the halo. The distribution of halo spin parameters has been extensively studied using N-body simulations (e.g. Efstathiou & Jones, 1979; Barnes & Efstathiou, 1987; Frenk et al., 1988; Catelan & Theuns, 1996; Cole & Lacey, 1996; Bullock et al., 2001; Bett et al., 2007). These studies generally agree in the shape of the distribution, finding that it only weakly depends on the halo mass and initial density fluctuations spectrum. By default, GALFORM uses the log-normal probability density function of $\lambda_{\rm H}$ given by Cole & Lacey (1996) to assign a spin parameter to each newly created halo (the criterion for halo creation was discussed in Section 4.2.1). Alternatively, the fit from the Millennium simulation obtained by Bett et al. (2007) can also be used.

The key assumption made GALFORM and many other semi-analytic models (e.g. Lacey & Cole, 1993; Kauffmann et al., 1994; Somerville & Kolatt, 1999; Hatton et al., 2003) is that the gas has initially the same angular momentum as the dark matter halo, and that angular momentum is conserved as the gas collapses. This assumption leads to a good agreement with observed disc sizes (e.g. Mo et al., 1998; de Jong & Lacey, 2000; Cole et al., 2000), though the typical scalelengths depend on other parameters as well (e.g. Cole et al. (2000) demonstrate the role of supernova feedback on the disc sizes). Note that numerical simulations struggle to reproduce the observed disc sizes, due to the angular momentum transfer between baryons and dark matter, possible due to numerical or limitations of the simulations (e.g. Okamoto et al., 2003; Governato et al., 2004).

As we will discuss in the following section, not all of the gas found in the halo will be capable of cooling and taking part in the formation of the galaxy. To obtain the angular momentum of the fraction of gas that does cool, the distribution of matter within the halo must be known. GALFORM models the halo dark matter density using the NFW profile (Navarro, Frenk & White, 1996, 1997):

$$\rho(r) = \frac{\Delta_{\rm vir}\rho_{\rm crit}}{f(a_{\rm NFW})} \frac{1}{r/r_{\rm vir}(r/r_{\rm vir} + a_{\rm NFW})^2} \qquad (r \leq r_{\rm vir}), \tag{4.7}$$

where $f(a_{\rm NFW}) = \ln(1 + 1/a_{\rm NFW}) - 1/(1 + a_{\rm NFW})$. The virial radius of the halo, $r_{\rm vir}$ is the radius at which the mean density in the halo is $\Delta_{\rm vir}$ times the critical density, $\rho_{\rm crit} = 3H^2/(8\pi G)$. The value of the virial overdensity $\Delta_{\rm vir}$ in GALFORM is given by the spherical top-hat collapse model (e.g. Coles & Lucchin, 2002, mentioned in Chapter 1) and is $\Delta_{\rm vir} = 178$ for an Einstein-de Sitter Universe ($\Omega_0 = 1$)[1]. Expressions used for evaluating $\Delta_{\rm vir}$ in low-Ω_0, open and flat universes were supplied by Lacey & Cole (1993) and Eke et al. (1996). The NFW profile has one free parameter, $a_{\rm NFW}$, which gives the scalelength measured in units of the virial radius which can be varied to provide the density profile fits for a wide range of halos found in N-body simulations. This free parameter also correlates with the halo mass and the analytic fit of this relation provided by Navarro et al. (1997) is used in GALFORM to set the halo density profiles.

GALFORM assumes that the rotational velocity is constant in concentric shells of material in the halo (remember that the density profile in Eq. (4.7) depends on r only) and always aligned in the same direction, which is a picture broadly consistent with simulations (e.g. Warren et al., 1992; Cole & Lacey, 1996). This, together with Eqs. (4.7) and (4.6), leads to the mean rotational velocity of the shells:

$$V_{\rm rot} = A(a_{\rm NFW})\lambda_{\rm H} V_{\rm H}, \tag{4.8}$$

where $A(a_{\rm NFW})$ is a dimensionless coefficient weakly dependent on $a_{\rm NFW}$ and $V_{\rm H} \equiv (GM/r_{\rm vir})^{1/2}$ is the circular velocity of the halo at $r_{\rm vir}$. The derivation of Eq. (4.8) is given in Appendix A of Cole et al. (2000).

The assumption of conservation of the angular momentum and Eq. (4.8) allow us to estimate the amount of angular momentum added to the galaxy by the cooling gas.

[1]We already mentioned this value in Chapter 3, when choosing the threshold overdensity to exclude halos from the density fields for reionization simulations

Note, however, that the gas distribution is not exactly equal to that of dark matter, given by Eq. (4.7). We return to this issue in the next section.

4.2.3 Gas cooling

The cooling of gas in dark matter halos is the foundation of current galaxy formation theory. As we discussed in Chapter 3, the assumption of gas as a perfect tracer of dark matter holds on large, linear scales. When an overdense region decouples from the Hubble flow, it begins the collapse eventually resulting in the formation of a virialized halo. The gas is thought to be heated by shocks associated with the virialization process to about the virial temperature of the halo (e.g. Peacock, 1999):

$$T_{\rm vir} = \frac{1}{2}\frac{\mu m_{\rm u}}{k} V_{\rm H}^2, \tag{4.9}$$

where μ is the mean molecular weight of the gas and m_u is the atomic mass unit. The gas is also thought to reach hydrostatic equilibrium. We will call this gas "hot" to distinguish it from the "cold" gas that is found in galaxies. A first order approximation for the density distribution of the hot gas would be to assume that it traces the dark matter, therefore that is distributed according to the NFW profile, Eq. (4.7). GALFORM takes a different distribution, one motivated by the results of high-resolution hydrodynamic simulations of galaxy cluster formation (e.g. Navarro et al., 1995; Eke et al., 1998; Frenk et al., 1999) which find the shock-heated gas to be less centrally distributed than the dark matter. The assumed distribution in the GALFORM is:

$$\rho_{\rm gas}(r) \propto 1/(r^2 + r_{\rm core}^2), \tag{4.10}$$

where $r_{\rm core}$ is the β-model parameter used to fit the results of the simulations (Cavaliere & Fusco-Femiano, 1976). Since the hot gas is less centrally concentrated, as seen in Eq. (4.10), assuming that it has the same mean rotational velocity as the dark matter, given by Eq. (4.8), will result in it having a slightly higher angular momentum than the dark matter. To avoid this, the gas is assumed to have a somewhat different (but still constant with radius) rotational velocity, $V_{\rm rot}^{\rm gas}$, defined so that the gas and dark matter have the same angular momentum within the virial radius.

With the known density distribution of gas and its temperature, assumed to be constant throughout the halo and given by Eq. (4.9), we can define the cooling time on

each radius as the ratio of thermal energy density to the cooling rate per unit volume, $\rho_{\rm gas}\Lambda(T_{\rm gas}, Z_{\rm gas})$:

$$t_{\rm cool}(r) = \frac{3}{2}\mu m_u \frac{kT_{\rm gas}}{\rho_{\rm gas}(r)\Lambda(T_{\rm gas}, Z_{\rm gas})}. \quad (4.11)$$

Here, $\rho_{\rm gas}$ is given by Eq. (4.10), $T_{\rm gas} = T_{\rm vir}$ is the gas temperature and $Z_{\rm gas}$ is the metallicity. The cooling function, $\Lambda(T_{\rm gas}, Z_{\rm gas})$, is provided by Sutherland & Dopita (1993). Using Eq. (4.11) we can introduce the cooling radius at time t since the formation of the halo, $r_{\rm cool}$, as the radius at which $t_{\rm cool} = t$. The cooling radius gives the amount of gas that can cool within the time t, hence losing pressure support leading to collapse. The amount of gas that actually collapses does not depend solely on the cooling time scale, but also on how much of it is able to reach the centre of the halo if it collapses under the influence of gravity alone. This time scale is estimated by the free-fall time (e.g. Lang, 1999, introduced in Chapter 1), which allows us to define the free-fall radius, $r_{\rm ff}(t)$, which limits the amount of gas that is able to fall onto the halo centre.

Finally, to compute the amount of gas that is able to collapse and be accreted onto the galactic disc within one GALFORM time step, Δt, the radius $r_{\rm min} = \min[r_{\rm cool}, r_{\rm ff}]$ is computed at times t and $t + \Delta t$ marking the shell of material that cools and collapses within Δt. If $M_{\rm shell}$ is the mass within this shell then:

$$\dot{M}_{\rm cool} = \frac{M_{\rm shell}}{\Delta t}, \quad (4.12)$$

is the cooling rate, the rate with which the galactic halo is supplied with fresh cold gas. This rate enters the system of differential equations that describe star formation, chemical enrichment and various feedback effects.

4.2.4 Star formation

The lack of a complete theory of star-formation may be the main missing element in understanding galaxy formation and evolution. The processes grouped under the star formation umbrella convert the cold gas to luminous stars, which in turn affects the state of the gas through various feedback processes. A simple treatment of all these processes has been shown to be sufficient for reproducing many observations of how the galaxies evolve. In the previous section, we showed how to obtain the amount of gas that cools and becomes available for star formation. Given the amount of cold gas available in the galaxy, $M_{\rm cold}$ (which is increased by accretion of cooling gas at the rate $\dot{M}_{\rm cool}$ introduced

in the previous section), GALFORM calculates the instantaneous star formation rate as:

$$\psi = M_{\rm cold}/\tau_*, \tag{4.13}$$

where τ_* is the star formation timescale. The timescale can be defined in different ways, some of which will be discussed in the following sections, but is in general a function of the properties of the galaxy and its halo. GALFORM assumes that some fraction R of the total mass taking part in the star formation is instantly recycled back into the ISM, so the stellar mass production rate is actually:

$$\dot{M}_* = (1 - R)\psi. \tag{4.14}$$

The star formation process is regulated by the injection of energy from stars into the ISM. The importance of such feedback processes was recognised early on in the development of the theory of galaxy formation. White & Rees (1978) noted that the faint-end slope of the luminosity function computed in their model is steeper than the observed one, hinting at the existence of processes that decrease the efficiency of star formation in low mass halos. One of the processes that can be invoked for this purpose is the ejection of cold gas from galaxies by supernova driven winds (e.g. Larson, 1974; Dekel & Silk, 1986; Bower et al., 2001). A detailed treatment of this procedure would need to include a multiphase ISM, track the collisions between cold gas clumps and their evaporation by supernova heating (e.g. McKee & Ostriker, 1977; Efstathiou, 2000). However, semi-analytic models include this process with simple parametric relations. In GALFORM, the amount of gas ejected from the disc is proportional to the star-formation rate (Cole et al., 2000):

$$\dot{M}_{eject} = \beta\psi, \tag{4.15}$$

where,

$$\beta = (V_{\rm disc}/V_{\rm hot})^{-\alpha_{\rm hot}}, \tag{4.16}$$

where $V_{\rm disc}$ is the circular velocity of the galactic disc at the half-mass radius, and $V_{\rm hot}$ and $\alpha_{\rm hot}$ are input parameters, the former having units of velocity. The $V_{\rm hot}$ parameter controls which galaxies are significantly affected by the feedback: when $V_{\rm disc} = V_{\rm hot}$ (i.e. $\beta = 1$), for every solar mass of stars formed another solar mass of cold gas is ejected

from the galactic disc. The dependence of $V_{\rm disc}$ comes from the amount of energy needed for the gas to escape the gravitational potential well of the galaxy, characterised by the circular velocity of the disc. For example, if the galactic potential well is assumed to be self-similar, the required energy to eject gas from the disc should scale as $V_{\rm disc}^2$ (Benson et al., 2003), implying $\alpha_{\rm hot} = 2$ in Eq. (4.16). In practise, the value of $\alpha_{\rm hot}$ is allowed to vary between models, e.g. Cole et al. (1994) had $\alpha_{\rm hot} = 5.5$, Cole et al. (2000) uses $\alpha_{\rm hot} = 2$ and Bower et al. (2006) use $\alpha_{\rm hot} = 3.2$.

The fusion processes that fuel stellar energy production change their composition and, together with supernova explosions, are responsible for the formation of elements heavier than helium (beyond the amounts produced in during early Universe nucleosynthesis). These heavy elements are commonly referred in astrophysics as metals. In GALFORM, the production of metals is given by the yield, p, which is a fraction of fraction of stellar mass formed that returns to the ISM as metals. A fraction, e, of the metals is ejected straight out of the galaxy into the hot gas component, while the remainder, $(1-e)$, goes into the cold gas in the galactic disc. The latter can still contribute to the enrichment of the hot gas due to the feedback driven ejection, Eq. (4.15). Note that the majority of GALFORM works assumes $e = 0$ (e.g. Cole et al., 2000; Benson et al., 2001; Baugh et al., 2005; Bower et al., 2006; Lacey et al., 2008). The yields remain an uncertain quantity so it is treated as an input parameter chosen to reproduce the observed metallicities in local elliptical galaxies. For example, Cole et al. (2000) originally assumed $p = 0.02$. Later models, e.g. Baugh et al. (2005), use the yields predicted by stellar nucleosynthesis models for a given choice of IMF.

4.2.5 Stellar population synthesis and dust extinction

In order to compare stellar populations of GALFORM galaxies to the observable quantities, we must assign luminosities to the stars. This is done with stellar population synthesis techniques, originally introduced by Tinsley (1972) and later developed by various groups (e.g. Bruzual & Charlot, 1993; Bressan et al., 1994; Bruzual & Charlot, 2003; Maraston, 2005).

In general, the stellar population synthesis starts with spectral energy distributions (hereafter, SED), $l_\lambda(t, Z)$ for stellar populations that have a given age, t, and metallicity, Z. The SED of a model galaxy is obtained by convolving l_λ with its star formation

history:

$$L_\lambda(t) = \int_0^t l_\lambda[t - t', Z(t')]\psi(t')\mathrm{d}t', \tag{4.17}$$

where $Z(t')$ is the metallicity of stars forming at time t' and $\psi(t')$ is the star formation rate at that time.

The key element of any stellar population model is the choice of the initial mass function (hereafter, IMF) of the stellar population. The IMF gives the number of stars of a given mass that form in an episode of star formation. It is commonly given in the following form:

$$\frac{\mathrm{d}N}{\mathrm{dln}m} \propto m^{-x}, \tag{4.18}$$

where N is the number of stars formed and m is the mass of the star. The parameter x is chosen by comparing with observations. Salpeter (1955) found $x = 1.35$ to be a good match for the solar neighbourhood for stars with masses $m \gtrsim 2\,M_\odot$. Kennicutt (1983) studied the IMF in disc galaxies finding $x = 0.4$ for $m < M_\odot$ and $x = 1.5$ for $m > M_\odot$. The Kennicutt (1983) IMF, found to also agree well with more recent studies of the local IMF (Kroupa, 2001), is the default choice in the GALFORM model. The choice of IMF is a crucial component in the production of ionizing photons, as we discuss in Section 4.3.3.

Also, we note that Cole et al. (2000) included the fact that a fraction of star forming mass will end up in brown dwarfs ($m < 0.1M_\odot$), which do not contribute any light to the stellar population SED. The fraction of brown dwarfs is specified by the parameter Υ as:

$$\Upsilon = \frac{(\mathrm{mass\,in\,visible\,stars\,+\,brown\,dwarfs})}{(\mathrm{mass\,in\,visible\,stars})}, \tag{4.19}$$

at the time a stellar population forms. By definition, $\Upsilon \geq 1$. The parameter is used to decrease the luminosities of the population by $1/\Upsilon$. Cole et al. (2000) show that a choice of a small value of $\Upsilon \lesssim 2$ is a good match to the observed local stellar populations. More recent models (Baugh et al., 2005; Bower et al., 2006) found that the inclusion of brown dwarfs is not necessary for reproducing observations, as their role is replaced by the additional feedback mechanisms introduced to match the observations of high-mass galaxy counts.

Finally, for the modelled galaxies to match the observations, it is necessary to include the effect of dust extinction. The presence of dust affects the luminosities and colors of

galaxies, especially the UV luminosities used to estimate the star formation rate. The original model of Cole et al. (2000) used the extinction tables of Ferrara et al. (1999), obtained from MC radiative transfer simulations that include absorption and scattering on dust particles. More recent versions of GALFORM employ a more detailed model of Silva et al. (1998), which better deals with star formation bursts and can compute the spectrum of energy re-radiated by dust in the far-infrared and submillimetre.

4.2.6 Galaxy mergers

The GALFORM model assumes that the elliptical galaxies and bulges in disc galaxies are a consequence of mergers between galaxies. The key concept for understanding the mergers is *dynamical friction*, discussed in detail by Chandrasekhar (1943). This term relates to the loss of momentum and kinetic energy of moving celestial bodies (be it stars or galaxies) due to gravitational interaction with surrounding matter. Dynamical friction caused by surrounding dark matter erodes galactic orbits which may lead to a collision (e.g. Binney & Tremaine, 1987; Sparke & Gallagher, 2007).

In the GALFORM model, when two dark matter halos merge, the central galaxy of the most massive halo automatically becomes the central galaxy of the new halo, while the others become satellite galaxies with randomly assigned orbits. The satellite galaxies merge with the central one after the time it takes for the dynamical friction to decay their orbits. This time depends on the initial specific energy and angular momentum of the orbit and on the galaxy and halo masses. The GALFORM scheme sets the merging timescale for each satellite using the results of Lacey & Cole (1993), who estimate the time for an orbit to decay in an isothermal halo:

$$\tau_{\rm mrg} = f_{\rm df}\Theta_{\rm orbit}\tau_{\rm dyn}\frac{0.3722}{\ln(\Lambda)}\frac{M_{\rm H}}{M_{\rm sat}}. \tag{4.20}$$

Here, $M_{\rm H}$ is the mass of the halo in which the satellite orbits, $M_{\rm sat}$ is the mass of the satellite galaxy plus the mass of the halo in which the galaxy *formed* (see Navarro et al., 1995, for more discussion) and $\ln(\Lambda)$ is the Coulomb logarithm which depends on the range of impact parameters of the collisions (Binney & Tremaine, 1987) and is taken to be equal to $\ln(M_{\rm H}/M_{\rm sat})$. The dynamical timescale of the halo is $\tau_{\rm dyn} \equiv \pi r_{\rm vir}/V_{\rm H}$. The factor $\Theta_{\rm orbit}$ describes the dependence of $\tau_{\rm mrg}$ on the orbital parameters:

$$\Theta_{\rm orbit} = [J/J_c(E)]^{0.78}[r_c(E)/r_{\rm vir}]^2, \tag{4.21}$$

where E and J are the initial energy and angular momentum of the satellite's orbit and $r_c(E)$ and $J_c(E)$ are the radius and angular momentum of a circular orbit with the same energy. The time it takes for every satellite to merge with the central galaxy is calculated using Eq. (4.20) resulting in a list of galaxy mergers. If a satellite does not merge during the halo's lifetime, it is assigned a new random orbit and thus new $\tau_{\rm mrg}$ in the next halo.

For every merger event, the GALFORM scheme has two scenarios depending on the ratios of the satellite and central galaxy masses[2], $M_{\rm sat}$ and $M_{\rm cen}$, based on the prescription used by Kauffmann & White (1993b) and Baugh et al. (1996). The first type of mergers, labelled *major* mergers, occur when $M_{\rm sat}/M_{\rm cen} \geq f_{\rm ellip}$, where $f_{\rm ellip}$ is an input parameter, the choice of which is motivated by numerical simulations of galaxy mergers (e.g. Barnes & Hernquist, 1992; Walker et al., 1996; Barnes, 1998). When this type of merger occurs, the disc is assumed to be destroyed, producing a bulge or elliptical galaxy. All the cold gas found in the disc at the time of the merger is converted into stars in a starburst which assumes a shorter timescale, related to the dynamical time of the formed spheroid. The second type, called *minor* mergers, occurs when $M_{\rm sat}/M_{\rm cen} < f_{\rm ellip}$. In the original Cole et al. (2000) model, this type of merger does not trigger starbursts. Instead, the stars of the satellite galaxy are added to the bulge of the central one and its cold gas is added to the central galaxy's main cold gas disc without changing the disc's specific angular momentum. The Baugh et al. (2005) model, which we will discuss in more detail shortly, allows for minor mergers to trigger starbursts if $M_{\rm sat}/M_{\rm cen} > f_{\rm burst}$ (while $f_{\rm burst} < f_{\rm ellip}$) and the gas fraction in the central galaxy is larger than $f_{\rm gas,crit}$. This choice is motivated by the simulations of Hernquist & Mihos (1995) who found gas-rich discs to be susceptible to SF bursts induced by the accretion of small satellites. Baugh et al. (2005) adopt $f_{\rm ellip} = 0.3$, $f_{\rm burst} = 0.05$ and $f_{\rm gas,crit} = 0.75$.

4.2.7 Photo-ionization feedback

The presence of ionizing radiation affects the formation of galaxies through two processes. First, the photo-heating of the IGM increases the pressure of the gas which both prevents its collapse in dark matter halos and evaporates the gas already collapsed into halos with shallow potential wells. The net effect of this process is that low mass halos have a significantly lower baryon fraction, $f_{\rm b}$, than the cosmic mean, $\langle f_{\rm b} \rangle = \Omega_{\rm b}/\Omega_0$.

[2]For this purpose, the mass of a galaxy includes only stars and cold gas.

The important role of the photo-heating in galaxy formation was first suggested by Doroshkevich et al. (1967) and later examined in the CDM framework by Couchman & Rees (1986). Since this initial discussion, many authors have focused on quantifying this effect using semi-analytic calculations (e.g. Babul & Rees, 1992; Efstathiou, 1992; Shapiro et al., 1994; Nagashima et al., 1999; Benson et al., 2002a,b; Somerville, 2002), idealised 1D simulations (e.g. Thoul & Weinberg, 1996; Kitayama et al., 2000; Mesinger & Dijkstra, 2008) and full 3D cosmological hydrodynamical simulations (e.g. Quinn et al., 1996; Weinberg et al., 1997; Gnedin, 2000a; Okamoto et al., 2008). The role of self-shielding of gas to ionizing radiation has also been explored using radiative hydrodynamics simulations (e.g. Susa & Umemura, 2004a,b). The exact value of the characteristic mass, M_c, below which galaxy formation is strongly affected by photo-heating is still under considerable debate. An important observational constraint for the choice of M_c is the number of satellite galaxies in the Milky Way (e.g Martin et al., 2007; Koposov et al., 2008), as they are thought to be significantly affected by the photo-heating suppression.

The second effect of ionizing radiation on galaxy formation comes from the ionization of gas found in halos. The ionization of hydrogen removes the most important cooling channel (the H atomic line radiative cooling, see cooling function in Sutherland & Dopita, 1993) in halos of $T_{\rm vir} \approx 10^4 K$, effectively suppressing star formation in them.

The role of photo-ionization feedback was examined with the GALFORM scheme in a series of papers by Benson et al. (2002a,b, 2003) and expanded on by Benson et al. (2006). The model of Cole et al. (2000) ignores the IGM pressure and assumes that the mass of gas that collapses into a halo of mass $M_{\rm tot}$ to be a equal to the cosmic mean baryon fraction, $M_{\rm gas} = (\Omega_{\rm b}/\Omega_0) M_{\rm tot}$. Benson et al. include the photo-heating with the prescription introduced by Gnedin (2000a), in which the collapsed baryon fraction is a function of the *filtering mass*, $M_{\rm F}$, which is a characteristic mass over which the baryonic density fluctuations are smoothed in the presence of dark matter derived using linear theory. The simulations of Gnedin (2000a) found that when $M_{\rm F} = M_{\rm tot}$, then $M_{\rm gas} = 0.5 \langle f_{\rm b} \rangle M_{\rm tot}$, and decreasing for $M_{\rm tot} < M_{\rm F}$. The existence of an analytic relation for $M_{\rm F}$ made it a favourite criterion for the effect of photo-heating in semi-analytic galaxy formation models (e.g. Benson et al., 2002a,b; Somerville, 2002), but later numerical simulations have found it to be somewhat of an overestimate (e.g. Hoeft et al., 2006; Okamoto et al., 2008). To compute the filtering mass, Benson et al. estimate a cosmically

averaged ionization and temperature fields from the Lyman-continuum luminosities of galaxies. The same estimate is included in the calculation of gas cooling, as mentioned above.

The work of Benson et al. has shown that the effect of photo-heating during reionization on the subsequent galaxy formation can be approximately reproduced by a simpler model than the one they employed - by assuming that all gas cooling is suppressed ($\dot{M}_{\rm cool} = 0$) for redshifts $z < z_{\rm cut}$ in halos with circular velocity $V_{\rm c} < V_{\rm cut}$. Later GALFORM models use this implementation to model the effect of reionization and we do the same in this thesis. The values of $z_{\rm cut}$ and $V_{\rm cut}$ are input parameters, but the exact choices are motivated by the observations (e.g. redshift of reionization suggested by the CMB polarisation data, Dunkley et al., 2009) and simulations (e.g. the filtering mass of Gnedin (2000a) corresponds to $V_{\rm cut} \approx 60\,{\rm km/s}$). We investigate the effect the choices of $z_{\rm cut}$ and $V_{\rm cut}$ have on the ionizing emissivity in Section 4.3.8.

4.2.8 Two fiducial GALFORM models

The model of Cole et al. (2000) has been further refined by a number of authors. In this section we present the key features of two such models, Baugh et al. (2005) and Bower et al. (2006), and compare them in terms of ionizing photon production. These two models can be labelled as "default" models as they have been a starting point for almost all following GALFORM related work (e.g. Bower et al., 2008; Font et al., 2008; Orsi et al., 2008; Lacey et al., 2009).

Baugh et al. 2005

The work of Baugh et al. (2005) (we refer to this model as Baugh05 hereafter) was motivated by the failure of previous GALFORM models (Cole et al., 2000; Granato et al., 2000; Benson et al., 2003) to reproduce the observed counts of submillimetre galaxies (SMGs), a population of high redshift galaxies detected by emission from warm dust in the rest-frame far-IR/submillimetre bands (see Blain et al., 2002, for a review), already mentioned in the introduction. They are thought to be young star-forming galaxies embedded in a thick clouds of dust. The dust envelope is heated by the UV radiation of the young massive stars and reemits the absorbed energy as IR radiation. Estimates of their star formation rate are not straight forward due to the UV part of the SED being

almost completely obscured. Also, it is possible that a part of the submillimetre emission is powered by active galactic nuclei (AGN), though X-ray observations suggest this is fraction is small, $\lesssim 10\%$ (Alexander et al., 2003).

The most significant change introduced in Baugh05 is the use of a *top-heavy IMF* during galaxy-merger induced star formation bursts. The galaxy merger model in GALFORM was introduced in Section 4.2.6. In the top-heavy IMFs, as hinted by their name, the number of massive stars produced is assumed to be significantly higher than what is observed in the solar neighbourhood. In practise, this is achieved by choosing a flatter IMF by modifying the parameter x in Eq. (4.18). The Baugh05 model uses an extreme value of $x = 0$ during SF bursts, but still employs the Kennicutt (1983) IMF for quiescent star formation (the Kennicutt IMF parameters were introduced in Section 4.2.5). The top-heavy IMF increases the SMG counts in two ways. First, as more massive stars are formed, the UV luminosity is significantly higher, e.g. the energy at 1500 Å released per unit stellar mass formed is about 4 times higher for the top-heavy IMF than for the Kennicutt one. Clearly, the top-heavy IMF is crucial for the ionizing photon emissivity, as we will discuss in Section 4.3.3. Second, the top-heavy IMF results in a higher metal yield p, which increases the absorption of UV radiation by dust. Combined, this results in a more efficient heating of the dust, boosting sub-mm emission. Baugh et al. (2005) find both of these effects to be crucial for reproducing the SMG luminosity functions: a simple increase of the UV luminosity is not enough.

To maximise the impact of the top-heavy IMF during bursts on luminosities of high-z galaxies, the merging discs must have a significant pool of cold gas available for star formation triggered by the burst. The Baugh et al. (2005) model accomplishes this scenario through its prescription for the star formation timescale, i.e. the denominator in Eq. (4.13):

$$\tau_* = \tau_{*0}(V_{\rm disc}/200\,{\rm km/s})^{\alpha_*}, \tag{4.22}$$

where τ_{*0} and α_* are constant parameters. This definition is in contrast from the one introduced in Cole et al. (2000), which depends on the dynamical time of the galactic disc. There are no strong arguments to prefer one prescription over the other, as both are constrained by the same observed gas fraction-luminosity relation at $z = 0$ (Fig. 9 of Cole et al. (2000)). Baugh et al. (2005) chose $\tau_{*0} = 8\,{\rm Gyr}$ and $\alpha_* = -3$. With this choice of parameters, the star formation timescale is roughly equal to the one defined by Cole et al.

(2000) at low redshifts, but it is much longer at high redshifts of interest for the SMGs. This results in a less efficient quiescent star formation which leaves more gas available for the star formation in bursts. The model employs a different SF timescale during SF bursts:

$$\tau_{*\mathrm{burst}} = \max[f_{\mathrm{dyn}}\tau_{\mathrm{dyn,bulge}}, \tau_{*\mathrm{burst,min}}], \tag{4.23}$$

where $\tau_{\mathrm{dyn,bulge}}$ is the dynamical time of the spheroid created in the merger that triggers the burst. The Baugh05 model sets $f_{\mathrm{dyn}} = 50$ and $\tau_{*\mathrm{burst,min}} = 0.2\,\mathrm{Gyr}$ to produce the best match to the present-day 60-μm luminosity function and the abundance of SMGs. The burst SF timescale is typically much shorter than the quiescent one, given by Eq. (4.22).

On a final note, the Baugh05 model includes feedback by supernova driven superwinds, introduced in Benson et al. (2003), to produce a better agreement of luminosity functions with the observations at the high luminosity end. This model component is not crucial for calculating the production of ionizing photons during reionization, as we discuss in Section 4.3.7.

Bower et al. 2006

The key new element introduced by Bower et al. (2006) (hereafter, Bower06) is the feedback due to injection of energy by AGN. This model, together with similar work of Croton et al. (2006), has shown that AGN feedback provides an answer to a number of outstanding questions in the theory of galaxy formation: what shapes the high mass end of the luminosity function at $z = 0$, why was the star formation rate in massive galaxies larger at higher redshift than it is in the local Universe (the so-called "cosmic downsizing" problem, e.g. Cowie et al., 1996) and what suppresses the cooling flows at the centres of rich clusters that are expected to form using the cooling time arguments discussed in Section 4.2.3.

The Bower06 model is significantly different from the Baugh05 model. The already mentioned AGN feedback replaces the superwind feedback as a tool for reproducing the high mass end of the luminosity function. In the model, the growth of the black holes at high redshift is primarily driven by disc instabilities, which we discuss in Section 4.3.6. Bower06 does not use a different IMF in bursts, employing Kennicutt (1983) throughout the model instead. To compute the star formation rate, the model uses the Cole et al. (2000) parametrisation for the star formation timescale, which depends on the galactic

disc dynamical timescale:

$$\tau_* = \eta_*^{-1} \tau_{\rm disc} (V_{\rm disc}/200\,{\rm km/s})^{\alpha_*}, \tag{4.24}$$

where η_* and α_* are dimensionless parameters and the disc dynamical timescale is defined as $\tau_{\rm disc} \equiv r_{\rm disc}/V_{\rm disc}$. Disc radius $r_{\rm disc}$ and rotational velocity $V_{\rm disc}$ are defined at the half-mass radius. The exact choice of parameters is $\eta_* = 0.002857$ and $\alpha_* = -1.5$. The Bower06 model also has significantly stronger supernova feedback, discussed in Section 4.3.4. The Bower06 model focuses on different problems in the theory of galaxy formation than Baugh05 and thus does not provide a satisfactory model for high redshift SMGs and Lyman-break galaxies, as we show in Section 4.4.

4.3 Lyman continuum emissivity from GALFORM

The nature of sources that dominate the Lyman continuum emissivity during reionization remains uncertain. A range of models, both semi-analytical (e.g. Fukugita & Kawasaki, 1994; Shapiro et al., 1994; Haiman & Loeb, 1997; Ciardi et al., 2000; Wyithe & Loeb, 2003; Benson et al., 2006) and numerical (e.g. Gnedin & Ostriker, 1997; Gnedin, 2000b; Sokasian et al., 2003; Ciardi et al., 2003; Iliev et al., 2006a), have shown that reionization can be achieved with stellar sources alone by $z \sim 6-15$. The notion of a Universe reionized only by high redshift galaxies hosting Population II stars is an attractive one due to them being far better understood than some more exotic source candidates such as mini-quasars (e.g. Haiman & Loeb, 1998; Valageas & Silk, 1999; Wyithe & Loeb, 2003) or Population III-type stars (e.g. Abel et al., 2002a; Bromm & Larson, 2004). If this is the case and galaxy formation is not drastically different at the high redshifts relevant to reionization study, the semi-analytic models such as GALFORM can be used to model the sources of LC radiation. Furthermore, using such detailed models is a large improvement over most current work that model sources with simple mass-luminosity relations applied to dark matter catalogues (e.g. Iliev et al., 2006a; McQuinn et al., 2007).

As we mentioned earlier, GALFORM was previously used to model reionization in Benson et al. (2006) (see also Benson et al., 2001, 2002a,b, for their earlier work on the subject). They improve the cooling calculation (described in Section 4.2.3) by introducing cooling due to molecular hydrogen, H_2, as well as more detailed treatment of photoionization heating feedback, and model reionization with a range of different models.

Our approach in this section differs from theirs in several respects. First, they do not attempt to include the effects of inhomogeneous reionization in their models, while we do this by coupling GALFORM with an RT code in Chapter 5. Because of this limit, Benson et al. focus their study on the total emissivity of ionizing photons and the global growth of the ionized fraction. As the proper modeling of HII region spatial distribution is one of our main motivations, we study in more detail where exactly the ionizing photons are coming from in GALFORM models. Second, at this stage, we chose not to include the mentioned improvements introduced in Benson et al. (2006). Instead, we examine the predictions of "default" models that have been calibrated in detail using lower redshift data. We will examine the effect of a more detailed treatment of basic physics in future work. Finally, we present comparison of GALFORM results with latest observations that have become available since the publishing of Benson et al. (2006).

This section is concerned with examining the effect of various GALFORM parameters on the production of ionizing photons. First, in Section 4.3.1 we compare the number of ionizing photons produced in Baugh05 and Bower06 and show that Bower06 does not produce enough photons to ionize the Universe by the redshift broadly suggested by observations. Next, we show the evolution of emissivity with redshift in the default Baugh05 model in Section 4.3.2. Here we also introduce two figures we will be using frequently in the following discussion. In the rest of the section, we discuss the role of starbursts (Section 4.3.3), supernova feedback (4.3.4), star formation timescale (4.3.5), disc instabilities (4.3.6), some sub-dominant effects (AGN and superwind feedbacks, metal production, Section 4.3.7) and photo-ionization feedback (4.3.8). We will return to some of the more important parameters in Section 4.4 where we compare the results of GALFORM with the latest observations of Lyman-break galaxies at $z \lesssim 10$.

4.3.1 The production of ionizing photons in the two default models

Here, we compare the production of ionizing photons in the two default models presented in the previous paragraphs. The main variable of interest is the ionizing *emissivity* of Lyman-continuum (LC) photons, i.e. the total number of photons capable of ionizing hydrogen ($h\nu \geq 13.6\,\mathrm{eV}$) emitted per unit time per unit co-moving volume:

$$\varepsilon = \frac{\sum_{i=1}^{N_{\mathrm{galaxies}}} \dot{N}_{\mathrm{LC,i}}}{V_{\mathrm{sample}}}, \tag{4.25}$$

where $\dot{N}_{\rm LC,i} = \int_{\nu_{\rm th}}^{\infty} L_{\nu,i}/h\nu {\rm d}\nu$ is the number of hydrogen-ionizing photons emitted from a single galaxy per unit time, $L_{\nu,i}$ is the luminosity per unit frequency ν of the said galaxy (in units of erg s^{-1}) and $N_{\rm galaxies}$ is the number of galaxies found in the volume $V_{\rm sample}$. The emissivity ϵ is a function of time and is intimately tied to the formation and evolution of galaxies.

To get a first order estimate of the reionization process, one can track the total number of ionizing photons produced up to some moment in time. Figure 4.1 shows how the total number of ionizing photons produced per hydrogen atom (both neutral and ionized) evolves with redshift in the two default models, using MC merger trees (Section 4.2.1) and assuming the same cosmological parameters as used in the Millennium [3] (Springel et al., 2005b). The minimum requirement for reionization is that at least one ionizing photon is produced per hydrogen atom in the IGM, marked by the bottom dashed line in Figure 4.1. It is clear that the Baugh05 model is much more effective in the production of ionizing photons: by $z \approx 10$ it produces ≈ 100 times more photons than the Bower06 model. The gap between the two models decreases with redshift and by $z = 5$ Bower06 produces only ≈ 15 times fewer photons than Baugh05. The minimum requirement of one photon per H atom is unrealistic, as it ignores all the possible sinks of ionizing photons starting with the escape of ionizing UV radiation from galaxies themselves. If we assume an escape fraction of 10%, in line with current simulations and observations mentioned in Chapter 1, then 10 photons must be produced to ionize a single hydrogen atom (marked by the top dashed line in Figure 4.1). In that case, the Bower06 model is not able to produce enough ionizing photons to ionize all hydrogen by $z \approx 6$, the end of reionization suggested by quasar data (e.g. Fan et al., 2002, see Chapter 1). Including recombinations and self-shielded structures increases the needed number of ionizing photons moving the redshift of reionization in the Bower06 model to even lower redshifts.

Following the result shown in Fig. 4.1, we conclude that the Bower06 model will not be able to produce enough photons to reionize hydrogen by $z \approx 6$. Therefore, we chose the Baugh05 model as the default model for our numerical reionization runs we

[3] $\Omega_0 = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.045$ and $h = 0.73$. Note that we chose to use this cosmology because it is the same cosmology employed by the Millennium-II (Boylan-Kolchin et al., 2009) simulation we will use for future reionization studies. The original Bower06 model is optimised for this cosmology, but Baugh05 assumed a slightly different one.

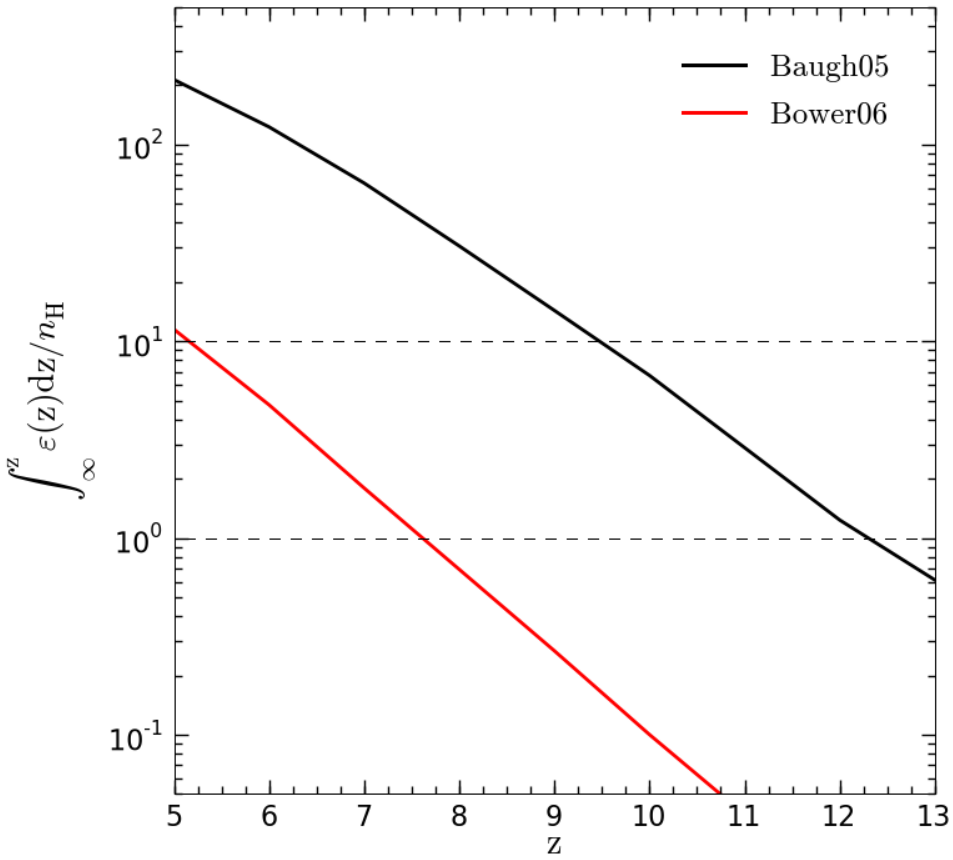

Figure 4.1: The number of ionizing photons produced per hydrogen atom up to redshift z in the two default GALFORM models, Baugh05 and Bower06. The horizontal dashed lines mark the minimum number of photons per H atom that must be produced to achieve reionization: in the most optimistic case, only one (bottom line), but 10 or more when the escape factor is taken into account (top line). The Baugh05 model produces ≈ 100 times more ionizing photons at $z \approx 10$ and reaches one photon per H atom $\Delta z \gtrsim 4$ earlier.

will present in Chapter 5 and the rest of this Chapter. The differences between the two models are many, but we can point to the major ones that are responsible for the vastly different LC photon production. First, the top-heavy IMF assumed in Baugh05 yields approximately 10 times as many photons produced during bursts (as predicted by stel-

lar population synthesis model of Bressan et al., 1998, for a typical stellar metallicity of $Z/Z_{\odot} \approx 10^{-2}$ in discs we find during EoR) than when Kennicutt IMF is assumed (see Section 4.3.3). Second, the Bower06 model assumes a significantly stronger supernova feedback, which decreases the number of low-mass galaxies (discussed in Section 4.3.4. Finally, the quiescent star formation timescale in Bower06 is significantly shorter than in Baugh05, which results in less star formation occurring in starburst mode (see Section 4.3.5). These parameters explain the bulk, but not all, of the difference in LC photon production. The role of these and other parameters will be discussed in the following section.

4.3.2 Evolution of ionizing emissivity in the Baugh05 model

As discussed in Section 4.2.8, we chose the Baugh05 model to be our reference model for the following work, due to Bower06 not producing enough ionizing photons to satisfy the current observational constraints on reionization. Here we show how the production of LC photon by halos of a certain mass evolves with redshift. The LC photon production as a function of halo mass will be the key variable we discuss in the rest of this section.

Figure 4.2 shows the evolution of the halo LC mass-luminosity relation as a function of redshift. The discussion in terms of halos instead of galaxies is chosen for easier comparison with other numerical reionization work (e.g. Iliev et al., 2005a; McQuinn et al., 2007) who estimate the LC luminosity directly from halo properties. We obtain the LC luminosity of each halo as the sum of luminosities of all of the galaxies it hosts. The mean of the distribution (thick lines) shows almost no evolution with redshift - the only significant feature is the drop-off at low masses ($M \lesssim 10^8 \mathrm{M}_{\odot}/h$) consistent with the increase of the minimum mass of halos that can cool by atomic hydrogen line cooling (i.e. halos with ($T_{\mathrm{vir}} \approx 10^4\,\mathrm{K}$) with decreasing redshift. A striking feature in the figure is the large difference between the mean and median mass-luminosity relations, which is caused by the fact that a small number of starburst galaxies dominate the production of LC photons. We focus on this issue in Section 4.3.3. Note that the median shows some evolution with redshift, with luminosity modestly decreasing with decreasing redshift. This is not matched in the mean which is dominated by starbursts. The shape of the mean mass-luminosity relation is also more complex than the linear or simple power-law relations assumed in works by other authors.

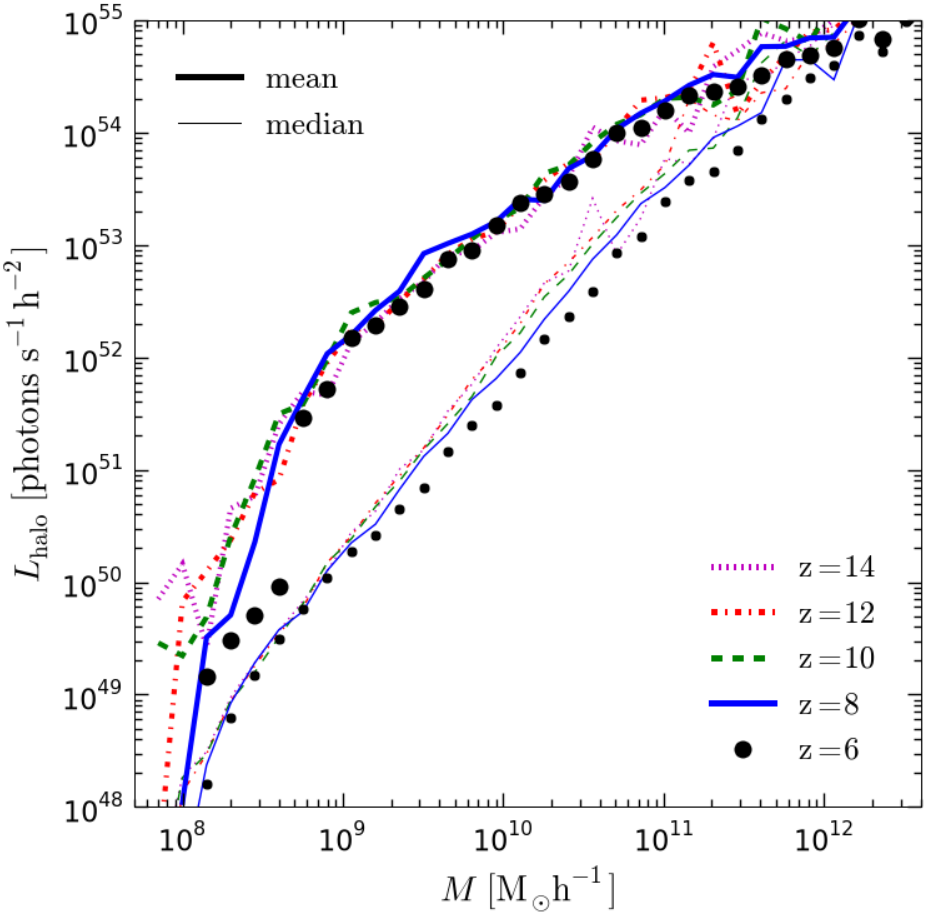

Figure 4.2: The LC luminosity of a halo as a function of its mass in the Baugh05 model at redshifts z =14 (magenta dotted lines), 12 (red dot-dashed), 10 (green dashed), 8 (blue solid lines) and 6 (black circles). The mean (thick lines) and median (thin lines) differ significantly due to the small number of starbursts dominating the LC photon production (discussed in more detail in Section 4.3.3). The mass-luminosity relation generally stays almost constant with redshift. The modest decrease with decreasing redshift of $L_{\rm halo}$ in the low mass end ($M \lesssim 5{\times}10^8 {\rm M}_\odot/h$) is because the mass of halos with $T_{\rm vir} = 10^4$ K (virial temperature needed for atomic line cooling) increasing with time. See text for more details.

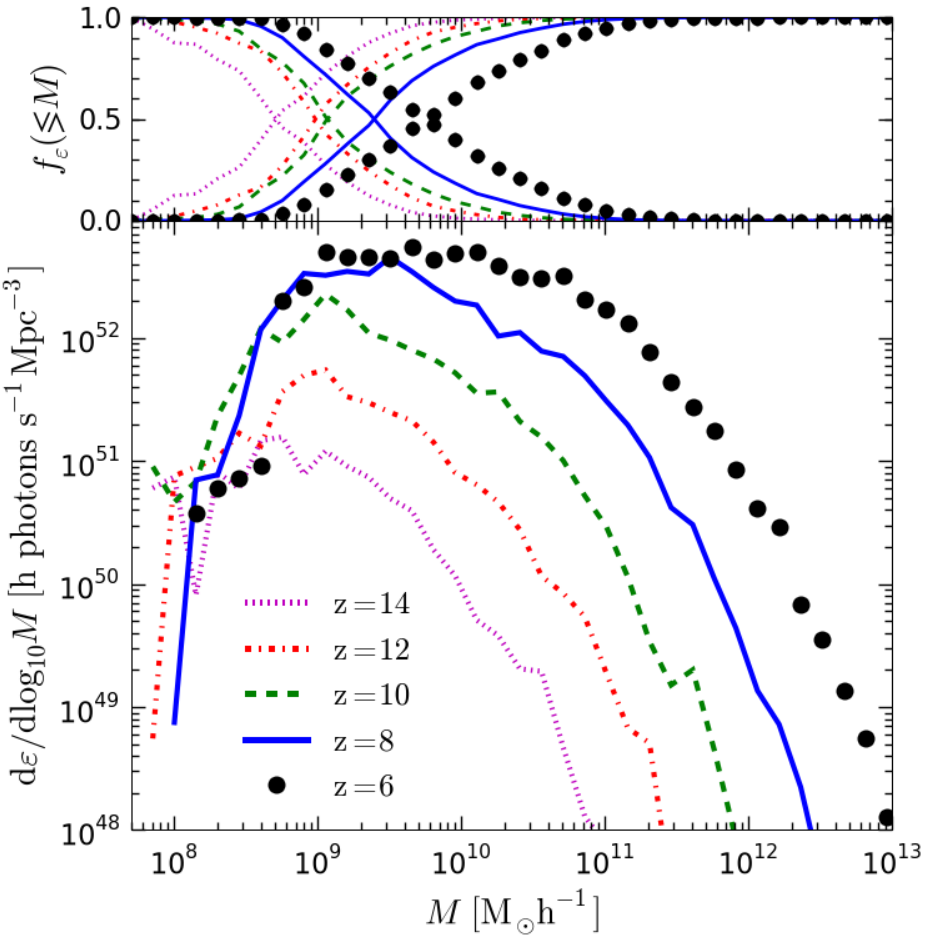

Figure 4.3: The LC emissivity per halo mass in the Baugh05 model as a function of redshift. Line styles correspond to results at different redshifts, as in Fig. 4.2. Since the mean mass-luminosity relation remains nearly constant (Fig. 4.2), the evolution of the emissivity is driven by the evolution of the halo mass function. The top panel shows the fraction of all LC photons produced by halos less (or more) massive than M. The intersection of the two lines for a given redshift gives the mass corresponding to the median halo mass for LC photon production. This typical mass for ϵ increases with decreasing redshift.

Figure 4.3 shows the differential LC emissivity (the number of LC photons produced by halos of a given mass) per halo mass as a function of redshift. Since the mean mass-luminosity relation does not evolve, as seen in Fig. 4.2, the increase of emissivity with

redshift seen in Fig. 4.3 is caused by the evolution of the halo mass function. A notable feature is therefore the increasing contribution to LC photon production from more and more massive halos. We demonstrate this in the top panel of Fig. 4.3, which shows the fraction of total LC emissivity produced by halos with masses smaller (larger) than M - the cross between the two lines (for a given redshift) gives the characteristic halo mass up to which 50% of LC photons are produced. The importance of low mass halos ($M \lesssim 10^9 \mathrm{M}_\odot/h$) clearly decreases with decreasing redshift. For example, halos of $M \lesssim 5 \times 10^8 \mathrm{M}_\odot/h$ are responsible for producing $\sim 50\%$ of LC photons at $z = 14$ (magenta dotted lines), but no more than $\sim 20\%$ at $z = 10$ (green dashed lines) and $\sim 2 - 3\%$ at $z = 6$ (black circles).

These two relations will be our main tool, together with the cumulative production of LC photons (as presented in Fig. 4.1) in the exploration of the importance of specific GALFORM parameters for ϵ in the following sections. We show the importance of the mass-luminosity relation of LC sources for HII region morphology in Chapter 5, Section 5.4.2.

4.3.3 Bursts and top-heavy IMF

As we discussed in Section 4.2.6, the mergers between galaxies are one of the crucial components of the GALFORM model. The bursts of star formation triggered by mergers may differ from the quiescent star formation, by assuming a different star formation time scale or IMF, as in the Baugh05 model (Section 4.2.8). To gauge the importance of both the IMF in bursts and the bursts themselves we ran three variations of the Baugh05 model: we turned off all SF bursts induced by mergers ($f_{\rm ellip} \gg 1$ and $f_{\rm gas,crit} > 1$), we allowed only major mergers to trigger bursts ($f_{\rm burst} = f_{\rm ellip}$) and we used the same solar neighbourhood, Kennicutt (1983) IMF for both bursts and quiescent star formation. Figure 4.4 shows the number of photons per hydrogen atom produced in these models as a function of redshift. The importance of bursts in the production of ionizing photons is clear: turning them off completely (the red line) results in about 10 times fewer photons produced than in the default Baugh05 model (black line) at all redshifts. If we assume that the reionization is achieved when 10 photons are produced for every atom[4] (top dashed line), the exclusion of bursts from the model results in $\Delta z \approx 4$ delay in the

[4]Note that we will use this fiducial value as the redshift of reionization throughout this chapter.

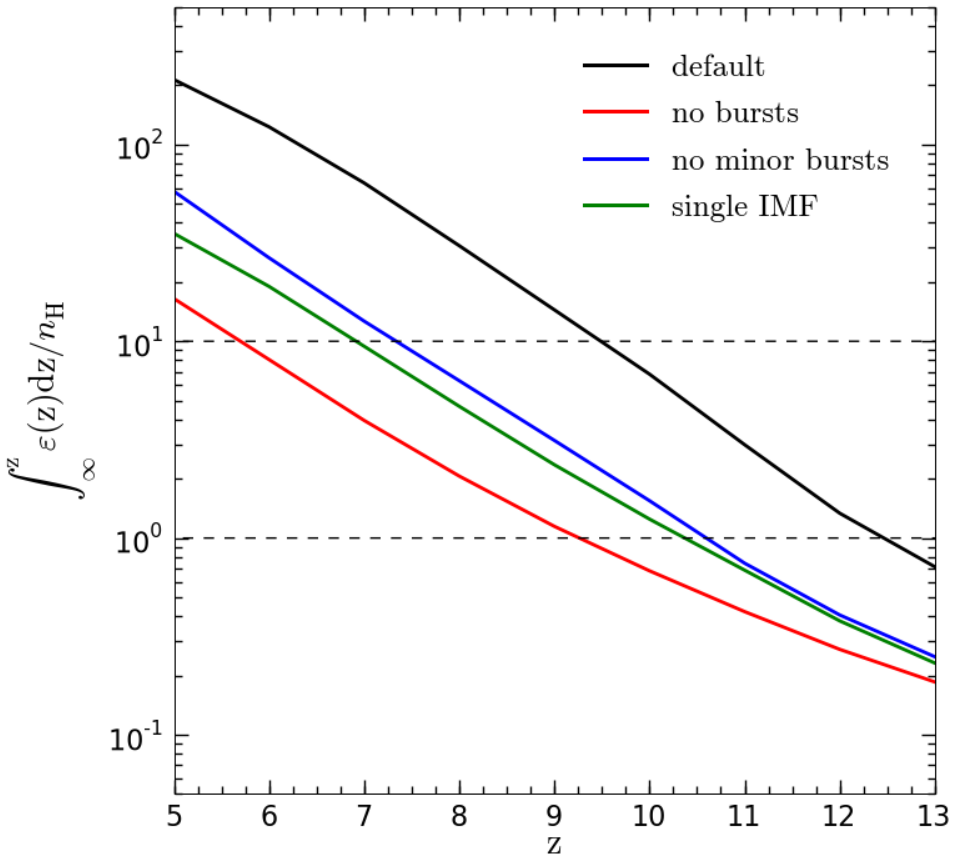

Figure 4.4: Number of LC photons per hydrogen atom produced up to a given redshift with respect to the variations of the burst parameters. Completely turning off the bursts (red line) reduces the total number of photons produced by ≈ factor 5-10, depending on the redshifts. About 35% of the photons in bursts are produced in major merger bursts alone (blue line). When only a single, Kennicutt (1983) IMF is used (green line), the bursts are responsible for producing $\approx 20 - 50\%$ of LC photons.

reionization redshift. The majority of these photons are due to the top-heavy IMF: when a single IMF is used (green line), only ≈ 2 times as many photons are produced as in the no bursts case and reionization is reached $\Delta z \approx 3.2$ later compared to the default model. Finally, minor merger bursts seem to be the more important type (also found by Baugh

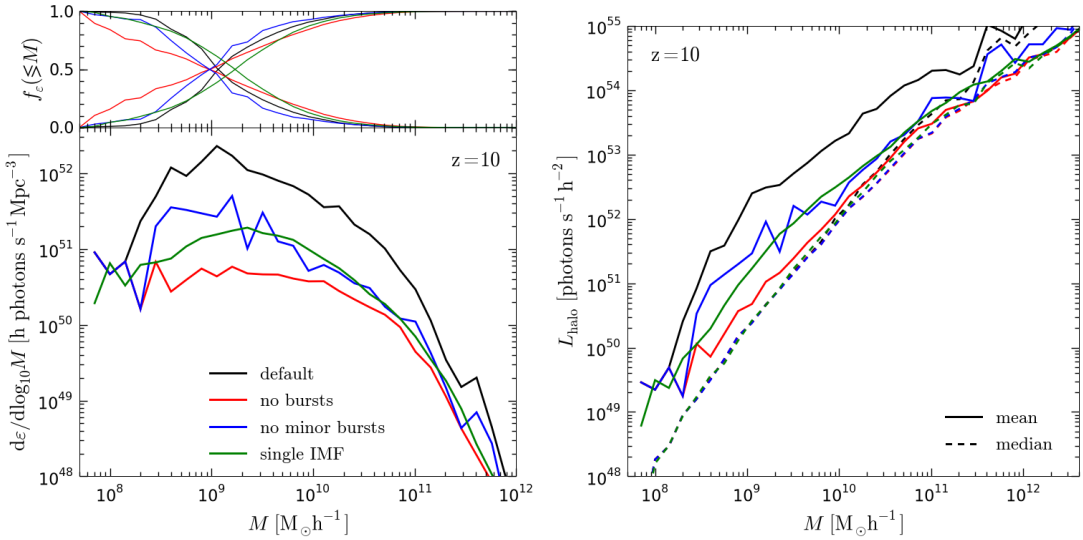

Figure 4.5: *Left:* The emissivity per halo mass with and without bursts, and with and without a top-heavy IMF. Turning off the bursts (red line) decreases the emissivity by more than an order of magnitude in halos with $M \approx 5\times10^8 - 5\times10^{10}\mathrm{M}_\odot/h$. The contribution of bursts themselves is smaller than the contribution of the top-heavy IMF (green line). Minor merger bursts (compare blue and black lines) contribute less than major merger bursts in the low mass halos ($M < 10^9\mathrm{M}_\odot/h$). *Right:* LC luminosity per halo mass. Reducing the importance of bursts reduces the skewness in the distribution as evidenced by the mean (solid lines) approaching the median (dashed lines). The median is generally unaffected by the burst parameters. See Fig. 4.6 for a closer look at the scatter in $L_{\rm halo}$.

et al. (2005) in explaining SMG number counts), as about 35% of all photons is produced in major mergers alone (blue line). The fact that major merger-only and a single IMF yield similar emissivities is a coincidence.

Further insight can be obtained by the plots in Fig. 4.5. The left panel shows the LC emissivity as a function of halo mass at $z = 10$. The values for different burst parameters agree at the atomic line cooling halo mass limit, $M \approx 10^8\mathrm{M}_\odot/h$, since no galaxies can form in smaller halos. A similar argument can be invoked to explain the dominance of major mergers in low mass halos (blue line at $M < 10^9\mathrm{M}_\odot/h$), as these have progeni-

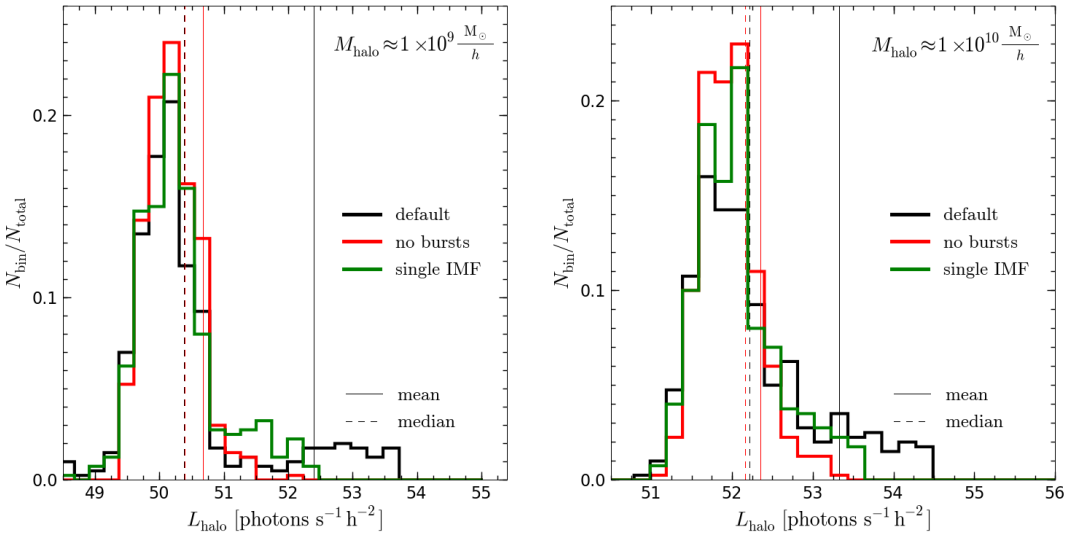

Figure 4.6: The fraction of halos with given LC luminosity $L_{\rm halo}$ in halos with mass $M \approx 10^9 {\rm M}_\odot/h$ (left panel) and $M \approx 10^{10} {\rm M}_\odot/h$ (right panel) in models with different burst parameters. When no bursts are included (red lines), the $L_{\rm halo}$ distribution is close to Gaussian, with very similar mean (vertical solid) and median (vertical dashed lines). The inclusion of bursts results in a small number of halos being significantly more luminous, more so in the default model (black lines), than in the one with a single IMF (green lines). This results in the shifting of the mean towards higher luminosities.

tors of similar masses and usually have only a single galaxy per halo. Interestingly, the characteristic halo mass marking the median of the emissivity (intersection of lines in the top panel) is not significantly affected by changing the burst parameters. A larger difference is seen in the relative contribution of different halo masses: the presence of bursts introduces a peak in the emissivity at moderate halo masses, $M \approx 10^9 {\rm M}_\odot/h$. The peak causes the majority of the photons to be produced in smaller range of halo masses. For example, the default Baugh05 model (black lines) produces 40% of all photons in the range $3 \times 10^8 \lesssim M \lesssim 10^9 {\rm M}_\odot/h$, while in the no bursts model (red lines), the same fraction is produced in the range $7 \times 10^7 \lesssim M \lesssim 10^9 {\rm M}_\odot/h$ (see top panel).

The right panel of Fig. 4.5 shows the ionizing luminosity of a halo as a function of its mass. The mean luminosity as a function of halo mass varies strongly with the burst

parameters, but the median is only slightly affected. This is again due to the merger induced starbursts dominating halo emissivities: a single bursting galaxy skews the mean toward higher luminosities even though the majority of galaxies have significantly smaller luminosities characterised by the median. This can clearly be seen when examining the luminosities in halos of approximately the same mass, as we do in Fig. 4.6. The majority of galaxies in all models have a close to Gaussian distribution centred close to the median (horizontal dashed lines). Models that include the starbursts feature a small "tail" of high luminosity halos, which causes the shift in the mean. The Poissonian nature of bursting halos, which in practise dominate the production of LC photons, should leave an imprint on the morphology of HII regions during reionization, as we show in Chapter 5. Coming back to the right panel of Fig. 4.5, notice that all models that have bursts in some form agree very well on the median of the distribution. This suggests that the median characterises the galaxies forming stars quiescently. As the only parameters we varied here relate to bursts, which occur in only a small fraction of halos at any time, the quiescent star formation is the same between models, thus explaining the equal medians. The bursts and, more importantly, the top-heavy IMF assumed in Baugh05 are crucial for achieving reionization before the end of the epoch suggested by the quasar data, $z \approx 6$ (Chapter 1). As Fig. 4.4 suggests, without the top-heavy IMF, reionization can barely be completed by that redshift. The criterion of 10 LC photons produced per H atom is quite liberal, as it corresponds to a relatively high escape fraction of 10% and neglects photon sinks completely. This means that the reionization will in full RT simulations probably be occurring later for the same source model parameters . The escape fraction remains unconstrained, so a faster reionization is still not out of the question, but it is safe to assume that the model without any bursts will not be able to reach reionization by $z \sim 6$ in a complete reionization model.

4.3.4 Supernova feedback

Supernova (SN) feedback is usually invoked in semi-analytic models of galaxy formation to produce a match with the low-luminosity end of the galaxy luminosity function. The galaxies in that luminosity range are the most likely sources of the majority of ionizing photons during reionization (since more massive galaxies are only appear later), so we expect the choice of SN feedback parameters to have a significant impact on the LC

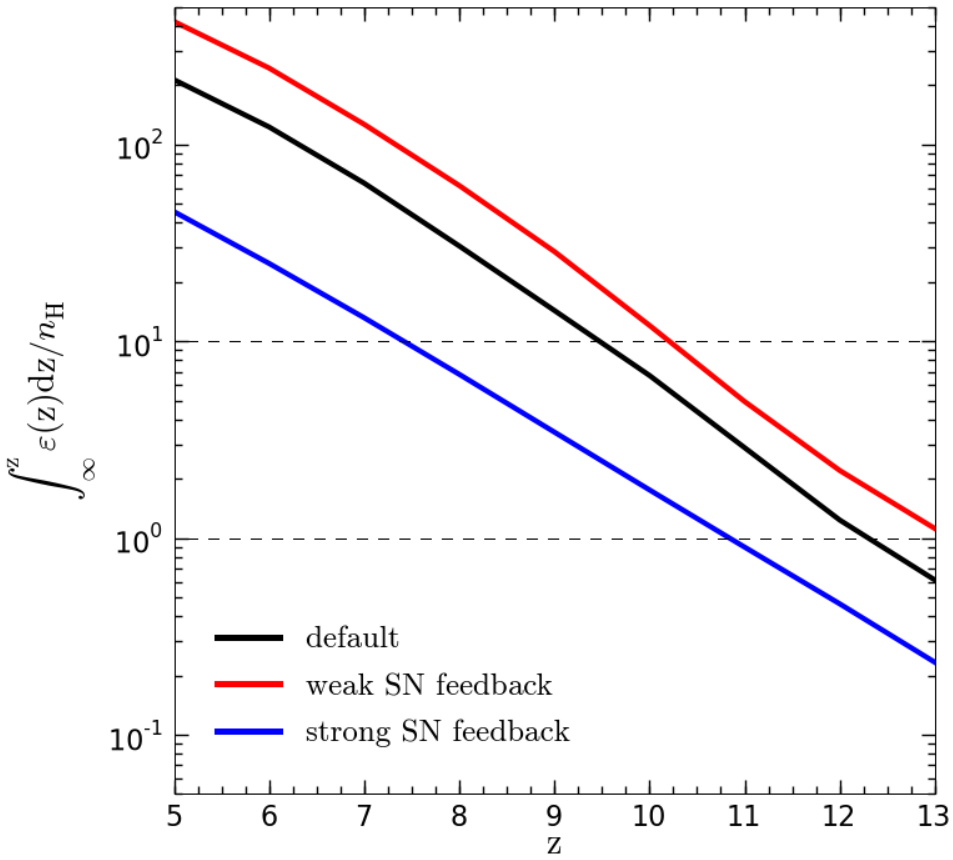

Figure 4.7: Number of LC photons per H atom emitted by redshift z for different SN feedback parameters. Decreasing the efficiency of SN feedback (red line) doubles the production of ionizing photons resulting in $\Delta z \approx 0.5$ earlier reionization. Increasing it to be similar to parameters in the Bower06 model results in $\Delta z \approx 2.5$ later redshift of reionization. See text for exact model parameters and further discussion.

emissivity.

The SN feedback implementation in GALFORM was introduced in Section 4.2.4. The parameters of Eq. (4.15) chosen for Baugh05 model are $\alpha_{\rm hot} = 2$ and $V_{\rm hot} = 300\,{\rm km/s}$. We compare this default selection with two models with different SN feedback parameters: the first, "weak" feedback model has $\alpha_{\rm hot} = 1$ and $V_{\rm hot} = 100\,{\rm km/s}$, while the second, "strong" feedback has $\alpha_{\rm hot} = 3$ and $V_{\rm hot} = 500\,{\rm km/s}$. The weak feedback model gives

almost identical results to no feedback at all (set with $\alpha_{\rm hot} = 3$ and $V_{\rm hot} = 1\,{\rm km/s}$) and the strong feedback parameters are close to the values in the Bower06 model ($\alpha_{\rm hot} = 3.2$ and $V_{\rm hot} = 475\,{\rm km/s}$). Figure 4.7 show the total number of LC photons produced up to a given redshift. By $z = 5$, the weak feedback model produces ≈ 2 times as many LC photons as the default Baugh05 model, while the strong feedback model produces ≈ 5 times fewer.

To illustrate the effect of the feedback parameters, we show the fraction of ionizing photons produced by halos with circular velocity smaller than some value of $V_{\rm disc}$ for different supernova feedback parameters in Fig. 4.8. This allows us to quantify the importance of SN feedback in each model, by finding the fraction of photons produced by galaxies with $V_{\rm disc} < V_{\rm hot}$, for which more cold gas is ejected from the galaxy by SN-driven winds than is added to the stellar mass by star formation. For the default Baugh05 model at $z = 10$ (black solid line), $\approx 50\%$ of LC photon production is occurring in galaxies with $V_{\rm disc} < V_{\rm hot}$, so may be significantly affected by SN feedback. For the weak feedback model (red solid line), $\approx 15\%$ of the total emissivity is produced in $V_{\rm disc} < V_{\rm hot}$ galaxies, while the strong feedback model (blue line) produces $\approx 70\%$ of LC photons in galaxies in the same range. Value of $V_{\rm hot}$ for each model is marked with a thin horizontal line.

Why does the strong feedback model shows a more significant effect on LC photon production than the weak feedback model in Fig. 4.7? First, note that despite the symmetric choice of parameters for the model variations, relative difference between the two models and the default model are not equal. For example, evaluating Eq. (4.15) at $V_{\rm disc} = 300$ km/s (the average $V_{\rm hot}$) for parameters for each model, we see that the strong feedback model ejects ≈ 4.5 times more cold gas than the default at that $V_{\rm disc}$, while the weak feedback model ejects ≈ 3 times less. These ratios increase with decreasing $V_{\rm disc}$. More importantly, Fig. 4.7 shows that the production of LC photons in the strong feedback model diverges slightly from the default with decreasing redshift, i.e. by $z = 14$, strong feedback model has produced ≈ 2 times fewer photons than the default, while by $z = 5$ it has produced ≈ 5 times fewer photons. This is not seen in the weak feedback model, which produces ≈ 2 times more LC photons than the default at all redshifts. This feature may be a consequence of the hierarchical buildup of galaxies. As seen in Fig. 4.8, the strong feedback model affects more massive galaxies than the other two models. The

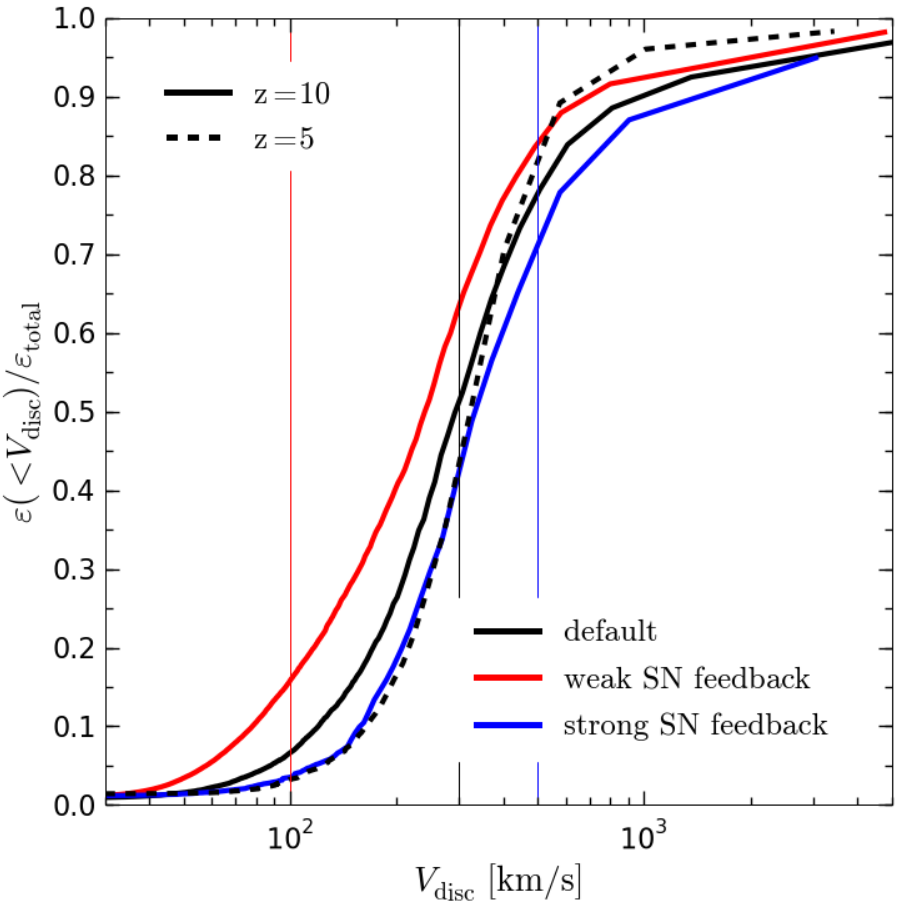

Figure 4.8: Fraction of total emissivity produced by galaxies with circular velocities at half-mass radius smaller than $V_{\rm disc}$. The figure demonstrates the effects of supernova feedback parameters (solid lines) on the relative importance of galaxy masses to the LC emissivity. Thin horizontal lines mark $V_{\rm hot}$ in each model. The growth of structure slightly reduces the relative importance of small galaxies (dashed and solid black lines, representing the default model at $z = 5$ and $z = 10$ respectively). Default model shown at $z = 5$ and $z = 10$, but variant models only shown at $z = 10$. See text for further discussion.

number of massive galaxies increases strongly with decreasing redshift (which may be seen in the evolution of the luminosity functions we present in Section 4.4.2), causing the production of LC photons to diverge in one model that affects the highest galaxy mass

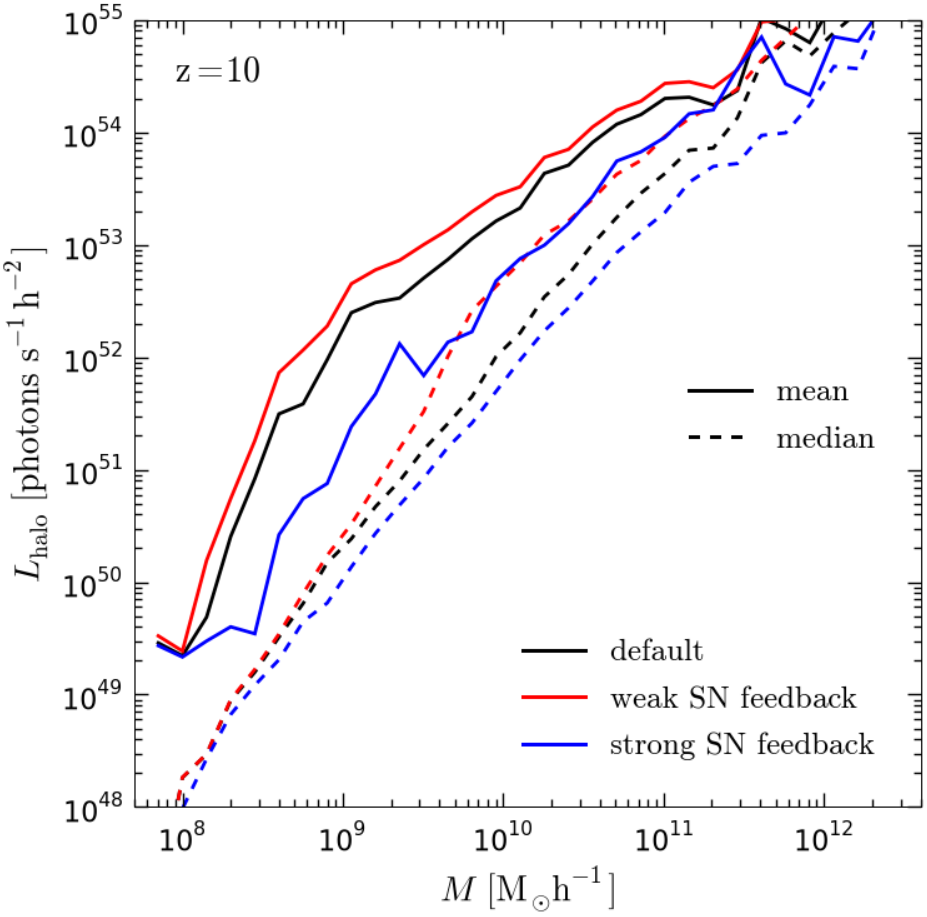

Figure 4.9: LC luminosity as a function of halo mass with respect to SN feedback parameters. Line styles are the same as in Fig. 4.7. See text for discussion.

range. On the other hand, the weak feedback model significantly affects only low mass galaxies which are the first ones to form in the hierarchical picture and whose number density does not evolve as strongly as for the more massive galaxies, which results in a roughly constant effect on LC emissivity. The effect of structure growth on the relevant galaxy mass for LC photon production can also be seen in Fig. 4.8, which shows the relative contribution of different $V_{\rm disc}$ galaxies at $z = 5$ and $z = 10$ in the default model (dashed and solid black lines, respectively). The relative importance of small galaxies to the LC emissivity decreases slightly: at $z = 5$, galaxies with $V_{\rm disc} < V_{\rm hot}$ produce $\approx 40\%$ of all LC photons, as opposed to $\approx 50\%$ the same $V_{\rm disc}$ galaxies produce at $z = 10$.

This is further elaborated in Fig. 4.9, where we present the dependence of halo LC luminosity on its mass for different SN feedback parameters. The decrease in SN feedback efficiency increases the median luminosity, which reflects the behaviour in non-bursting galaxies, with $M \gtrsim 10^9 \mathrm{M}_\odot/h$ (median, dashed red line). For less massive halos, the distribution remains largely unaffected since SN feedback is effective in small galaxies even in the weak feedback model (leading to a smaller difference between median values of weak feedback and default models, red and black dashed lines respectively, in that mass range). The strong feedback model suppresses the median luminosity in halos of all masses more strongly (blue dashed line). More importantly, the mean luminosity (blue solid line) is more severely affected by the increase in feedback efficiency than the median, most likely due to less cold gas being available for starbursts (as strong feedback will eject more cold gas during the quiescent star formation mode) combined with the strong feedback in bursts themselves. Note that the SN feedback efficiency parameters are assumed to be the same in both burst and quiescent star formation in the Baugh05 model, in spite of the change in the IMF between the two modes.

We have shown that the SN feedback can be a significant process for regulating ionizing photon emissivity, though not as strong as the boost in photon production due to the top-heavy IMF discussed in Section 4.3.3. The SN feedback in Baugh05 is not a dominant effect on the production of LC photons, as evidenced by the fact that further decrease of SN feedback strength does not result in a large shift in reionization redshift, as seen in Fig. 4.7. The increase of efficiency, on the other hand, may have a much more severe impact. The difference in SN feedback parameters can therefore explain a significant fraction of the difference in the emissivities of Baugh05 and Bower06 models seen in Fig. 4.1, which is not fully explained by just the different IMF in bursts. The parameters chosen by Bower06 yield a slightly stronger feedback than the strong feedback model presented in this section.

4.3.5 Star formation timescale

We have discussed the star formation timescales when comparing the two Durham default GALFORM models in Section 4.2.8. The choice directly affects the star formation rate, defined in Eq. (4.13). The Baugh05 model defines τ_* with Eq. (4.22), where it depends only on the V_{disc} of the galaxy, while the Bower06 model uses Eq. (4.24), which

also includes the dependence on the dynamical time of the galactic disc, $\tau_{\rm dyn}$. As discussed in Baugh et al. (2005), the use of Eq. (4.22) is motivated by the fact that it gives much longer τ_* than Eq. (4.24) at high redshifts, thus leaving more cold gas in the disc available for star formation in the burst mode. In this section we show the effect the choice of τ_* has on the ionizing photon emissivity.

Figure 4.10 shows the star formation time scale as a function of $V_{\rm disc}$ using different prescriptions for τ_*. The default Baugh05 model (black points) computes τ_* using Eq. (4.22) with τ_{*0} =8 Gyr and $\alpha_* = -3$. The default model is compared with two variants that also employ Eq. (4.22): the shorter timescale has τ_{*0} =4 Gyr, $\alpha_* = -2$ and the longer timescale has τ_{*0} =16 Gyr, $\alpha_* = -4$. (blue and green points, respectively). The final model replaces the default τ_* relation of Baugh05 with the one from Bower06, Eq. (4.24), with the same parameters as in the Bower06 model: $\eta_* = 0.002857$ and $\alpha_* = -1.5$ (red points). The scatter seen in the model that uses Bower06 τ_* is due to the added dependence on $r_{\rm disc}$ in Eq. (4.24). As seen in Fig. 4.10, the Baugh05 shorter and longer timescale variants result in only show a difference in τ_* for galaxies with $V_{\rm disc} \lesssim 200\,{\rm km/s}$ when compared to the default model. As shown in Fig. 4.8, these galaxies are responsible for only $\approx 25\%$ of total LC photons produced at $z = 10$, so the LC emissivity should not be significantly affected by this change, as we will demonstrate shortly. On the other hand, the use of the Bower06 prescription for τ_*, results in a much more significant departure from the default Baugh05 model: τ_* is on average 3 orders of magnitude shorted for all values of $V_{\rm disc}$. This difference can have a much more significant effect on the LC emissivity.

Figure 4.11 shows that our preliminary expectations were correct. Variation of parameters in Eq. (4.22) has a very small effect on the production of ionizing photons when compared to the default values. On the other hand, the use of the Bower06 relation for τ_* results in two times fewer photons produced by $z = 5$ than in the default model, and a delay in reionization redshift is $\Delta z \approx 1$.

Further insights can be obtained from panels in Fig. 4.12, which show the LC emissivity and luminosity with respect to halo mass. Again, the variation of Eq. (4.22) has a minimal effect on emissivity (left panel). Still, the right panel shows that the consequences are not negligible: the median luminosity of individual halos is significantly affected by the change in SF timescale for halos with $M \lesssim 10^{10}{\rm M}_\odot/h$. The mean, on

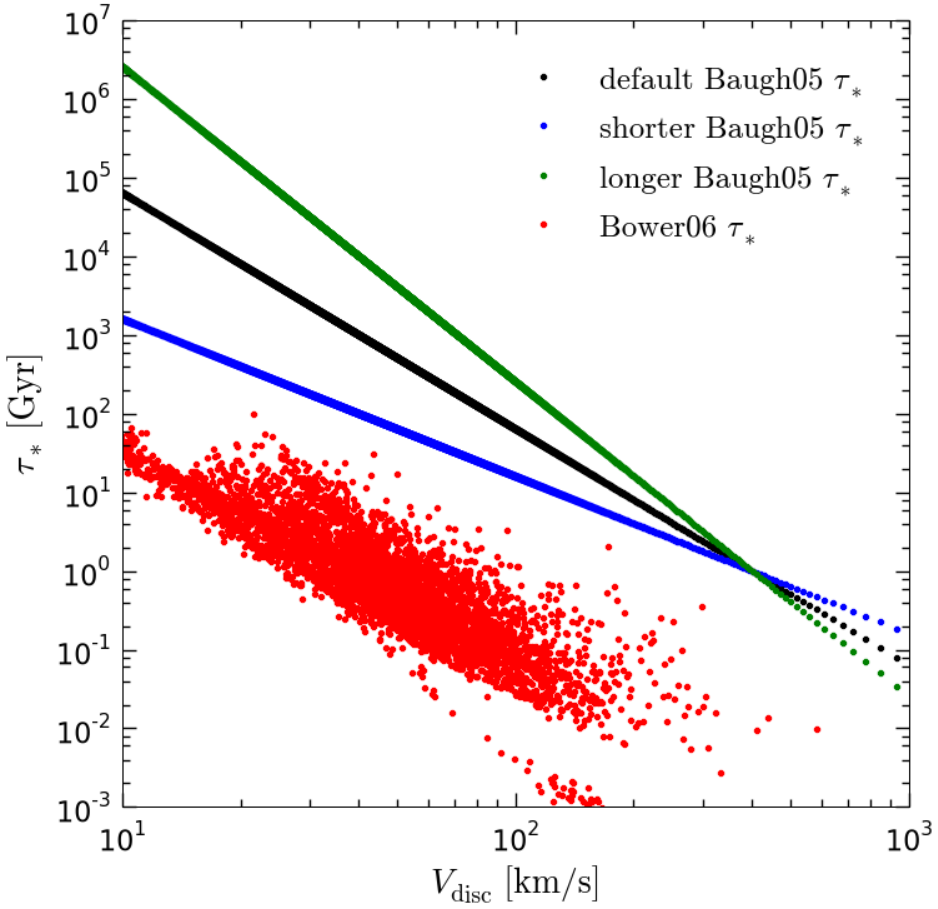

Figure 4.10: The star formation timescale as function of disc circular velocity at half-mass radius for different definitions. The values plotted use definitions given by Eq. (4.22) (black, blue and green points) and Eq. (4.24) (red points), while the values of $V_{\rm disc}$ and $r_{\rm disc}$ are taken from galaxies at $z = 10$ in the default Baugh05 model. The scatter in Bower06 τ_* results is due to the dependence on $r_{\rm disc}$. The SF timescale at this redshift is, on average, ≈ 3 orders of magnitude shorter with the Bower06 model prescription than in the default Baugh05 model. See text for exact model parameters and further discussion.

the other hand, remains unaffected which explains the unchanged emissivities. This last effect is not surprising when one remembers that a different SF timescale is assumed during bursts that are responsible for the mean luminosity, Eq. (4.23), and which is com-

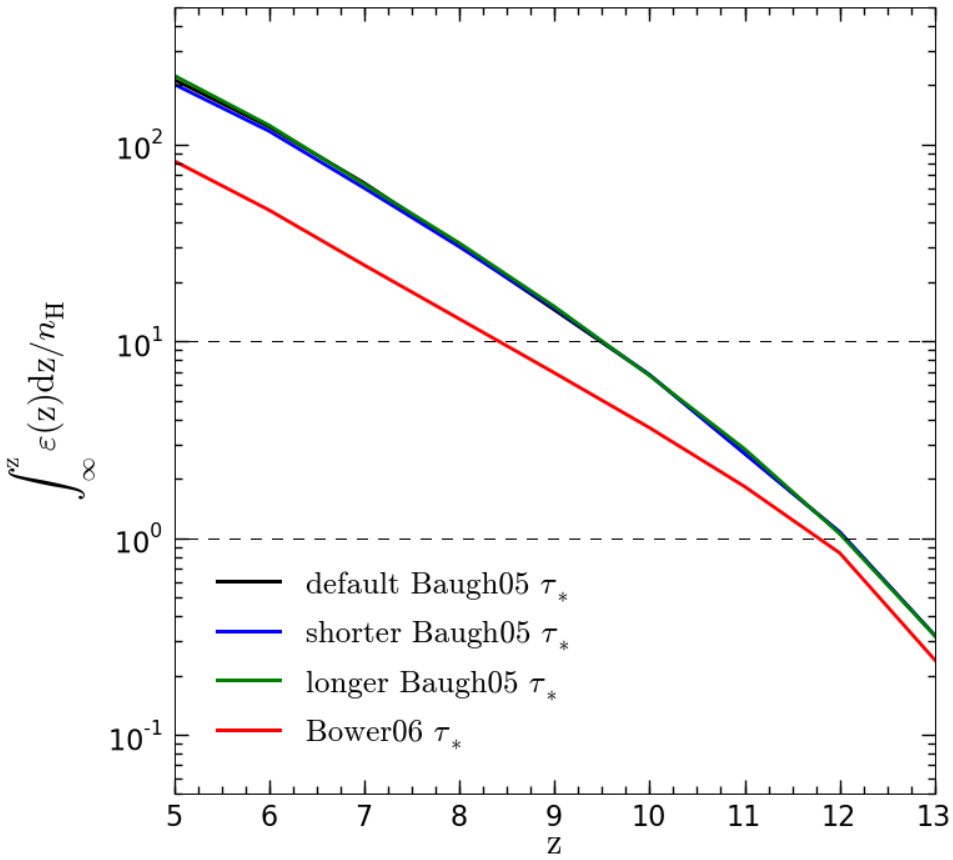

Figure 4.11: Number of LC photons per H atom produced up to redshift z with respect to different star formation timescales. The change of parameters of Eq. (4.22) (blue and green lines) results in a negligible differences from the default Baugh05 model (black line). On the other hand, the use of Eq. (4.24) which introduces the dependence on the disc dynamical time, results in $\Delta z \approx 1$ delay in reionization redshift (chosen here to be achieved when 10 photons are produced per H atom).

pletely unrelated to the SF timescale in discs we are varying here. We tested the effect of the timescale in bursts by varying $f_{\rm dyn}$ in Eq. (4.23) by a factor of two[5] and found

[5]The relatively small factor we varied the timescale by is justified due to the burst timescale being already significantly shorter (of order ~ 10 Myr) than the quiescent one. Note that stronger variations of the timescale may result in a more significant effect on LC emissivity.

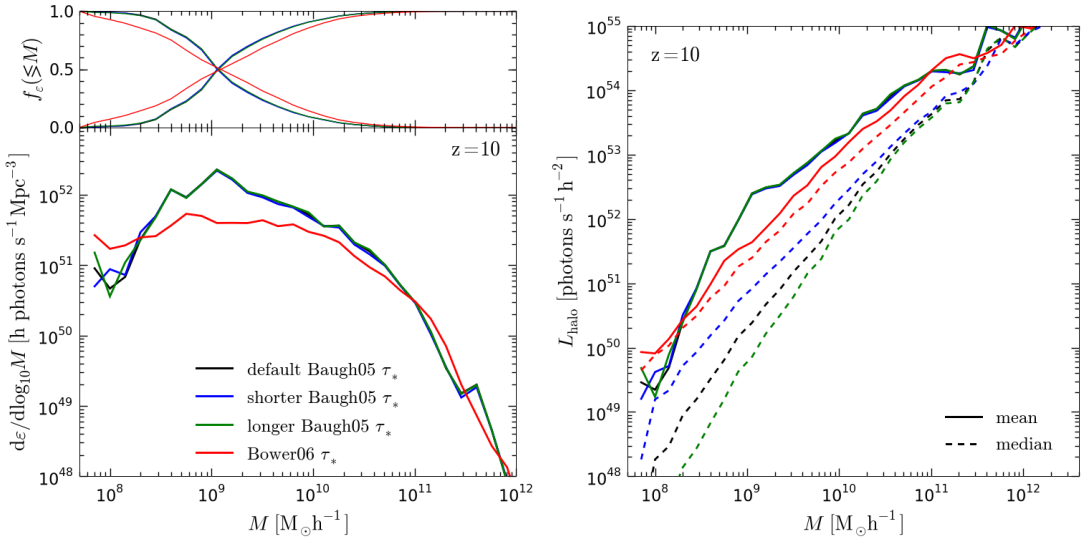

Figure 4.12: The LC emissivity per halo mass (left) and LC luminosity of individual halos as a function of their mass (right) with respect to different SF timescale prescriptions. See text for discussion.

a negligible difference in the LC emissivity. The quiescent star formation affects star formation in bursts only by depleting the cold gas reservoir between burst events, as we mentioned earlier. Therefore, a more severe decrease of the quiescent star formation time scale should affect the LC luminosity during bursts. This is exactly what happens when we use the Bower06 SF timescale prescription (red lines in Fig. 4.12). The Bower06 SF timescale model further increases the median halo LC luminosities (red dashed line in the right panel of Fig. 4.12) over almost the whole range of halo masses. At the same time, the mean LC luminosity is decreased (though not as severely as the increase in the median), resulting in a reduction of the LC emissivity over the same mass range (red line in the left panel). Note that the emissivity is higher in the lowest mass halos ($M \lesssim 2 \times 10^8 \mathrm{M}_\odot/h$) in the Bower06 τ_* model than in the others. This is due to the negligible contribution of SF bursts in those halos caused by the lack of smaller mass progenitors, as discussed in Section 4.3.3. In those halos, the emissivity is dominated by the quiescent star formation, so the shorter SF timescale actually boosts the production of LC photons.

Finally, Fig. 4.13 explicitly shows how the switch to SF timescale of Eq. (4.24) affects star formation in bursts. In the default model (black lines), $\approx 50\%$ of stars are produced

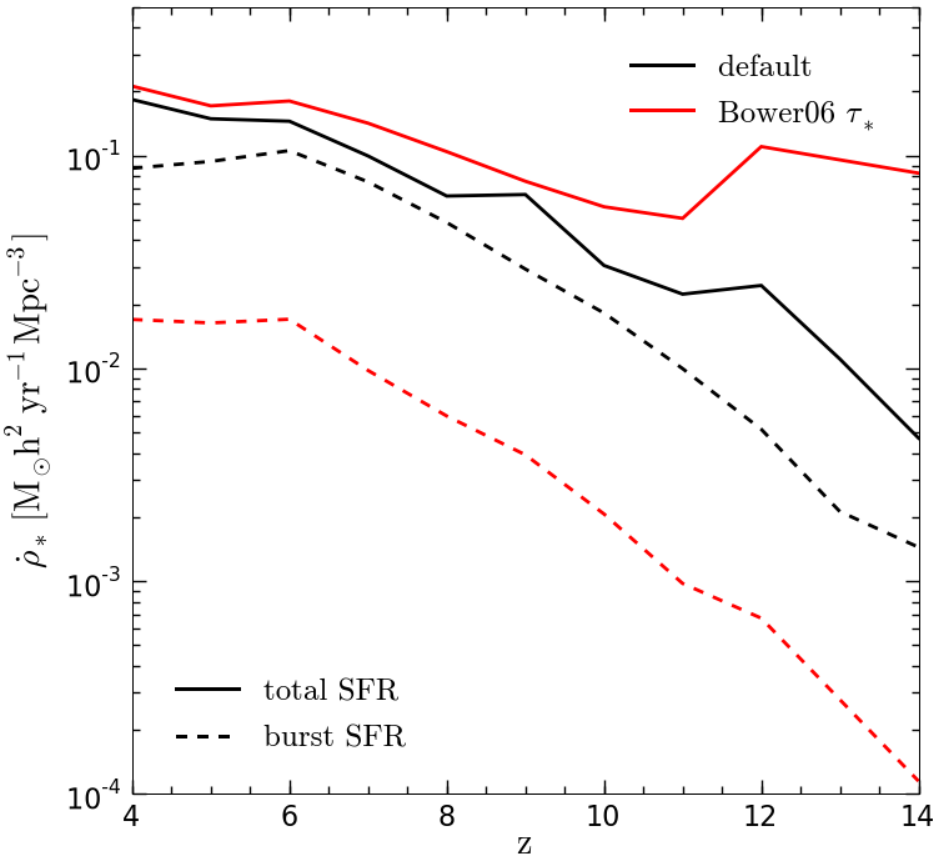

Figure 4.13: Star formation rate density as a function of redshift in models with two different SF timescale prescription. The total star formation rate in the Bower06 τ_* model (red solid line) is somewhat higher than in the default model (black solid line), especially at higher redshifts. A much larger difference is seen in the star formation rate in bursts (dashed lines). This is due to the shorter SF timescale increasing the efficiency of quiescent star formation, reducing the cold gas reserve available for bursts.

in the burst mode at redshifts $4 \lesssim z \lesssim 10$. In contrast, the Bower06 τ_* model forms no more than $\approx 10\%$ of stars in bursts over the same redshift range. Therefore, even though the total SF rate density is slightly higher in the Bower06 τ_* model, the reduced efficiency of star formation in bursts decreases the total LC emissivity produced.

The main conclusion of this section is that a significant increase of the star formation rate in the quiescent mode of the Baugh05 model may lead to a decrease of the total LC photon emissivity. This may seem counterintuitive only if one forgets that the model employs a top-heavy IMF for star formation in bursts. As a final note, we point out that the effects discussed here are caused by the change in the- timescale alone, i.e. the dependence on disc dynamical time in Bower06 τ_* prescription does not play a role. The same effects can be achieved with a Baugh05 τ_* chosen to match the values of Bower06 τ_*.

4.3.6 Disc instabilities

The question of dynamical stability of galactic discs has been the subject of intense study from the early 1970s (e.g. Toomre, 1964; Ostriker & Peebles, 1973; Toomre, 1977, 1981; Efstathiou et al., 1982; Binney & Tremaine, 1987; Christodoulou et al., 1995). In short, when a disc becomes sufficiently massive for its self-gravity to dominate the gravitational potential of the combined galaxy bulge and halo, it becomes unstable to small perturbations introduced by e.g. small satellite mergers or dynamical friction with dark matter substructures. This instability leads to the formation of a bar in the galactic disc and, eventually, to the conversion of the disc into a spheroid. The criterion for the stability of the disc introduced to GALFORM in Cole et al. (2000) is based on the work of Efstathiou, Lake & Negroponte (1982) who used numerical simulations to study the evolution of idealised disc galaxies with central spheroids. As a simple first-order estimate of the disc stability, they introduced the following relation:

$$\epsilon \equiv \frac{V_{\rm max}}{(GM_{\rm disc}/r_{\rm exp})^{1/2}}. \tag{4.26}$$

In the GALFORM scheme, $V_{\rm max}$ is the circular velocity at the disc half-mass radius, $r_{\rm disc}$, $M_{\rm disc}$ is the mass of the disc alone and $r_{\rm exp} = 1.68 r_{\rm disc}$ is the disc exponential scalelength. Note that in the original implementation, $V_{\rm max}$ is assumed to be maximum circular velocity on the rotation curve. In essence, Eq. (4.26) is the square root of the ratio of mass of the whole galaxy (with a significant contribution from the central spheroid) to that of the disc alone within the disc half-mass radius. The choice of ϵ value needed for disc stability, $\epsilon_{\rm stable}$, in GALFORM models is informed by the result of Efstathiou et al. (1982), who found that the discs are stable when $\epsilon \gtrsim 1.1$. At each time step, Eq. (4.26) is evaluated

for every galaxy. If $\epsilon \leq \epsilon_{\rm stable}$ is found, the model assumes that the stellar disc evolves into a bar and then to a spheroid with all the stars from the disc becoming part of the bulge (Combes et al., 1990; Combes, 2000). Note that more recent simulations, show that this may not be the best approximation (e.g. Agertz et al., 2009). From the LC emissivity perspective, the most important point is that the disc instability is assumed to trigger a burst of star formation with the same prescription as the ones triggered by mergers. Disc instabilities are not a standard component of Cole et al. (2000) model, nor of the Baugh05 model. We were motivated to consider them here, because the Bower06 model finds the disc instabilities to be crucial for the growth rate of black holes that drive the AGN feedback at early redshift (see Fig. 2 in Bower et al., 2006)), which we will briefly discuss in the following section.

Figure 4.14 shows the effect of the introduction of disc instabilities into the default Baugh05 model. The choice of $\epsilon_{\rm stable} = 0.8$ comes from the Bower06 model and can be considered a somewhat modest instability criterion (taking into account that, from the definition of ϵ in Eq. (4.26), the minimum possible value of ϵ is $\epsilon \approx 0.62$). In the second model, we double $\epsilon_{\rm stable}$ to 1.6 in order to introduce a more extreme instability criterion (we increased the criterion up to $\epsilon_{\rm stable} = 3$ without causing further significant changes to the emissivity). The increase in the emissivity in the whole halo mass range is entirely due to the more common triggering of SF bursts i.e. a larger fraction of star formation occurs with a top-heavy IMF. In our redshift range of interest ($6 \lesssim z \lesssim 15$), the instability-triggered bursts are much more common than the merger induced ones. Indeed, if we switch off the merger triggered bursts in the models with disc instabilities (in the same way described in Section 4.3.3) there is practically no difference in the LC emissivity relative to the model that includes both merger and disc instability caused bursts. As seen in the right panel of Fig. 4.14, the strong criterion for instabilities results in burst SF mode absolutely dominating in all halos ($> 99\%$ of star formation is in the burst mode at all redshifts, evidenced by the fact that median is almost equal to the mean). On the other hand, when using the weaker criterion, the instabilities are somewhat less common in the least massive halos, $M \lesssim 5 \times 10^9 {\rm M}_\odot/h$. The use of a stronger instability criterion shifts the characteristic halo mass for LC emissivity (median emissivity mass, top panel in the left plot), to a slightly higher mass than seen in the default or weaker instability models.

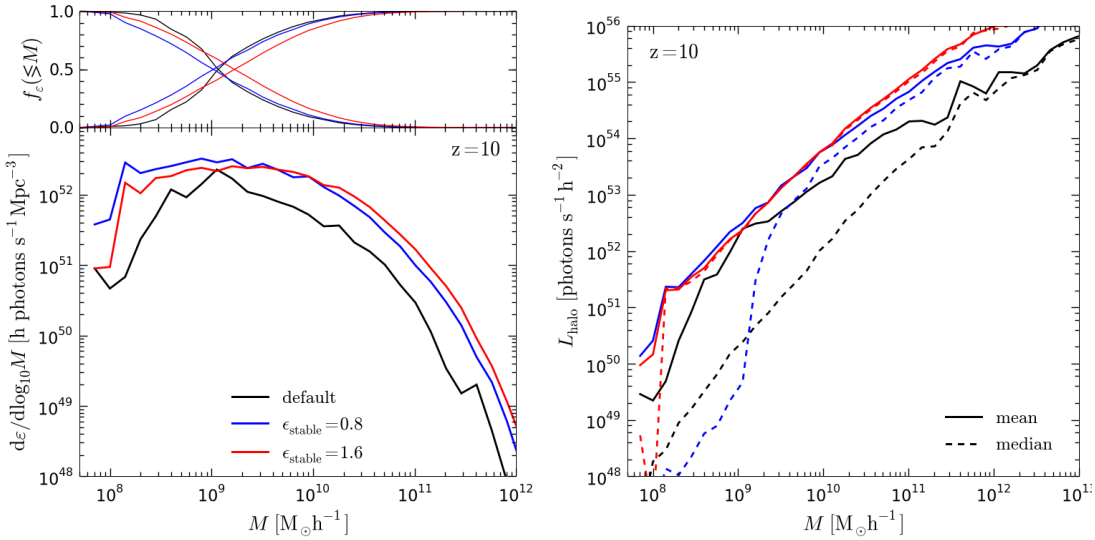

Figure 4.14: LC emissivity (left) and LC luminosity (right) as a function of halo mass with the inclusion of disc instabilities to the Baugh05 model. The increase in emissivity in all halo masses is due to the triggering of SF bursts by the disc instability prescription, resulting in even more star formation in the burst mode than in the default Baugh05 model. For the strong instability model (red lines), the instabilities are triggered efficiently in all halo masses (compare mean and median in right panel, red dashed and solid lines). In the more modest instability model (blue lines), discs are somewhat more stable in low mass halos (see blue dashed line for $M \lesssim 5 \times 10^9 \mathrm{M}_\odot/h$).

Figure 4.15 shows the increase in the photon production due to the addition of disc instabilities to the model. The two models with instabilities produce almost the same number of photons by $z = 5$, but the weaker instability criterion model has a somewhat larger emissivity at early redshifts. This is due to the slightly stronger contribution of smallest halos in the weak instability model, seen in Fig. 4.14. As massive halos become more important at lower redshifts, the gap between the two instability models closes. The difference between these models and the default one also decreases with decreasing redshift: by $z = 13$, the weak instability model produced ≈ 7 times as many LC photons as the default model, but by $z = 5$ it produced only ≈ 2 times as many.

The inclusion of disc instabilities in the Baugh05 model results in a significant boost

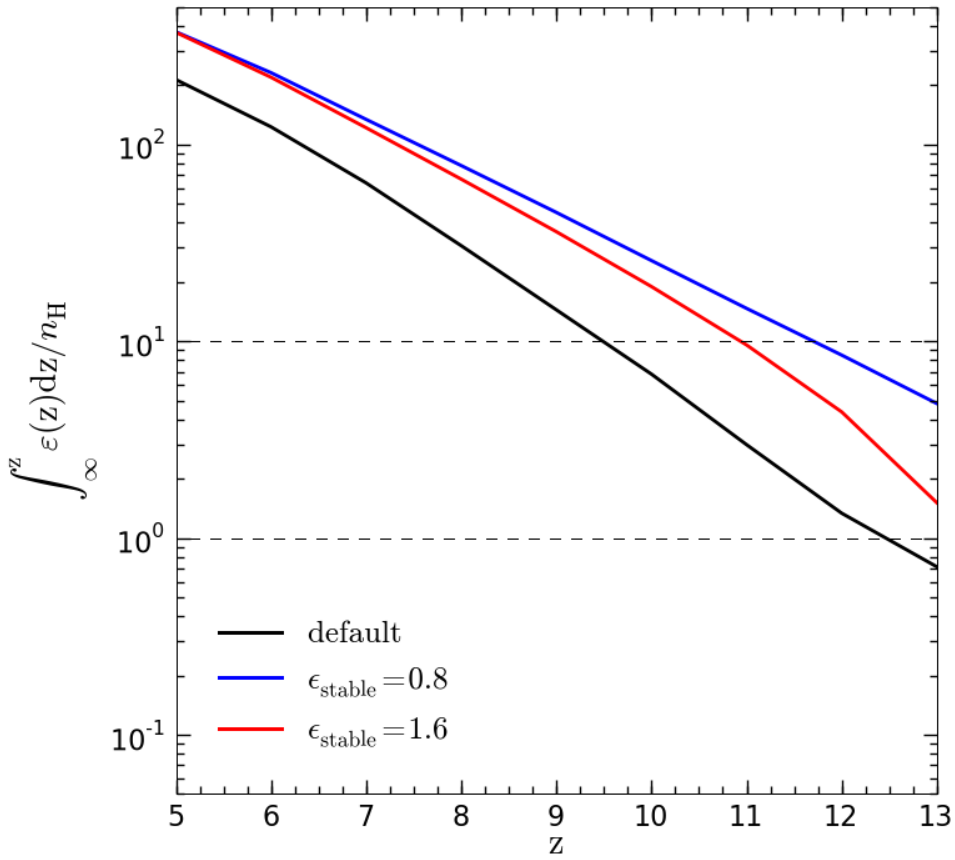

Figure 4.15: Number of LC photons per H atom produced by redshift z in Baugh05 models with disc instabilities. The models with disc instabilities produce ≈ 3 times as many photons as the default models by $z = 10$. This emissivity boost results in a $\Delta z \approx 1.5$ earlier redshift of reionization. The weaker instability model has a slightly higher emissivity at higher redshifts, possibly due to more cold gas being available for bursts, as seen in Fig. 4.14.

to the LC emissivity, especially at higher redshifts. This is the result of a more common triggering of SF bursts, during which the top-heavy IMF is assumed, which is the most important component of the model in terms of the ionizing emissivity as we discussed in Section 4.3.3. The inclusion of disc instabilities to the model also has important consequences beyond those we discussed here. Most crucial are certainly the effect on the

distribution of disc sizes and morphological types of galaxies, due to the conversion of discs to spheroids. Cole et al. (2000) found that the inclusion of disc instabilities in their default model results in a slightly better agreement of disc scale-lengths with observations (see Fig. 8 of their paper). They also showed that the disc instabilities significantly shift the distribution of galaxy morphological types at lower luminosities towards elliptical galaxies. For more details, see their original paper as well as a recent study by Benson & Devereux (2009).

4.3.7 Sub-dominant model parameters

We also tested the effect of a wide range of other physical processes modelled in GALFORM on the ionizing emissivity. Here we briefly mention some of the ones we found not to have a significant effect on the production of LC photons during reionization.

AGN and superwind feedback

The most recent updates to the semi-analytic models have focused on reproducing the observed break in the bright end of the luminosity function (the shape of the faint end can largely be explained by the supernova feedback discussed in Section 4.3.4). At the present day, the cooling function drops quickly for halos roughly more massive than the hosts of Milky Way-size galaxies[6] (e.g. Sutherland & Dopita, 1993) leading to the much longer cooling times and introducing the break in the luminosity function. Still, this cooling argument is not sufficient to explain in detail the position of the observed bright-end break therefore suggesting the existence of additional physical processes.

For this purpose, the Baugh05 model employs a model for supernova-driven superwinds that remove the cold gas from the galactic disc completely out of the halo (regular supernova feedback from Section 4.3.4 only returns it to the hot halo gas reservoir), previously implemented in GALFORM by Benson et al. (2003). See Baugh (2006) or Lacey et al. (2008) for the exact implementation. The Bower06 model uses the previously mentioned AGN feedback for the same purpose. In this implementation, gas cooling is switched off in halos that are quasi-hydrostatically cooling (i.e. halos in which the gas cooling time is longer than the free-fall time) if the energy lost to cooling is smaller than

[6]The mass of the Milky way halo is $M_{\rm vir} \approx 2 - 3 \times 10^{12} {\rm M}_{\odot}$ (Xue et al., 2008), with the corresponding virial temperature $T_{\rm vir} \approx 10^{6} {\rm K}$.

some fraction (a model parameter) of the Eddington luminosity of the central black hole. See Bower et al. (2006) and references therein for a more detailed discussion about AGN feedback and this particular implementation.

We tested the importance of both feedback prescriptions on the production of LC photons during reionization, $z \gtrsim 6$. We found both effects to have a negligible effect on the LC emissivity, which is expected as these processes generally affect more massive halos that form at significantly lower redshifts. The AGN feedback, particularly, makes no difference whatsoever when added to the Baugh05 model. Bower et al. (2006) found that, for the AGN feedback to be effective at all, the disc instabilities must be included to provide a channel for mass-buildup of black holes. Even with the disc instabilities, which we discussed in Section 4.3.6, the AGN feedback effect on LC emissivity is non-existent at our redshifts of interest. On the other hand, the superwind feedback makes a measurable difference, albeit a very modest one: a Baugh05 variant without superwinds begins to produce slightly more LC photons after $z \approx 8$, which results in $\approx 70\%$ more LC photons produced by $z = 5$. The redshift of reionization (10 LC photons produced per H atom) is practically unaffected since it is reached before $z \approx 8$ (see the default Baugh05 model result in e.g. Fig. 4.1). We therefore conclude that these GALFORM parameters are unimportant for the production of LC photons during reionization. Note however that the ionizing radiation from AGN may provide a contribution to the total LC emissivity. We will explore the GALFORM predictions for quasar production of LC photons in reionization context in future work.

Metallicity

The chemical composition of the gas can affect the ionizing luminosities for two reasons (e.g. Baugh, 2006): first, the rate at which the gas cools is a function of metallicity (Sutherland & Dopita, 1993) and, second, the stellar populations formed out of more metal-enriched gas will be redder and somewhat less luminous in UV (see Fig. 4 in Cole et al., 2000). Also, the dust extinction increases with ISM metallicity. In GALFORM, the metal abundance is given as a fraction of total mass of a galactic component that is made out of elements heavier than helium, $Z = M_{\rm Z}/M$. We tested the importance of Z on the LC emissivity with two Baugh05 variant models: one in which we assumed $Z = 0$ when evaluating the cooling rate and another in which no metals are produced (yield

$p = 0$). The former has practically no effect on the LC emissivity at $z \geq 5$. This is because the mean metallicity of the hot halo gas (which is the component for which the cooling is computed) remains close to primordial at these high redshifts e.g. $\langle Z/Z_{\odot} \rangle \approx 10^{-3}$ at $z \approx 9$. Switching off metal production completely by setting $p = 0$ results in a modest boost in the LC emissivity, producing about 50% more LC photons at all redshifts and having $\Delta z \approx 0.3$ earlier redshift of reionization. As the cooling is not affected by these low halo gas metallicities at high-z, this effect comes from the slightly higher luminosities of $Z = 0$ stellar populations. Still, this observed difference is very modest compared to other previously discussed parameters.

4.3.8 Photo-ionization feedback parameters

Lastly, we discuss the effect of the photo-ionization parameters, described in Section 4.2.7, on the evolution of the LC emissivity. While this type of feedback is clearly a crucial feature in reionization modelling, at this stage of the discussion, where we focus only on the GALFORM predicted emissivities prior to the end of reionization, the choice of $z_{\rm cut}$ and $V_{\rm cut}$ parameters is irrelevant. This is simply because the GALFORM implementation we use here assumes instantaneous reionization, which occurs at redshift $z_{\rm cut}$. This means that the photo-ionizing feedback affects galaxy formation only *after* reionization. The two default GALFORM models set similar values of the reionization parameters: $z_{\rm cut} = 6$ in both Baugh05 and Bower06, while $V_{\rm cut} = 60\,{\rm km/s}$ in the former and $V_{\rm cut} = 50\,{\rm km/s}$ in the latter. The value of $z_{\rm cut}$ was motivated by the presumed end of reionization suggested by quasar Gunn-Peterson data (e.g. Fan et al., 2002), while $V_{\rm cut}$ is motivated by the filtering mass implementation of Benson et al. (2002a,b) we mentioned previously. Also, they both find that the exact choice of these parameters does not have a significant impact on predictions for the low redshift observables they consider.

As we will discuss at more length later, we adopt the same implementation of the effect of photo-ionization feedback in our full RT scheme in Chapter 5. The most important difference is that there we do not assume a uniform instantaneous reionization, but compute a reionization redshift for each individual halo using the RT calculation. Before discussing the effect of photo-ionization feedback in such a scenario, it is valuable to briefly address the effects the choice of the two parameters has on the LC emissivity.

Figure 4.16 shows the effect of the choice of $z_{\rm cut}$ (left panel) and $V_{\rm cut}$ (right panel)

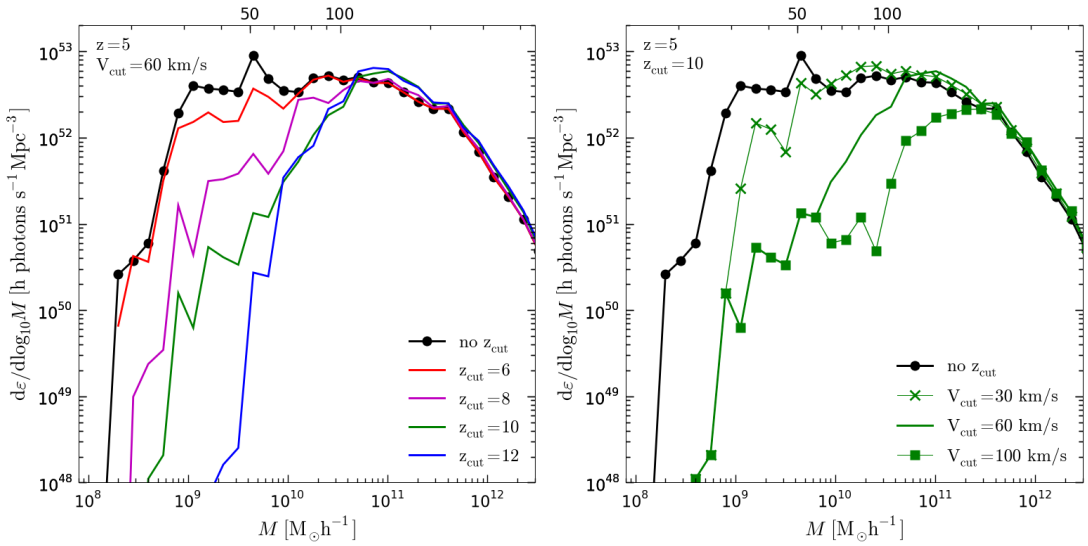

Figure 4.16: Lyman-continuum emissivity per halo mass with varying redshift of reionization, $z_{\rm cut}$ (left panel) and halo circular velocity affected by the photo-ionization feedback (right panel). The scale on top of both plots gives the circular velocity of halos (in $\rm km\,s^{-1}$) at $z = 5$. See text for discussion.

in the default Baugh05 model. The top axis in both panels shows the circular velocity of halos at the redshift[7], $z \approx 5$. The first thing to notice is that the emissivity from halos with $V_{\rm c} > V_{\rm cut}$ is also affected by the photo-ionizing suppression. This is due to the fact that the suppressed gas cooling in small mass halos affects the luminosity in their (more massive) descendants down the hierarchy, both in terms of stellar mass and the frequency of galaxy mergers. The left panel also demonstrates the delayed nature of photo-ionization suppression in the GALFORM prescription: the emission is only slightly affected when $z_{\rm cut} = 6$ (red line), one unit of redshift higher than the plot, $z = 5$. Only when the redshift of reionization is $z_{\rm cut} \gtrsim 12$, is the emissivity from halos with $V_{\rm c} < V_{\rm cut}$ completely suppressed (blue line). This is important to note because the majority of numerical models that take photo-ionization into account when modelling reionization sources assume that the emissivity of suppressible halos goes to zero as soon as they are ionized. The left panel of Fig. 4.16 shows this is not necessarily an accurate assumption, as it takes 3 units of redshift ($\approx$ 550 Myr) to drop the emissivity of all but the smallest

[7]The mass of halo with circular velocity $V_{\rm c}$ decreases with increasing redshift as $M(V_{\rm c}, z) \propto (1+z)^{-3/2}$ in our redshift range of interest.

halos by $z = 5$ (compare black and magenta lines in the affected mass range).

The right panel of Fig. 4.16 shows how the choice of the critical circular velocity, introduced in Section 4.2.7, affects the LC emissivity. Assuming $z_{\rm cut} = 10$, the default value of $V_{\rm cut}$ in the Baugh05 model, $V_{\rm cut} = 60\,{\rm km/s}$ (green solid line), results in a 50% decrease in LC emissivity at $z = 5$ when compared with the model without any photo-ionization feedback (black line with circles). If we assume a value suggested by more recent numerical simulations (e.g. Okamoto et al., 2008) of $V_{\rm cut} = 30\,{\rm km/s}$ (green line with X symbols), the emissivity is decreased by only 10% at the same redshift (and even less at earlier times). This result suggests that photo-heating has only a small effect on the course of reionization, a point which we demonstrate in Chapter 5. Note also that, since all runs in this panel use the same merger trees, the mass at which the results of the three suppressed runs converge ($M \approx 10^9 {\rm M}_\odot/h$) marks the halos which are completely suppressed for all considered values of $V_{\rm crit}$.

Here we discussed the effect of photo-ionization feedback on LC emission in GALFORM alone, meaning that a single redshift of reionization is assumed. We discuss the importance of the photo-ionization feedback process on the progression of inhomogeneous reionization in Chapter 5.

4.4 Observed galaxies at the epoch of reionization

A crucial advantage of using sophisticated semi-analytic galaxy formation models such as GALFORM to predict the sources of ionizing radiation is the ability to compare the predictions of the model against other observational data. Indeed, as we mentioned earlier, almost all GALFORM model parameters are constrained through such comparisons (e.g. Cole et al., 2000; Baugh et al., 2005; Bower et al., 2006). The past few years saw a substantial increase in the samples of known high redshift galaxies, especially in the $z \sim 4 - 6$ range where tens of thousands of galaxies with luminosities $M_{\rm AB} \lesssim -16$ or $\sim 0.01 L^*_{z=3}$ are now known (Bouwens et al., 2008b). Much of this progress is due to the deep multi-wavelength data collected by Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS, Ford et al., 2003) and Subaru's Suprime-Cam (Miyazaki et al., 2002). The installation of the new WCF3[8] camera on the HST in May 2009 allowed for a

[8] http://www.stsci.edu/hst/wfc3

new push into even higher redshifts, with stronger constraints on the luminosity function (LF) at $z \sim 7$ (e.g. Oesch et al., 2009; Ouchi et al., 2009; Wilkins et al., 2009) and first estimates made on the LF up to $z \sim 10$ (Bouwens et al., 2009b). These observations provide, for the first time, a direct look at a galaxy population around the time of cosmic reionization.

In this section we will compare the predictions of the GALFORM models (namely, the default Baugh05 model and some variations discussed in Section 4.3) with the galaxy sample obtained from the Hubble Ultra Deep Field 2009 (HUDF09) fields using the Lyman-break technique. The HUDF09 program (GO11563) aims to create three ultra deep fields using the near-IR filters on the WCF3 camera, one positioned over the original HUDF field (Beckwith et al., 2006) and the other two over the HUDF05 fields (Oesch et al., 2007). The data from three near-IR bands, Y_{105}, J_{125} and H_{160}, in conjunction with the previously obtained optical data from the same fields, allows the selection of $z \sim 7$ z_{850} (Oesch et al., 2009), $z \sim 8$ Y_{105} (Bouwens et al., 2009a) and $z \sim 10$ J_{125} (Bouwens et al., 2009b) dropouts to up to ~ 29 apparent magnitude at 5σ detection. In this work, we focus only on the Lyman-break selected galaxies, ignoring the other promising techniques for identifying high-z galaxies, namely the direct detection of their Lyman-α line emission (e.g. Partridge & Peebles, 1967a,b; Djorgovski et al., 1985; Cowie & Hu, 1998; Hu et al., 1998; Steidel et al., 2000; Stiavelli et al., 2001; Hu et al., 2004; Taniguchi et al., 2005). The properties of these Lyα-emitters have been explored with GALFORM in detail by Delliou et al. (2005, 2006) and Orsi et al. (2008). We focus on Lyman-break selected galaxies because, at this time, this technique has yielded detections at higher redshifts than the Lyman-α technique. GALFORM has previously been used to study the Lyman-break galaxies at lower redshifts than here by various authors: Baugh et al. (1998, 2005); Lacey (2006); Gonzalez et al. (2010, in preparation); Lacey et al. (2010, in preparation)

We structure this section as follows. In Section 4.4.1 we briefly introduce the Lyman-break technique for selecting high-z galaxies. In Section 4.4.2, we compare the results of the GALFORM models with the HUDF09 data and discuss the evolution of the luminosity function and its implications for the ionizing emissivity. We discuss both the LF and the parameters of the Schechter fits through the GALFORM LF predictions.

4.4.1 The Lyman-Break technique

In essence, the Lyman-break technique refers to the selection of actively star-forming galaxies at high redshifts using photometric measurements. The technique and its applications have been reviewed by Giavalisco (2002). For more details see also the first successful uses of the technique by Steidel & Hamilton (1992, 1993) and Steidel et al. (1995), where the authors attempted to detect galaxies in the field of QSOs with known optically thick absorbers (which may be caused by massive galaxies, e.g. Wolfe et al., 1986; Lanzetta et al., 1991; Steidel et al., 1994), or in the randomly chosen fields (Steidel et al., 1996a,b). The Lyman-break technique rests on two assumptions about the spectra of high-z galaxies: (2) there is a pronounced break in the galaxy SED at 912 Å, the Lyman-limit and (2) the spectra are roughly flat ($f_\nu \propto \nu^0$) in the far-UV redward of the Lyman limit, where the flux is dominated by emission from massive stars. The former feature comes from the very efficient absorption of LC ($\lambda \leq 912\mathring{A}$) radiation in stellar atmospheres, as well as in the ISM (Heckman, 2000) and IGM (Steidel et al., 1995, 1999; Madau, 1995). The detection of such galaxies by colour requires a minimum of three broadband filters, chosen so that the redshifted Lyman-break feature falls approximately between the bluer two. This means that the galaxy will be detected in the two redder bands, while being almost invisible in the bluest band: this "disappearance" from the bluest band is responsible for coining the name "dropout galaxies" for the objects selected using this technique. As an example, (Steidel & Hamilton, 1992, 1993) used a set of specially designed filters, $U_nG\mathcal{R}$, to detect star-forming galaxies in $3.0 \lesssim z \lesssim 3.5$ range. At those redshifts, the U_n filter is mostly shortward of the Lyman-break while G is set so that it is not severely blanketed by the Lyman-α forest leading up to $1216\mathring{A}$. Steidel et al. (1995) argued that for a galaxy to satisfy the previously stated criteria, the $U_n - G$ colour should show a strong break ($U_n - G \geq 1.6$), while the colour redward of the break should be blue ($G - \mathcal{R} \leq 1.2$). Finally, an additional criterion is applied to minimise contamination of the sample by intrinsically red galaxies at lower redshifts, $U_n - G \geq 1 + (G - \mathcal{R})$. The sample may also be contaminated by individual stars in the field, which may be distinguished from galaxies by their spectral features. These criteria define a region on the colour-colour diagram which holds all of the prospective galaxies at the expected redshift. The efficiency of this selection method was shown to be very high by subsequent spectroscopic confirmations, with $\approx 85\%$ of U_n-dropout candidates

confirmed as galaxies at the desired redshift (Steidel et al., 1996b, 1999). This makes it an excellent tool for finding large numbers of galaxies at a specific redshift with a relatively small investment in terms of observation time, compared to a complete spectroscopic survey.

The technique remains largely the same when applied to higher redshifts ($z \gtrsim 6$). The main difference is that the Lyman-α forest line blanketing essentially removes all of the flux blueward off the Lyman-α emission line (e.g. Bunker et al., 2009). In this case, the Lyman-α line is used as location of the break instead of the LC break. Also, as we mentioned earlier, for $z \gtrsim 7$ the dropout feature is redshifted beyond the visible range, requiring the use of a near-IR camera such as the WFC3/IR used for the HUDF09 project or the upcoming James Webb Space Telescope[9] (JWST).

4.4.2 Comparing GALFORM results with HST results

The data against which we compare the GALFORM results is compiled from various sources. For $z \sim 4$, 5 and 6 we use the data compiled by Bouwens et al. (2007) from various HST ACS deep fields: HUDF (Beckwith et al., 2006), HUDF NICMOS parallels fields (HUDF-Ps Bouwens et al., 2004; Thompson et al., 2005), HUDF05 (Oesch et al., 2007) and GOODS (Giavalisco et al., 2004). For higher redshifts, we use the previously discussed HUDF09 results at $z \sim 7$ (Oesch et al., 2009), $z \sim 8$ and 10 (Bouwens et al., 2009a). For the details on how the observational data was reduced, we direct the reader to these papers.

Luminosity function of dropouts and GALFORM parameters

We begin the comparison by looking at the luminosity function (LF) at various redshift. In essence, this is the number density of galaxies, $\mathrm{d}n$, with luminosity between L and $L + \mathrm{d}L$. It is a fundamental observable, in a sense that it is mostly[10] independent of assumptions about the theoretical models (Bouwens et al., 2008b). Figure 4.17 compares the observations at redshifts 6, 7, 8 and 10 with the predictions of GALFORM models. The observed galaxy densities were converted from the concordance cosmology (h = 0.7, $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$) to the Millennium cosmology used in the GALFORM models (h =

[9] http://www.jwst.nasa.gov/

[10] A cosmological model must be assumed to interpret the observed distances and volumes.

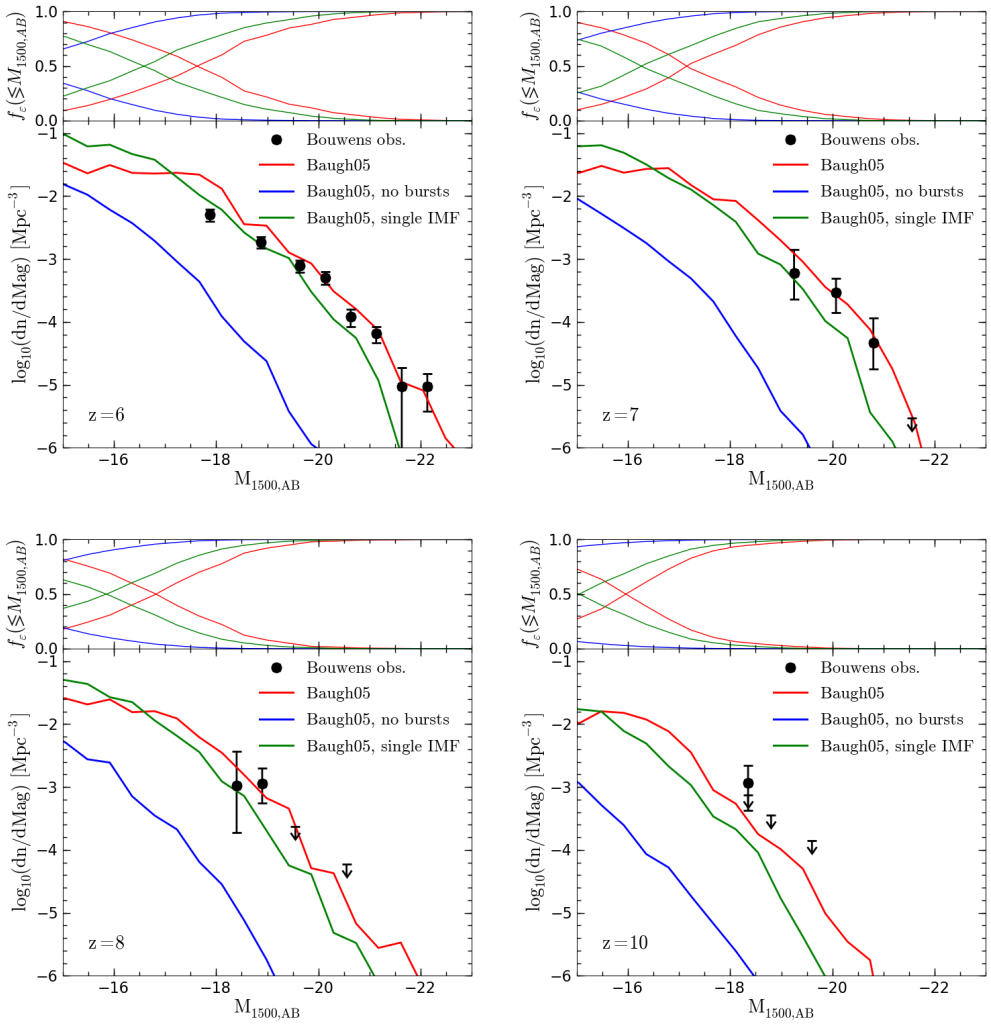

Figure 4.17: UV luminosity functions (1500 Å) at $z \sim$ 6,7,8 and 10 (top to bottom and left to right). The observational data with 1σ error bars are given by black symbols, and are compared against the default Baugh05 (red line), Baugh05 without top-heavy IMF in bursts (green line) and Baugh05 without any bursts (blue line). The downward arrow symbols mark 1σ upper limits on the observed LF. Top panels show the fraction of ionizing photons produced by galaxies with UV luminosities smaller or larger than $M_{1500,\mathrm{AB}}$. The default Baugh05 model shows a good agreement with observational data at all redshifts. See text for discussion.

0.73, $\Omega_0 = 0.25$, $\Omega_\Lambda = 0.75$), resulting in $\approx 5\%$ lower LFs than the ones published in the original papers. The UV luminosities are represented by absolute magnitudes[11] at 1500 Å in the rest frame. The dependence on h was accounted for in both number densities and luminosities of galaxies in each bin (assuming $h = 0.73$). As the plots demonstrate, the default Baugh05 model results are in good agreement with the presently available observations. This model was previously shown to be in good agreement with Lyman-break galaxies at $z \sim 3$ (Fig. 3 in Baugh et al., 2005; Gonzalez et al., 2009) and $z \sim 6$ (Gonzalez et al., 2010, in preparation). However, this is the first time a comparison with $z > 6$ samples has been performed. The UV LF is significantly underestimated in the model without bursts (blue lines), as one might expect from the large difference in the production of LC photons as see in e.g. Fig. 4.4. The difference between the default Baugh05 (red lines) and version without the top-heavy IMF in bursts (green lines) is smaller, especially when one remembers the large difference in the ionizing photon emissivity (Figs. 4.4 and 4.5) found between the two models. The more modest difference seen in the 1500 Å LFs is caused by the stronger dust extinction in the default model, due to more efficient metal production caused by the higher yields in bursts. When the dust extinction is not considered, the difference in galaxy LFs of the two models becomes much more prominent, but then both models significantly exceed the observed LF. With the dust extinction included, the top-heavy IMF model is only slightly better in matching observations than the one without it.

The top panels in Fig. 4.17 show how the production of LC photons shifts to less luminous galaxies with increasing redshift: while the observed range of galaxy luminosities accounts for $\approx 50\%$ of emissivity at $z = 6$ ($M_{1500,\mathrm{AB}} \lesssim -18$), at $z = 10$, the observed galaxies sample only a few percent of the total emissivity. This results emphasises the potential inaccuracy when estimating the LC emissivity from observations such as these alone. We remind the reader that the fractions shown in the top panels correspond to the total production of LC photons in *each model*, but this does not reflect the large differences in emissivities produced between the different models. Note that the Baugh05 model have not been changed in any way to fit this most recent data.

Figure 4.18 compares the default Baugh05 and Bower06 in terms of the predicted Lyman-break galaxy LFs at high-z. At lower redshift ($z = 5$, left panel), the predictions

[11]The AB magnitude system (Oke & Gunn, 1983) is used throughout this thesis.

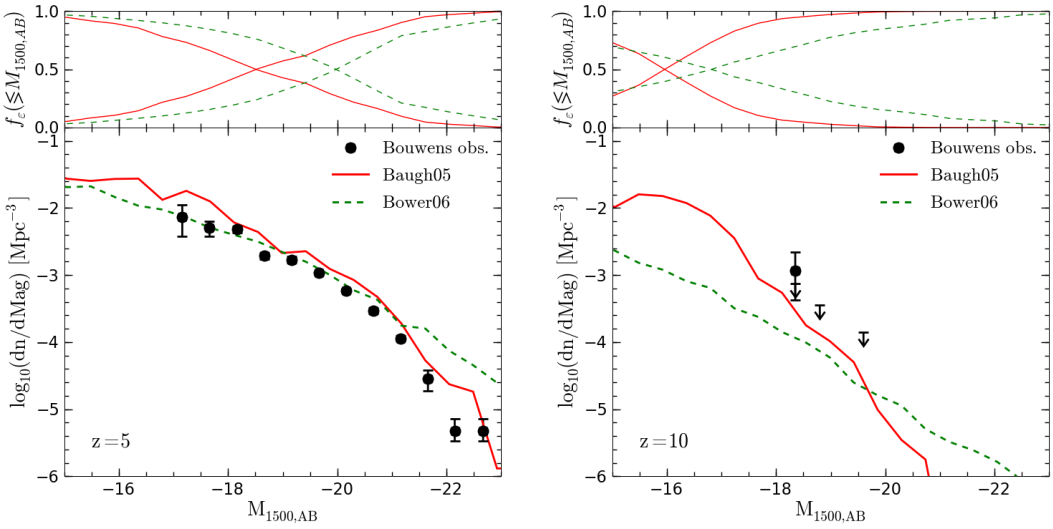

Figure 4.18: UV luminosity functions at $z \sim 5$ (left panel) and 10 (right panels) predicted by default Baugh05 (red solid lines) and Bower06 (green dashed lines) models. Observational data constraints are shown with black symbols (downward arrows give upper limits only). The Bower06 model seems to match slightly better the observed low-luminosity end at $z = 5$, but it over-predicts the high-luminosity end. At $z = 10$, the Bower06 model predicts many fewer low luminosity and many more high luminosity galaxies than Baugh05, as expected by the large difference in LC emissivity between the two models (see Fig. 4.1).

at lower luminosities are fairly similar between the two models, with Bower06 even providing a slightly better fit to the observed LF. However, there is a big difference in the high-luminosity end, where the Bower06 model does not show the exponential drop-off in the LF: the model's predictions lack the drop-off in all $z \geq 5$ predictions, having a simple power-law shape. The observed data features an obvious drop-off, which exposes the inadequacy of the Bower06 model for matching the observations of galaxies at high redshifts. At even higher redshifts ($z = 10$, right panel), the difference in the predicted number of low luminosity galaxies between the two models also becomes clear. Since the low luminosity galaxies are the dominant sources of LC photon production at these redshifts (see the fraction of ionizing photons produced by the Baugh05 model, top panel), this feature certainly contributes to the large difference in LC emissivity produced by

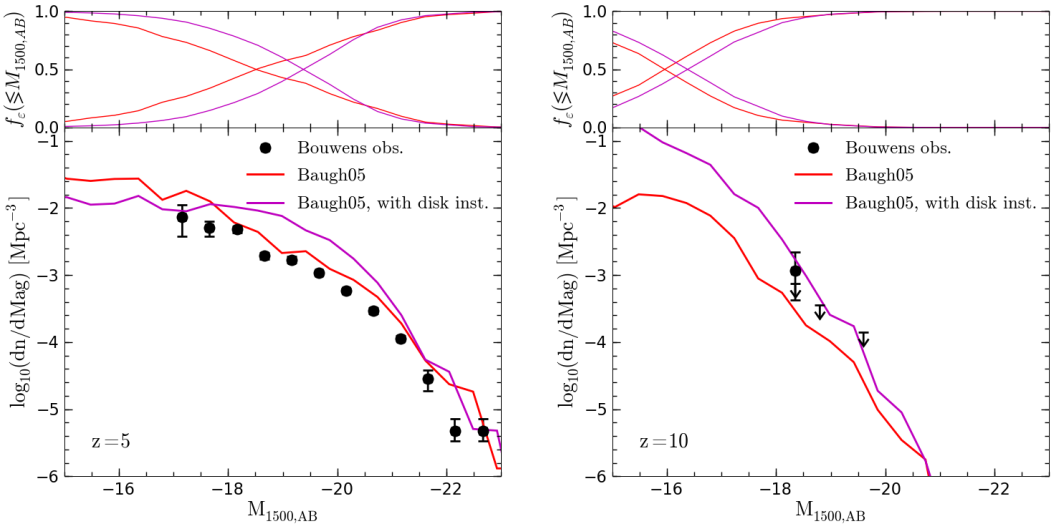

Figure 4.19: UV luminosity functions at $z \sim 5$ (left panel) and 10 (right panel) compared to Baugh05 model with disc instabilities (magenta lines). More common starbursts in this model result in overestimated number densities at $z \sim 5$. The lack of observational constraints on the low-luminosity end at $z \sim 10$ prevents the complete discounting of the disc instability model.

the two models (Fig. 4.1). The better agreement between the two models at $z = 5$ is also matched by a smaller difference in emissivity at that redshift, mentioned in Section 4.2.8, but it is too late for the Bower06 model to produce significant number of ionizing photons to meet the observational constraints on the epoch of reionization.

In Section 4.3.6, we have shown that the disc instability process can be more effective in triggering starbursts than galaxy mergers, resulting in more gas converted to stars with the top-heavy IMF, thus boosting LC emissivities (see Figs. 4.14 and 4.15). Figure 4.19 shows how the inclusion of disc instabilities in the default Baugh05 model affects the predicted UV LF. The disc instability model used here assumes $\epsilon_{\rm stable} = 0.8$. At $z = 10$ (right panel), the same redshift as Fig. 4.14, the predicted galaxy LF is boosted over the whole range of luminosities, especially for the low-luminosity end. The explanation is the same as the one we put forward in the case of the low-mass end of the LC emissivity distribution: disc instabilities trigger bursts in the smaller halos where the galaxy mergers do not occur due to the lack of progenitors of smaller masses. The

model somewhat exceeds the upper limits suggested by the observational data in the observed range ($-18 \lesssim M_{1500,\mathrm{AB}} \lesssim -20$). The current lack of data points at lower luminosities means that the shape of the low-luminosity tail cannot be constrained at this redshift. At lower redshifts, such as $z = 5$ (left panel), the disc instabilities result in a much smaller difference from the default model. For $M_{1500,\mathrm{AB}} \gtrsim -18$, the disc instability model even produces a somewhat lower galaxy LF. This is a feature also mirrored in the emissivity, where we find a significant drop for low mass halos ($M \lesssim 10^9 \mathrm{M}_\odot/h$) at the same redshift. The most likely cause for this behaviour is that, when disc instabilities are present, starbursts are triggered early and often and consume the cold gas pool thus reducing the efficiency of both quiescent star formation and later bursts. More massive halos which cool gas more efficiently remain unaffected and the model with disc instabilities boosts the LF at higher luminosities at $z = 5$ even more efficiently than at $z = 10$. Future attempts to unify the Baugh05 and Bower06 models into a single picture of galaxy formation will most likely have to include disc instabilities since they appear to be crucial for the buildup of black hole mass at early redshifts necessary for efficient AGN feedback, as we mentioned in Section 4.3.6. This comparison against the observed population of high redshift galaxies discourages a simple inclusion of disc instabilities in the Baugh05 model, suggesting that the potential unified model will need to be more than a sum of the parts of the original ones.

Throughout this thesis we have assumed the same ΛCDM cosmology as was used in the Millennium simulation (Springel et al., 2005b). More recent measurements, namely the ones made by the *Wilkinson Microwave Anisotropy Probe*[12] (WMAP) mission, suggest a somewhat different set of parameters for the 5-year dataset. From the perspective of structure formation, the largest difference is introduced by the lower value of σ_8, the present-day mass variance on 8 Mpc/h scale used to normalise the linear power spectrum. The lowering of σ_8 has the effect of delaying structure formation (e.g. McQuinn et al., 2007), which can be crucial for the emissivity during reionization when the source number density is already low. Figure 4.20 compares the UV LF of the Baugh05 model in our default (Millennium) cosmology ($\sigma_8 = 0.9$ and $n_s = 1$; other values quoted earlier) with the best-fit cosmology with 5 years of WMAP data (WMAP5; $\Omega_m = 0.273$, $\Omega_\Lambda = 0.726$, $\Omega_b = 0.0456$, $h = 0.705$, $n_s = 0.96$ and $\sigma_8 = 0.811$; see Hinshaw et al., 2009).

[12]http://map.gsfc.nasa.gov/

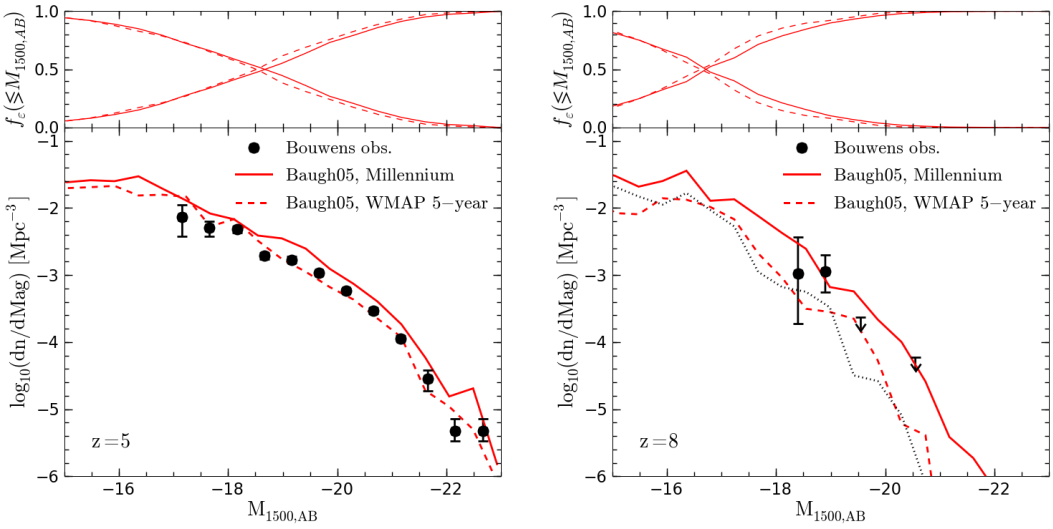

Figure 4.20: UV luminosity functions at $z \sim 5$ (left panel) and 8 (right panel) for the default Baugh05 model, but with two different sets of cosmological parameters. The lower σ_8 in the WMAP 5-year cosmology (red dashed lines) results in the delayed structure formation when compared to the Millennium one (red solid lines). This is demonstrated by comparing WMAP5 result at $z = 8$ (right panel) with the Millennium result at $z = 10$ (dotted black line). Use of the WMAP5 cosmology results in a better agreement with more abundant observations at lower redshifts ($z = 5$, left panel). See text for further discussion.

At $z = 5$ (left panel), where the UV LF is well constrained, the use of WMAP5 cosmology provides a better fit to the observational data. At higher redshifts such as $z = 8$ (right panel) or $z = 10$ (not plotted) the difference with the Millennium cosmology results are larger, but both models fit the sparse observational data. Further improvement of the galaxy sample at high-z will better constrain this comparison. The dotted black line in the right panel gives the UV LF of the Millennium cosmology model at $z = 10$, added to demonstrate the delay in structure formation caused by lower σ_8: the halo mass function in the WMAP5 cosmology is the similar at z as at some earlier redshift z' in Millennium cosmology, when $\sigma_8' > \sigma_8$. The observed difference introduced in the galaxy number densities result in the $\Delta z \approx 1.5$ delay in redshift of reionization between the two models. Note that we have not adjusted any of the original Baugh05 parameters beyond the

cosmology. Standard GALFORM practise requires the model parameters to be adapted to the used cosmology, therefore the discussion about which cosmology is appropriate to be used is not complete.

Schechter function fits through the luminosity functions

Observational studies commonly fit the LF data with Schechter functions, (Schechter, 1976):

$$\Phi(L)\mathrm{d}L = n^* \left(\frac{L}{L^*}\right)^{\alpha} \exp\left(-\frac{L}{L^*}\right) \mathrm{d}\left(\frac{L}{L^*}\right). \tag{4.27}$$

Fitting this function through the available data reduces the properties of the LF to three quantities: the characteristic luminosity L^* which marks the transition from a power law representing low luminosities to an exponential for the high luminosities, slope α of the power law at $L \ll L^*$ and the normalization n^* (units of number density). It is useful to express Eq. (4.27) in terms of absolute magnitude. Remembering that $L_1/L_2 = 10^{-0.4(M_1-M_2)}$ we get:

$$\Phi(M)\mathrm{d}M = n^*(0.4\ln 10)10^{0.4(\alpha+1)(M^*-M)} \exp\left(-10^{0.4(M^*-M)}\right) \mathrm{d}M, \tag{4.28}$$

where M^* is the characteristic absolute magnitude with the same interpretation as L^* in Eq. (4.27).

It is interesting to see how the Schechter fits through the GALFORM model LFs compare to those predicted by observations. To that end, we fit Schechter functions to the LFs obtained from GALFORM models using a non-linear least square method (e.g. Bates & Watts, 1988) using the SciPy numerical package (Jones et al., 2001). Before fitting, the GALFORM LF results were truncated at both high and low end to remove spurious values stemming from numeric issues (a drop-off at the lowest luminosities and occasional erratic values in highest luminosity bins). The fit takes into account the Poisson count errors predicted by GALFORM. Note that this is a somewhat simpler procedure than the one used to fit the HUDF09 data, where the fit is not done on the predicted LF, but on the raw data. See Bouwens et al. (2007) and references therein for details on their fitting procedure.

Before proceeding with the discussion of our results, we must point out that the comparison of Schechter fits does not provide more information than a direct comparison of

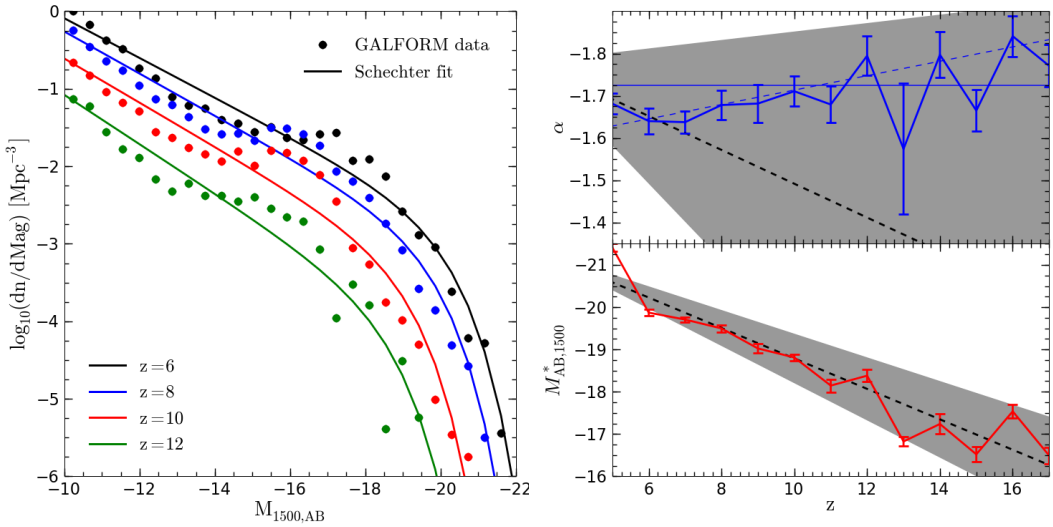

Figure 4.21: *Left:* Schechter fits (lines) through the default Baugh05 model UV LF prediction (circles) at $z = 6$ (black), 8 (blue), 10 (red) and 12 (green). *Right:* Evolution of M^* (bottom) and α (top) Schechter function parameters with redshift. Solid lines with error bars represent fits through the GALFORM data, while the shaded areas give the empirical fits (mean value shown by dashed black lines) from Bouwens et al. (2008b). Dashed blue line in the top panel is a linear fit through GALFORM α values, while the horizontal thin solid blue line marks the mean α. See text for discussion.

LFs does. The fits we perform here are done over a significantly wider range of luminosities than the range covered by current observations. As there is not reason to expect the Schechter fits to perfectly match the observed LF (as can be seen in left panel of Fig. 4.21), we must be careful not to over-interpret these results. Our main motivation here is to demonstrate how the predictions of the model unconstrained by practical limitations, compares to the currently available data.

The evolution of individual parameters of the Schechter function offers some insight into the structure formation process. Notably, the evolution of M^* (i.e. brightening of galaxies) at high redshifts is expected from the evolution of the halo mass function, e.g. the number density of present-day L^* galaxy halo hosts ($\sim 10^{12}\mathrm{M}_\odot/h$) increases by approximately three orders of magnitude between z=10 and 3, while remaining relatively

constant for $z < 3$ (see e.g. Fig. 2 of Springel et al., 2005b). On the other hand, the evolution of α can be very important for reionization, i.e. the slope of the power-law part of the LF indicates the number of low luminosity galaxies that are thought to be the dominant contributors to the LC emissivity as we discussed previously.

Figure 4.21 shows the redshift evolution of the UV LF in described terms of the Schechter function and its parameters. The left panel shows the LF and the corresponding Schechter fits as a function of redshift. It is clear that the Schechter functions are only approximate fits to the LF data. However, some global trends in the evolution of the LFs can be discussed in terms of the Schechter fit parameters, given in the right panel. The evolution of M^* in the considered redshift range is strong, as expected: it increases from ≈ -17 at $z = 16$ to ≈ -20 at $z = 6$ (bottom in right panel). The shaded areas in the right panels mark the range of the empirical fit to the data done by Bouwens et al. (2008b) (see their Eq. 3). The results of the default Baugh05 model show excellent agreement with the empirical relation for the evolution of M^*. The M^* parameter is the only parameter constrained by observations, because it is defined by the high-luminosity end of the LF which is the only part currently observed. Still, as we pointed out before, our model fits cover a much wider range of luminosities, so the perceived perfect match to observations seen in this figure, must be at least partly accidental.

On the other hand, the evolution of the slope of the low-luminosity end of the LF is not so clear. By eye, as seen in the left panel of Fig. 4.21, α evolves weakly at $z \geq 6$, if at all. A direct look at the values of α (top of right panel) shows some weak linear evolution with redshift (dashed blue line is a linear fit through the α values). Arguably, the error bars are sufficiently large not to be able to completely exclude the no-evolution case (demonstrated by the horizontal blue solid line, which shows the mean α). The empirical fit from Bouwens et al. (2008b) is unconstrained for the slope parameter α evolution, so a comparison with GALFORM results would be pointless (we still include the predicted values in the figure for completeness, grey shaded areas). We do note that the values of α from the Baugh05 model agree with the individual values observed at $z \lesssim 7$, $\alpha \approx -1.7$. It was previously argued that a steep low-luminosity end of $\alpha \lesssim -1.6$ is needed for low luminosity galaxies to be capable of reionizing the Universe (Lehnert & Bremer, 2003; Beckwith et al., 2006; Stark et al., 2007). This is true in the Baugh05 model as well, as the majority of ionizing photons is coming from the low-luminosity end of the LF (see top

panels of Fig. 4.17. Note, however, that the value of α in our fits strongly depends on the luminosity range over which we performed the fit. A look at the left panel of Fig. 4.21 shows that the slope of the low-luminosity end of GALFORM LFs is not a simple power-law. More precisely, there is a flat "plateau" around $M_{1500,\mathrm{AB}} \approx (-14) - (-15)$ which separates two power-law like slopes. This feature is caused by the presence of starbursts, that only boost the number densities of galaxies with $M_{1500,\mathrm{AB}} \lesssim -14$. It is therefore clear that the value of α we obtain in our fits may significantly change if we restrict the luminosity range to better match the current observations, which are generally limited to $M_{1500,\mathrm{AB}} \lesssim -17$.

The effects of reionization on the LBG luminosity function

As a final point, it may be interesting to see whether the Lyman-dropout LFs can be used to constraint the reionization in some way. For that purpose, in the left panel of Fig. 4.22 we compare the observed UV LF at $z = 6$ with Baugh05 variants with different photo-ionization feedback parameters, z_{cut} and V_{cut}. In the original Baugh05 (red lines), $z_{\mathrm{cut}} = 6$, which means that it is not affected by any photo-ionization feedback at $z = 6$. If the redshift of reionization is assumed to be earlier, $z_{\mathrm{cut}} = 10$ or 12, there is clear evidence of the suppression of the power-law slope resulting in an underestimate of the observed LF at $M_{\mathrm{AB},1500} \gtrsim -20$. Note that this suppression decreases the importance of less luminous galaxies in the LC emissivity (top panel). If we were sure that $V_{\mathrm{cut}} = 60$ km/s, then this would certainly suggest reionization at $z \approx 6$. However, as we discussed earlier, this value of V_{cut} is probably an overestimate, with a more probable value being $V_{\mathrm{cut}} \approx 30$ km/s. In that case, the observed LF is consistent with a redshift of reionization of $z \approx 10$, the value suggested by the CMB polarisation data (Komatsu et al., 2009). In terms of the Schechter fit slope (right panel), the default model (red line) and $z_{\mathrm{cut}} = 10$, $V_{\mathrm{cut}} = 30$ km/s show approximately the same evolution with redshift. Note that the suppression of α is significantly delayed with respect to the assumed reionization redshift, e.g. when $z_{\mathrm{cut}} = 12$ (green line), the slope does not start to diverge from the default until $z \approx 8$. This is yet another feature caused by the photo-ionization feedback prescription in GALFORM, that, as we will argue in Chapter 5 makes source suppression generally unimportant for the progress of reionization. We cannot argue that these results provide any strong constraints on the reionization process. However, we can claim

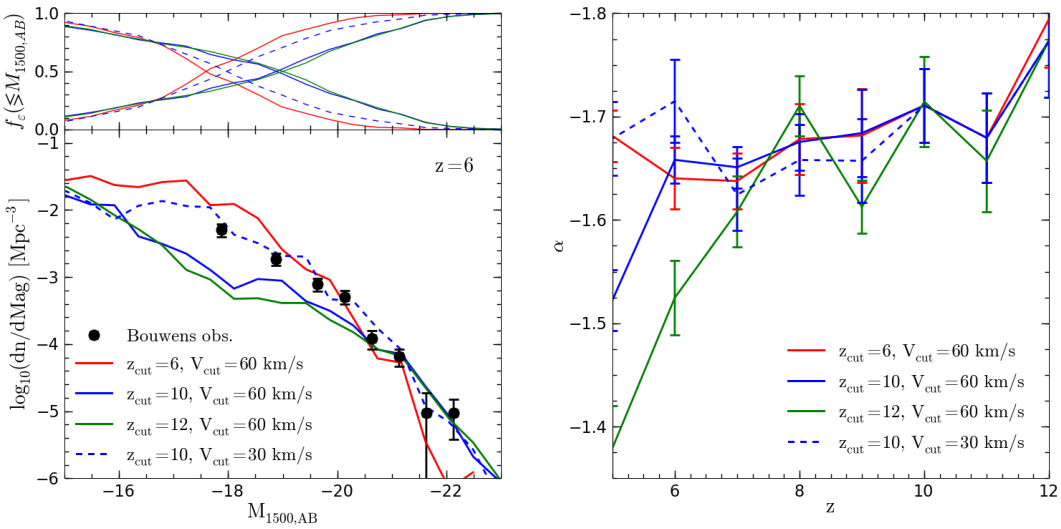

Figure 4.22: *Left:* UV LF at $z = 6$ with respect to varying photo-ionization feedback parameters in the Baugh05 model. Earlier reionization (blue and green solid lines) may depress the low-luminosity LF away from the observed values if $V_{\rm cut} = 60\,{\rm km/s}$. When $V_{\rm cut} = 30\,{\rm km/s}$, the earlier reionization (blue dashed line) does not have a strong effect and the result generally agrees with the default Baugh05 result (red solid line) and observations (circles). *Right*: Evolution of the Schechter slope parameter α with redshift in the same models. The default Baugh05 and $z_{\rm cut} = 10$, $V_{\rm cut} = 30\,{\rm km/s}$ models evolve similarly. The effect of suppression on the low-luminosity end slope is significantly delayed: $z_{\rm cut} = 12$, $V_{\rm cut} = 60\,{\rm km/s}$ does not start to significantly diverge from the default model until $z \approx 8$.

that the comparison with the galaxy LFs can provide yet another avenue for constraining the reionization process when a more detailed models of the relevant physics becomes available.

4.5 Comparison of merger tree generating techniques

Throughout all the previous discussion in this chapter, we have exclusively used merger trees generated with the Monte-Carlo algorithm described in Section 4.2.1. The relatively

small computational cost of this technique allowed us to explore global average photon production in a complete halo sample over the full relevant mass range. On the other hand, for the inhomogeneous reionization simulations that include radiative transfer that we will discuss in Chapter 5, we use the other technique described in Section 4.2.1, which generates the merger trees from the outputs of N-body simulations. This must be done to create the merger trees that are consistent with the density field used for a specific RT run, i.e. to obtain the correct positions for sources.

The use of N-body simulations, due to a much higher computational cost in comparison to the MC technique, results in more strict limitations on mass and time resolution for the merger trees. Indeed, the need to resolve halos of $M \approx 10^8\,\mathrm{M}_\odot/h$ in simulation boxes of $\gtrsim 100\,\mathrm{Mpc/h}$ is one of the key technical problems in numerical simulations of reionization. These conditions have been reached only recently (Trac & Cen, 2007; Iliev et al., 2008; Shin et al., 2008). An alternative used by some authors (e.g. McQuinn et al., 2007; Zahn et al., 2007; McQuinn et al., 2009) is to put the unresolved small halos in by hand, following the analytic prescriptions similar to the ones used to generate the MC merger trees. This is a promising technique when simple source prescriptions that do not require halo merger histories are used. For use with GALFORM, we would need the added halos to be attached to the existing N-body merger trees but, unfortunately, a technique that accomplishes this efficiently has still not been developed. As we will demonstrate in this section, none of the N-body simulations we use in the full RT simulations of Chapter 5 provides a complete sample of the LC emissivity. Keeping this in mind, in Section 4.5.1 we will compare the emissivity obtained from N-body merger trees with the ones obtained from MC trees to estimate the fraction of the LC photon production we will be missing. This, as we will demonstrate, does not simply depend on the mass resolution of the N-body runs, but also on the parameters of the GALFORM model used. Finally, in Section 4.5.2 we test the time resolution of the merger trees that is necessary to produce a converged emissivity.

4.5.1 LC emissivity from N-body merger trees

To test the effects of box size and mass resolution of N-body simulations on the LC emissivities predicted by GALFORM we use a set of simulations obtained with the "lean" version of Gadget-2 (Springel, 2005a). We give the parameters of the simulations in Ta-

Simulation	L_{box}	$N_{\rm DM}$	$m_{\rm DM}$	$M_{\rm halo,min}$
	[Mpc/h]		$[10^5\,{\rm M}_\odot/{\rm h}]$	$[10^8\,{\rm M}_\odot/{\rm h}]$
L10N1024	10	1024^3	0.65	0.013
L20N512	20	512^3	41.35	0.83
L20N1024	20	1024^3	5.17	0.1
L50N1024	50	1024^3	80.78	1.62
L140N1024	140	1024^3	1773.23	35.47

Table 4.1: Parameters of the cosmological N-body runs discussed in Chapters 4 and 5. The labelling scheme was introduced in Chapter 3. The given values are as follows: size of the computational box, L_{box}; number of dark matter particles, N_{DM}; mass of dark matter particles, $m_{\rm DM}$; minimum halo mass, $M_{\rm halo,min} \equiv 20 \times m_{\rm DM}$, given by the chosen minimum FOF group size of 20 particles. The Cosmological parameters [Ω_0, Ω_Λ, $\Omega_{\rm baryon}$, σ_8, n_s, h] used for all runs are [0.25, 0.75, 0.045, 0.9, 1, 0.73].

ble 4.1. Some of these runs have been used in Chapter 3 (Table 3.1), where we introduced the labelling scheme.

The number of halos at some redshift is commonly expressed in the form of a differential halo mass function (MF), like the one shown in Fig. 4.23. The halo mass function ${\rm d}n/{\rm dlog}_{10}M$, is defined as the number of halos of mass M per unit volume per $\log_{10}$ unit mass interval. An analytical form of this function can be, to a first approximation, derived using the Press-Schechter formalism (see Eq. (1.9)). Many improved analytic descriptions have been developed since (e.g. Sheth & Tormen, 1999; Jenkins et al., 2001; Sheth & Tormen, 2002; Warren et al., 2006; Reed et al., 2006). In Fig. 4.23, we compare the halo MFs of our N-body runs to that of Jenkins et al. (2001), who proposed a "universal" form of the halo mass function obtained by fitting N-body simulations and showed it to agree well with the Sheth & Tormen (1999) analytic function up to $z \approx 5$. Later work

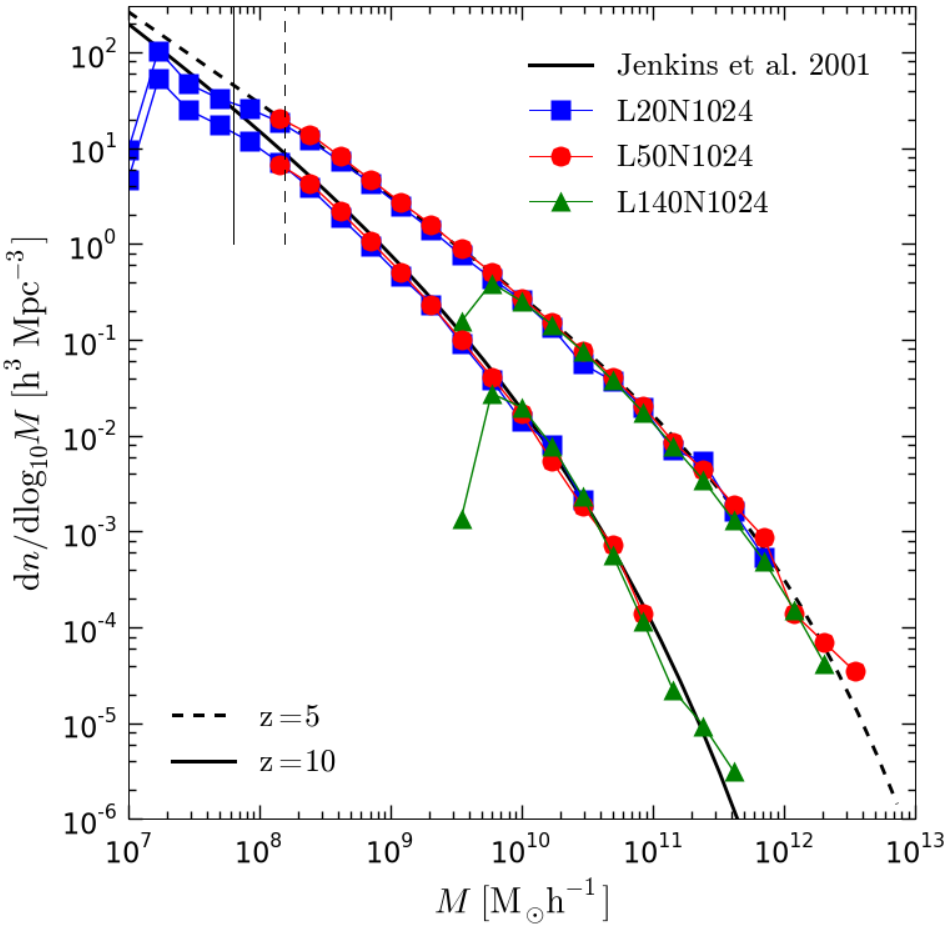

Figure 4.23: Number of halos per halo mass (halo mass function) of three N-body simulations, L20N1024 (blue squares), L50N1024 (red circles) and L140N1024 (green triangles) at redshifts 5 and 10. The results are compared to the Jenkins et al. (2001) mass functions (black solid and dashed lines) used to predict the number of halos at each redshift in the GALFORM run. The vertical thin lines mark the minimum halo mass necessary for atomic line cooling in the halo. In the fully resolved ranges, all the N-body runs agree well among themselves and with the Jenkins et al. predictions.

further improved on their results, especially at higher redshifts (Reed et al., 2006). Still, here we use the form proposed by Jenkins et al. (2001) because this is the mass function we employ when generating the merger trees with the Monte-Carlo technique in

GALFORM.

The smallest halo that can be found in an N-body simulation is given by the mass resolution, i.e. the mass of a single particle, and by the properties of the group finder. For all the simulations of this thesis we use the friends-of-friends group finder (FOF Davis et al., 1985) with the linking length of $b = 0.2$ and we consider a group to be a halo if it has ≥ 20 particles. The particle masses and corresponding minimum halo masses for individual simulations are shown in the last two columns of Table 4.1. The 20 particles criterion per halo is chosen to extract as many halos as possible from the N-body data and may be somewhat optimistic, as evidenced by the large deviations of the smallest halo mass bins from the theoretical expectation in Fig. 4.23 (when high confidence of halo detection is needed, a higher group size is used, e.g. ≥ 100 in Reed et al., 2006). The number of high mass halos is dictated by the box size, since the halo mass function can also be interpreted as a probability of finding a halo of a given mass in unit volume. The effect of the box size is clear when comparing L50N1024 and L140N1024 at $z = 10$ in Fig. 4.23. Finally, the halo counts from each N-body simulation will differ from theoretical predictions due to finite box size which limits the maximum wavelength in the power spectrum of density fluctuations (e.g. Sirko, 2005; Bagla & Prasad, 2006; Power & Knebe, 2006).

When a simple mass-luminosity relation is used to obtain the ionizing emissivity from the N-body data, such as the linear relation of Iliev et al. (2006a), the halo mass function provides all the necessary data to model photon production. In that case, the fraction of total emissivity resolved by a single N-body run is completely defined by the minimum and maximum halo mass range on the halo mass function. The same is not true in the GALFORM model, as can be seen in the left panel of Fig. 4.24 which compares the emissivity of the default Baugh05 model obtained using the N-body trees from three different simulations with the trees computed using the Monte-Carlo technique. For example, the minimum halo mass in the L50N1024 simulation is $1.62 \times 10^8 \mathrm{M}_\odot/h$ (Table 4.1) but, in the left panel of Fig. 4.24, it does not match the emissivity of the L20N1024 run (which has $M_{\mathrm{halo,min}} = 10^7 \mathrm{M}_\odot/h$) until $M \approx 10^9 \mathrm{M}_\odot/h$, even though the halo mass functions agree very well between the two simulations in that mass range (see Fig. 4.23). This discrepancy shows that the emissivity depends on more than the simple number of halos at any given redshift: in this case, it depends strongly on the starbursts triggered

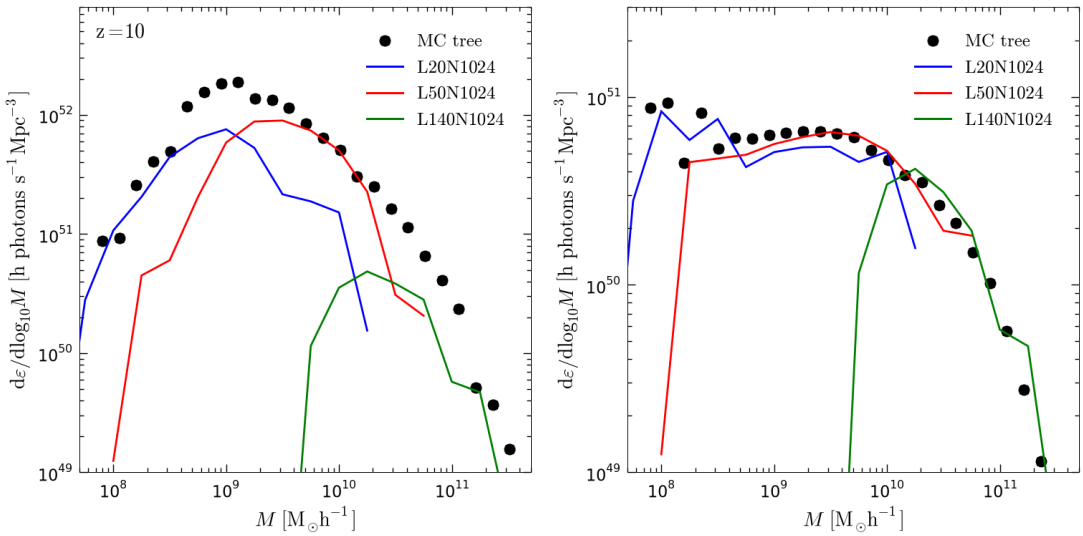

Figure 4.24: The LC emissivity as a function of halo mass in the default Baugh05 (left panel) and Baugh05 without starbursts (right panel) models at $z = 10$ for three N-body runs and an MC tree run. The N-body runs are shown with the same styles as in Fig. 4.23. In the default model (left), the N-body runs resolve different ranges in the emissivity, but the ranges do not correspond directly to the minimum and maximum halo masses resolved by each simulation (Fig. 4.23). The reason for this is the importance of galaxy mergers in triggering starbursts, as evidenced by the model without them (right panel). The discrepancy between the Monte-Carlo and N-body runs is also probably due to different merger rates from the two techniques.

by galaxy mergers. The last claim is backed up by the right panel of Fig. 4.24 which shows the results using the same merger trees as in the left panel but in a Baugh05 model variant without any starbursts. Here, the mass ranges where the emissivity is properly resolved correspond to the fully resolved halo mass function ranges of each simulation.

The left panel of Fig. 4.24 also shows a discrepancy between the results of N-body and MC tree runs. Some comparisons of the two techniques have been done before, (e.g. Helly et al., 2003; Harker et al., 2006), but the focus of those studies was on lower redshifts and on other properties of the galaxy populations (such as the optical or near-IR luminosity functions), not the LC emissivity. When the starbursts are excluded from the

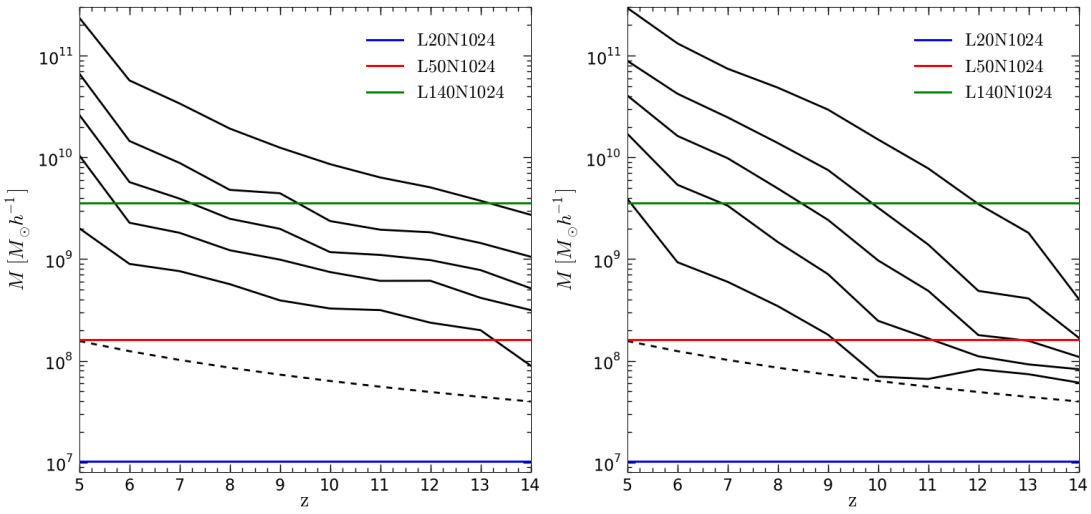

Figure 4.25: The halo mass ranges relevant for LC photon production as a function of redshift in the default Baugh05 (left panel) and Baugh05 without starbursts (right panel) models. The black solid curves mark halo masses that account for a fraction, f, of the total emissivity produced at a given redshift z, $M \geq M(z, f)$; from top to bottom, the fractions are $f = (0.1, 0.3, 0.5, 0.7, 0.9)$. The dashed black line marks the mass of halos with $T_{\rm vir} = 10^4 K$, the temperature necessary for atomic line cooling, i.e. the minimum mass of halos capable of forming galaxies. The coloured lines mark the minimum resolved halo masses of the three N-body simulations used (Table 4.1), same runs and colour scheme as in Fig. 4.23. See text for further discussion.

Baugh05 model (right panel), the discrepancy between the two merger tree creation procedures disappears, pointing to the different halo merger rates in the two techniques as the source of the problem. We will do a more detailed study of the merger tree properties at high redshift and their effect on the galaxy formation in the future.

Regardless of the differences between N-body and MC tree emissivities in the Baugh05 model, the shape of the emissivity's dependence on halo mass is approximately the same. This allows us to discuss the importance of halos of a certain mass to LC photon production in relative terms. In Fig. 4.25, we show the fraction of the total LC emissivity produced in halos with mass smaller than M as a function of redshift (we were previ-

ously showing the fractions at a single redshift in the top panels of the emissivity plots in Section 4.3). For a halo to be capable of cooling efficiently enough to form galaxies, it has to have $T_{\rm vir} \gtrsim 10^4 K$. The halo mass that corresponds to that virial temperature, given by Eq. (1.7), is shown with the dashed line in both panels of Fig. 4.25. The solid lines mark the halo masses up above which a specified fraction, f, of LC photons is produced at redshift z, $M(z, f)$. For example, halos in the mass range bounded by the dashed and the bottom solid line account for the first 10% of the photons produced. Comparing $M(z, f)$ ranges with the minimum halo mass resolved by an N-body simulation, and ignoring for the moment the importance of galaxy mergers in the default Baugh05 model, we can see what fraction of LC emissivity is missed by each N-body simulation due to its limited mass resolution. The horizontal coloured lines mark the mass limit of the simulations we are discussing. The limiting halo mass of the L20N1024 run (blue line) is well below the $M(z, T_{\rm vir} = 10^4 K)$ halo mass limit, suggesting that the low mass halo contribution to the total emissivity is completely resolved at all redshifts - this is indeed the case, as seen in Fig. 4.24. By comparison, L50N1024 run (red line) resolves the limiting halo mass only at $z = 5$, while L140N1024 (green line) resolves practically nothing in terms of LC photon production at $z = 14$ and $> 50\%$ of the photons only for $z \lesssim 7$ in the default Baugh05 model (left panel).

Comparison of the two panels of Fig. 4.25 shows how the relevant mass range changes with respect to the GALFORM model employed. A model with significant contribution from merger-induced starbursts (left panel) produces the majority of the photons (e.g. 80%, produced in the mass range defined by the top and bottom solid lines) in a smaller halo mass range than the one without them (right panel). Also, in the former case, the small mass halos, $M \leq 10^8 {\rm M}_\odot/h$, contribute a negligible amount of photons to the total emissivity regardless of their high numbers (as seen in Fig. 4.23), as the progenitors of these halos are not capable of forming galaxies ($T_{\rm vir} < 10^4 K$) resulting in no galaxy mergers that could trigger starbursts. When the bursts are ignored (right panel), the contribution of $M \lesssim 10^8 {\rm M}_\odot/h$ halos becomes much more relevant: they produce $\approx 50\%$ of the total emissivity at $z = 14$ and $\approx 20\%$ at $z = 10$.

The ratio $M(z, f)/M(z, T_{\rm vir} = 10^4 K)$ also holds some useful information. As seen in Fig. 4.23, with decreasing redshift the number of high mass halos grows significantly faster than the number of small mass ones, a feature of the hierarchical buildup of struc-

ture. Keeping this in mind, if the increase of emissivity is dominated by the growth of structure, then the ratio $M(z, f)/M(z, T_{\rm vir} = 10^4 K)$ would increase with time, possibly more so for high than for low f. This is indeed the case in the no starbursts model (right panel), where the solid lines clearly diverge from the dashed one with decreasing redshift. On the other hand, in the default model (left panel), the same ratio increases much more slowly, suggesting that the evolution of emissivity does not directly correlate with that of the halo MF. Note that the ratio changes discontinuously at $z = 6$ in both models: this is the choice of $z_{\rm cut}$ in the default Baugh05 model for the redshift at which the photo-ionization feedback is turned on (see Section 4.3.8).

Finally, we ask the question of how much does the limited resolved emissivity range in the N-body runs affect the progress of reionization. Figure 4.26 shows number of LC photons produced per H atom up to some redshift by the default Baugh05 model using full mass range MC trees (black solid line) and N-body trees from the previously used simulations (coloured lines). The very limited mass resolution of L140N1024, as seen in Fig. 4.25, causes a severe delay in the resulting ionization of the Universe: this run produces a single photon per hydrogen atom $\Delta z \approx 3.7$ later than the MC tree run, and 10 photons per atom $\Delta z \approx 3$ later. Clearly, the production of photons speeds up at lower redshifts for the L140N1024 trees, as more sources are being resolved as seen in Fig. 4.25. The other two N-body tree runs obtain results much closer to the MC tree ones. An interesting feature is the cross-over of photon buildup between L20N1024 and L50N1024. This is, of course, due to the increasing contribution to emissivity from larger mass halos. At early times ($z \gtrsim 12$), low mass halos ($M \lesssim 10^9 {\rm M}_\odot/h$) dominate the emissivity and the L20N1024 simulation, which fully resolves that halo mass range, almost matching the results of the MC tree run. The small box size of the L20N1024 run means that more massive halos are not found in that simulation and the photon production quickly starts to diverge from the desired result with decreasing redshift. An opposite scenario takes place in the L50N1024 simulation: while the low mass halos are not resolved, the larger ones that become increasingly important are, and the result gradually converges on the desired one (MC tree) at lower redshifts. Note that we showed in the left panel of Fig. 4.24 that the use of MC merger trees in the default Baugh05 model results in slightly higher emissivities than the ones obtained using N-body trees, therefore a perfect agreement in the total number of LC photons produced is not expected.

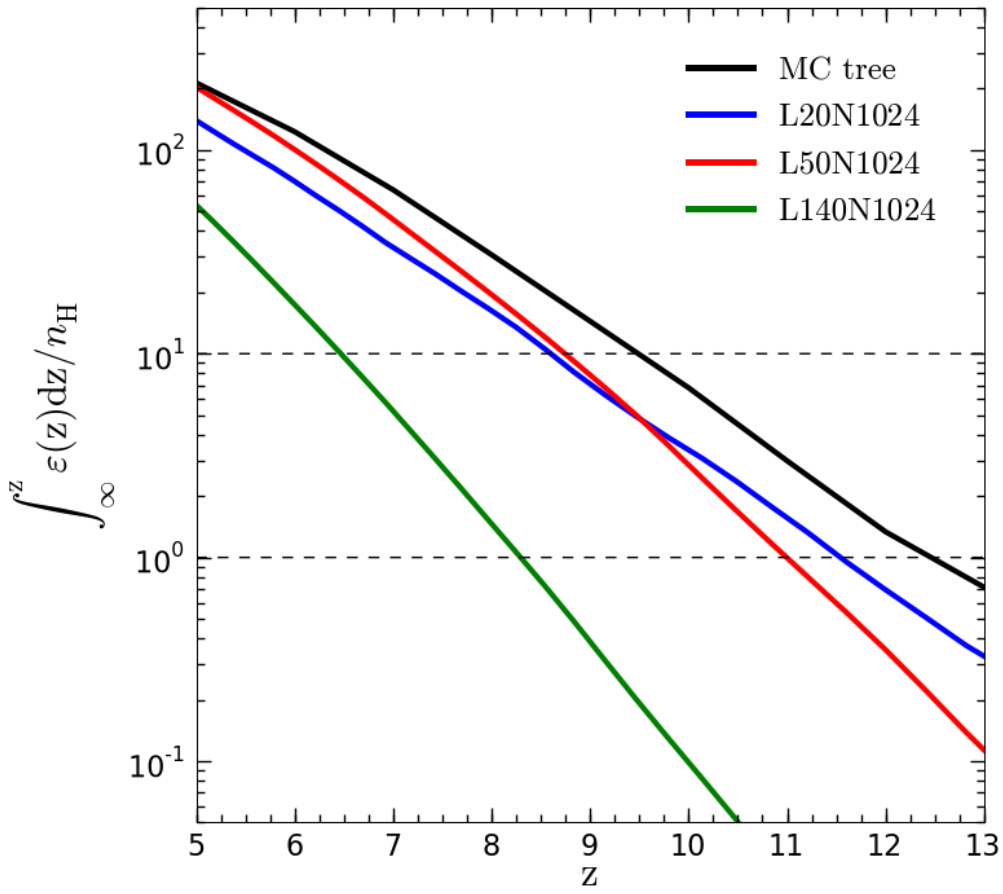

Figure 4.26: The number of LC photons produced per H atom by redshift z using the default Baugh05 model with different merger trees. The Monte Carlo merger trees (black line) resolve the full halo mass range relevant for LC emissivity, while the three N-body trees (coloured lines) resolve different sections of the mass range, as seen in Fig. 4.24. The small box run (L20N1024, blue line) resolves the small halo mass range which is dominant during the early stages of the reionization process ($z \gtrsim 12$), but does not resolve higher mass halos that become relevant at later times. The medium range box (L50N1024, blue line) misses some of the smallest halos at early times, but makes up for it by resolving more massive ones later on.

Coincidentally, the L20N1024 and L50N1024 runs produce 10 photons per H atom at about the same time. This means that we cannot choose the optimal run out of the two

on the reionization redshift alone. What is clear, though, is that the limits in the resolved emissivity of both runs leads to a significant delay in the redshift of reionization, $\Delta z \approx 1$, when compared to the results of the MC tree run. Still, this difference is not larger than what is obtained by varying some of the more important GALFORM model features, as we have shown in Section 4.3. This leads us to conclude that the limited resolution of our N-body simulations will not prevent us from obtaining relevant results in the full RT simulations we perform in Chapter 5.

4.5.2 Time convergence of N-body merger trees

Thus far, we have focused solely on the issues connected to the resolving of the halo MF. Another aspect of the source model is the time resolution of the merger trees: how many steps of the tree, $N_{\rm steps}$ mentioned in Section 4.2.1, must be used to obtain a converged estimate of the production of LC photons during reionization. When using Monte Carlo trees, the computational cost of increasing the time resolution of merger trees used in GALFORM is small (remember that a very fine merger tree is always created when using this technique). On the other hand, the time resolution of the N-body merger trees is defined by the available outputs of the parent simulation. This becomes an issue especially when one wishes to employ very large simulations (in terms of particle numbers), when the limit to $N_{\rm steps}$ becomes dependent on having available disk space. As we intend to employ one such simulation in our future work, the Millennium-II run (Boylan-Kolchin et al., 2009), we need to determine the minimum number of outputs that the N-body trees must have to provide a reasonably converged estimate of LC emissivity.

To find the answer to this question, we use the L20N512 simulation (Table 4.1). This simulation has approximately the same mass resolution as L50N1024, but in a smaller box, therefore it is not adequate for resolving the full halo mass range discussed in the previous section. Here, we are not interested in what fraction of emissivity is properly resolved, but only in obtaining a converged result from an individual simulation. The L20N512 run is adequate for this purpose.

The L20N512 run produced 100 snapshots between $z = 5$ and $z = 26.71$, equally spaced in $\log_{10}a$, $\Delta\log_{10}a \approx 0.007$. This spacing is similar to the one employed in the Millennium simulations at lower redshifts (Springel et al., 2005b; Boylan-Kolchin et al., 2009). In physical time, the longest time step is $\Delta t \approx 27$ Myr at $z = 5$, while $\Delta t \approx 11$

Myr at $z \approx 10$, which is somewhat shorter than time steps used in other reionization numerical models (e.g. Iliev et al., 2006a; McQuinn et al., 2007). To test the effect of time resolution, we create more than one set of merger trees from this simulation: one that uses halos found in every snapshot and three others which use every second, fourth and eight snapshot.

Figure 4.27 shows the cumulative production of LC photons of the default Baugh05 GALFORM scheme employing these 4 sets of merger trees. The majority of the difference in the photon production seen in this figure is due to the decreasing frequency of galaxy merger-triggered starbursts, which we showed to be crucial for the LC emissivity (Section 4.3.3). This effect is more pronounced at early redshifts[13] where the galaxies, and therefore their mergers, are much less frequent. The convergence trend with the increasing number of steps is clear. The difference between the highest and lowest time resolution runs (12 and 100 steps, green and black lines respectively) is substantial, resulting in $\Delta z \approx 1.5$ delay in reionization redshift for the coarser one. On the other hand, using every second snapshot to construct the merger trees (red line in Fig. 4.27) has a very small effect, with only $\approx 20\%$ less photons produced and resulting in a modest $\Delta z \approx 0.1$ delay when compared to the highest time resolution tree.

The difference in results of the two highest time resolution runs in Fig. 4.27 is much smaller than what may result due to changes of GALFORM model parameters (as discussed in Section 4.3) or due to cosmic variance in the N-body runs themselves. We therefore conclude that ≈ 50 N-body merger tree steps, equally spaced in $\log_{10}a$, is adequate for resolving the evolution of emissivity in the $5 \lesssim z \lesssim 25$ redshift range, $\Delta\log_{10}a \approx 0.013$. We have employed this time resolution in all of the N-body GALFORM runs throughout this thesis. Note that we have also checked the dependence of the ionized fraction in the cosmologically expanding density field on the number of discrete snapshots used to represent it (both in the Stromgren sphere test described in Section 3.5 and in a number of previously mentioned N-body runs). We find that the number of snapshots needed to obtain a converged result from a coupled scheme with GALFORM and RT depends completely on the convergence in GALFORM alone, i.e. the number of snapshots needed for RT is much lower.

[13] $z > 15$, not shown in Fig. 4.27.

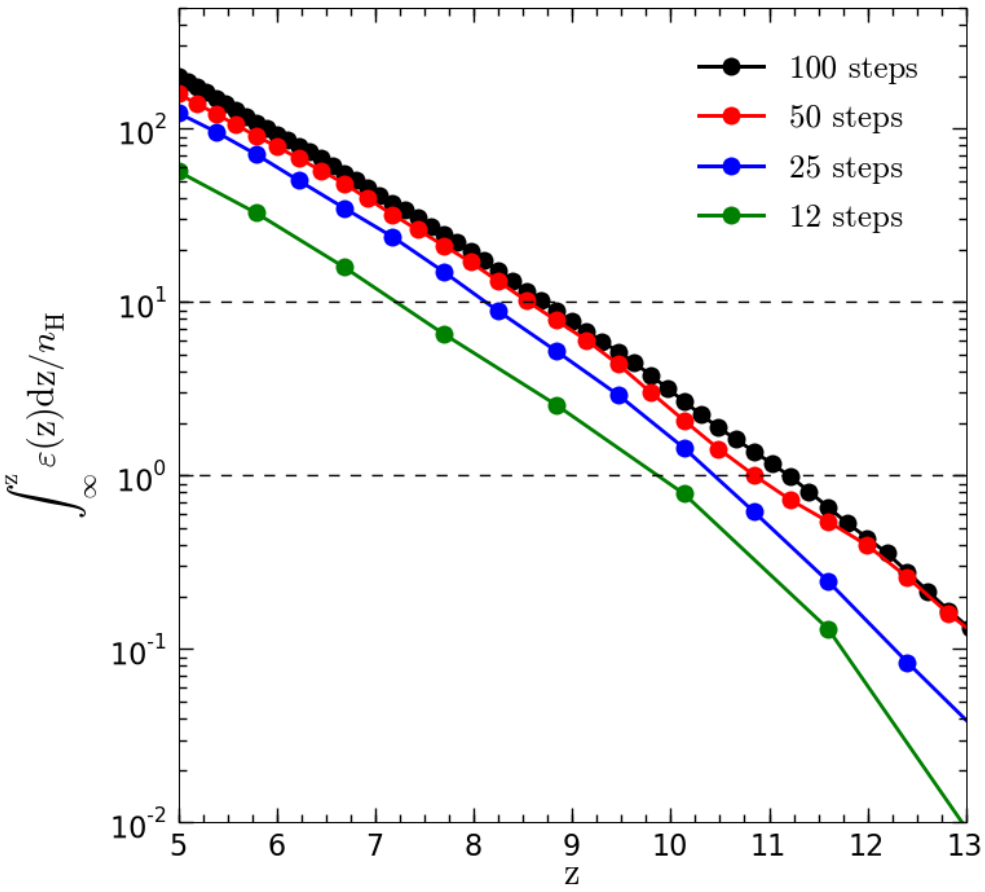

Figure 4.27: The number of LC photons per hydrogen atom produced up to redshift z by the default Baugh05 model employing merger trees constructed using different numbers of L20N512 simulation snapshots. The snapshots are spaced equally in $\log_{10}a$ between $z = 5$ and $z = 26.7$. The production of LC photons converges with increasing number of tree steps over this redshift range. The difference between the highest time resolution run (100 snapshots in z range, black line) and one with two times longer steps (50 snapshots, red line) is no more than $\approx 20\%$ in the plotted redshift range, resulting in a $\Delta z \approx 0.1$ difference in reionization redshift. This result suggest that 50 tree steps provide sufficient time resolution for LC emissivity in our redshift range of interest.

4.6 Conclusions

This chapter presented the LC emissivity generated by galaxies during the presumed reionization period ($6 \lesssim z \lesssim 15$), as predicted by the GALFORM semi-analytic galaxy

formation and evolution model. The Monte Carlo merger tree generation technique that is a part of the GALFORM scheme allowed us to explore the effects of various model parameters on LC emissivity without the limitations introduced by the resolution issues in N-body simulations. We chose to focus on one of the standard Durham GALFORM models, developed by Baugh et al. (2005). We showed that the other fiducial model, by Bower et al. (2006), does not produce enough ionizing photons to finish reionization by $z \approx 6$.

The key element of the Baugh05 model is the assumed top-heavy IMF during bursts of star formation, triggered by galaxy mergers or disc instabilities. The use of the top-heavy IMF instead of the single, Kennicutt (1983) IMF in bursts introduces a $\Delta \approx 2.5$ shift in the reionization redshift, the largest difference caused by a single model parameter (see Fig. 4.4). The top-heavy IMF is therefore crucial for achieving reionization by $z \approx 10$, the redshift suggested by the CMB τ_e data (Komatsu et al., 2009; Dunkley et al., 2009). We must emphasise that the top-heavy IMF was not originally included to produce enough LC photons during reionization, but to provide a match with observed sub-mm galaxies. The case is further strengthened by Nagashima et al. (2005a,b), who found that the use of the top-heavy rather than the solar neighbourhood IMF in bursts results in much better agreement to the observed metallicities in intracluster gas in clusters and stars in elliptical galaxies. The Baugh05 model was also shown to provide a good match for the Lyα-emitters in $z \sim 3-9$ range (Delliou et al., 2005, 2006; Orsi et al., 2008), Lyman-break galaxies in $z \sim 3-10$ (Baugh et al., 2005; Gonzalez et al., 2009, see Section 4.4) and infrared Spitzer[14] observations (Lacey et al., 2008). The large importance of the top-heavy IMF suggests that the reionization is practically completely driven by bursts, thus making reionization sources more luminous and less frequent (see Fig. 4.6). This scenario is different from the ones commonly examined by numerical RT simulations (e.g. Ciardi et al., 2003; Sokasian et al., 2003; Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007) and may introduce a specific imprint on the morphology of HII regions during reionization, which we discuss in Chapter 5.

Supernova feedback is also an important regulator of LC emissivity. Figure 4.7 shows that stronger SN feedback parameters (close to the ones employed by Bower06) introduces a $\Delta z \approx 2$ delay in reionization redshift, nearly as much as the top-heavy IMF.

[14]http://www.spitzer.caltech.edu/

A weaker SN feedback has a significantly smaller effect ($\Delta z \approx 0.5$ earlier reionization), which suggests that the SN feedback used in Baugh05 are quite modest in suppressing the LC emissivity. The largest boost in the production of LC photons is found in models that include disc instabilities ($\Delta z \approx 2$ earlier than default for the strong disc instability model, see Fig. 4.15). The disc instabilities are more effective in triggering bursts than galaxy mergers, which results in more star formation occurring in the burst mode with the top-heavy IMF.

We compared the predictions of the Baugh05 model with the recently published Lyman-break galaxy LFs reaching $z \sim 6 - 10$ (Bouwens et al., 2008a, 2009a,b). We found that the model provides a good match for the observed LBGs in this redshift range. The differences between GALFORM models, e.g. Baugh05 and Bower06 in Fig. 4.1, are not as clear as when comparing LC photon production. This is because the observed magnitude range covers only the high-luminosity end of the LF. At redshifts $z \gtrsim 7$, the observed LF accounts for less than 20% of the total photon production in the default Baugh05 model (see e.g. top panels in Fig. 4.17). This suggests that the LC emissivity predictions derived from the Schechter fits of the observed data (e.g. Bouwens et al., 2008a) may be significantly off, i.e. low-luminosity LF slope parameter α is not well constrained. However, the Schechter fits through our GALFORM data are in general agreement with the current observations (Fig. 4.21), so those estimates may not be too wrong.

The Baugh05 model LC emissivities suggest (if one is willing to allow the use of a top-heavy IMF in bursts) that reionization can easily be achieved with Population II-type sources alone. However, some reasonable variations of the model (e.g. the solar neighbourhood IMF in bursts) may indeed result in such a scenario. Of course, the emissivities alone do not determine reionization, as the inclusion of photon sinks (recombinations, minihalos, Lyman-limit systems) will have a strong effect on the process. A full model may therefore be much more photon starved and maybe even require the inclusion of sources currently thought to be sub-dominant, such as Pop III stars or mini-quasars, in order to complete reionization within the redshift range suggested by observations.

Chapter 5

Inhomogeneous reionization with SimpleX and GALFORM

5.1 Introduction

In Chapter 4, we discussed in some detail the issue of modeling high redshift galaxies, which are thought to be the most likely sources of ionizing radiation that causes reionization (e.g. Madau, 1999; Barkana & Loeb, 2002; Loeb, 2006; Trac & Gnedin, 2009). Discussion of the ionizing emissivity forms a general basis of analytic and semi-analytic models of the reionization process (e.g. Haiman & Loeb, 1997; Miralda-Escude et al., 2000; Haiman & Holder, 2003; Wyithe & Loeb, 2003; Furlanetto et al., 2004; Benson et al., 2006). What these models usually miss is the spatially inhomogeneous nature of the reionization process, which can only be accurately captured using numerical RT simulations. Reionization is characterised by two distinct stages, as discussed by e.g. Gnedin (2000b). First, individual HII regions are formed around the early sources of ionizing radiation. As more sources are formed and the emissivity increases, the HII regions grow and begin to overlap, exposing the remaining neutral gas to even more intense radiation. The final overlaps mark the end of reionization, after which any region in space can "see" ionizing radiation from a large number[1] of sources, i.e. one can begin discussing a more uniform photo-ionization background (e.g. Haardt & Madau, 1996;

[1] After reionization, the mean free path of ionizing radiation is governed by the number density of Lyman-limit systems (e.g. Miralda-Escude, 2003).

Mesinger & Furlanetto, 2009). This process cannot be modelled properly without the inclusion of radiative transfer. First, RT is necessary to accurately compute the efficiency of recombinations and other ionizing photon sinks in the cosmological density field we previously discussed in Section 3.4.1 This is crucial for finding out how many ionizing photons are actually necessary to complete reionization, a number we were choosing fiducially in Chapter 4. More importantly, RT is crucial in modeling the evolution of galactic HII regions that are going to be directly observed by future hydrogen 21cm line observations by telescopes such as LOFAR[2], PAPER[3] or SKA[4]. The exact way the HII regions propagate during reionization also affects the process itself: the photo-ionizing radiation feedback may suppress the formation and emissivity of sources found in ionized regions, creating a self-regulating loop (Iliev et al., 2007).

The current lack of constraints on many components of the reionization process, most notably on the nature of the high-z galaxies and other ionizing radiation sources, requires any model to be highly computationally efficient in order to allow the exploration of the vast parameter space. In this chapter, we discuss one such model we created by coupling the GALFORM galaxy formation scheme (Chapter 4) with the SimpleX RT code (Chapter 3). GALFORM provides us with a much more detailed model of ionizing radiation sources than the ones commonly used in the literature (e.g. Gnedin, 2000b; Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007), while the computational efficiency of SimpleX allows the production of many realisations of the reionization process in a relatively short time. The concurrent running of the two components allows for the time and space-varying photo-ionization feedback to be included in the model. Our use of a semi-analytic galaxy formation model to provide ionizing radiation sources is similar to the scheme by Ciardi et al. (2003). In contrast to that work, our model includes the previously mentioned suppression of sources by photo-ionizing feedback and has a much higher spatial resolution due to the adaptive nature of the SimpleX Voronoi grid (discussed in Chapter 3). It is also significantly faster due to both the computational efficiency of our RT technique and the parallelization of our code. The adaptive computational grid is also a major improvement over many other models (Iliev et al., 2006a; McQuinn et al., 2007; Trac & Cen, 2007), but a more important distinction from those

[2]http://www.lofar.org/

[3]http://astro.berkeley.edu/ dbacker/eor/

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

works is the detailed model of the relevant galaxy population provided by GALFORM.

This chapter is organised as follows. In Section 5.2, we discuss the numerical implementation of our method. The issues stemming from the spatial resolution of the RT grid are discussed in Section 5.3. In Section 5.4 we present some early results obtained with the model. We discuss the topological progression of reionization in our model (i.e. which overdensity regions are ionized first) and the role of the density field in Section 5.4.1. In Section 5.4.2, we show how our choice of the source model affects the typical sizes of HII regions during reionization. Finally, in Section 5.4.3 we gauge the importance of source suppression on the evolution of the mean ionized fraction in our model.

5.2 Method

The link between the source population (computed by GALFORM) and the propagation of ionizing radiation (modelled using SimpleX) is the cosmological density field. We represent the continuous evolution of the density field with a series of snapshots from dark matter-only N-body simulations, as is commonly done in reionization simulations (e.g. Ciardi et al., 2000; Iliev et al., 2006a; McQuinn et al., 2007). The baryons are assumed to trace the dark matter, as discussed in Section 3.4.1. We also exclude all the particles with overdensities $\Delta > 100$ from the density field, as a way of removing dark matter halos for reasons also discussed in Section 3.4.1. The N-body data is used to construct halo merger trees for use with GALFORM, as discussed in Section 4.2.1. The number and time spacing of snapshots needed to obtain a converged result is dictated by the necessary time resolution of the merger trees, discussed in Section 4.5.2. The use of static N-body snapshots significantly reduces the cost of a single model run, as it allows for the density field to be computed only once and reused for the RT runs. Note, however, that some authors chose a more consistent approach by evolving the density and ionization fields concurrently (e.g. Trac & Cen, 2007). For all the simulations we use in this chapter, we produce ≈ 50 snapshots between redshifts ≈ 25 and 5, with output times corresponding to an equal spacing on the $\log_{10} a$ scale, $\Delta \log_{10} a \approx 0.013$.

The key element of the model is the "communication" between the two components (GALFORM and SimpleX), added to model space and time-dependent photo-ionization feedback. The algorithm of the method is as follows (see the diagram in Fig. 5.1). First,

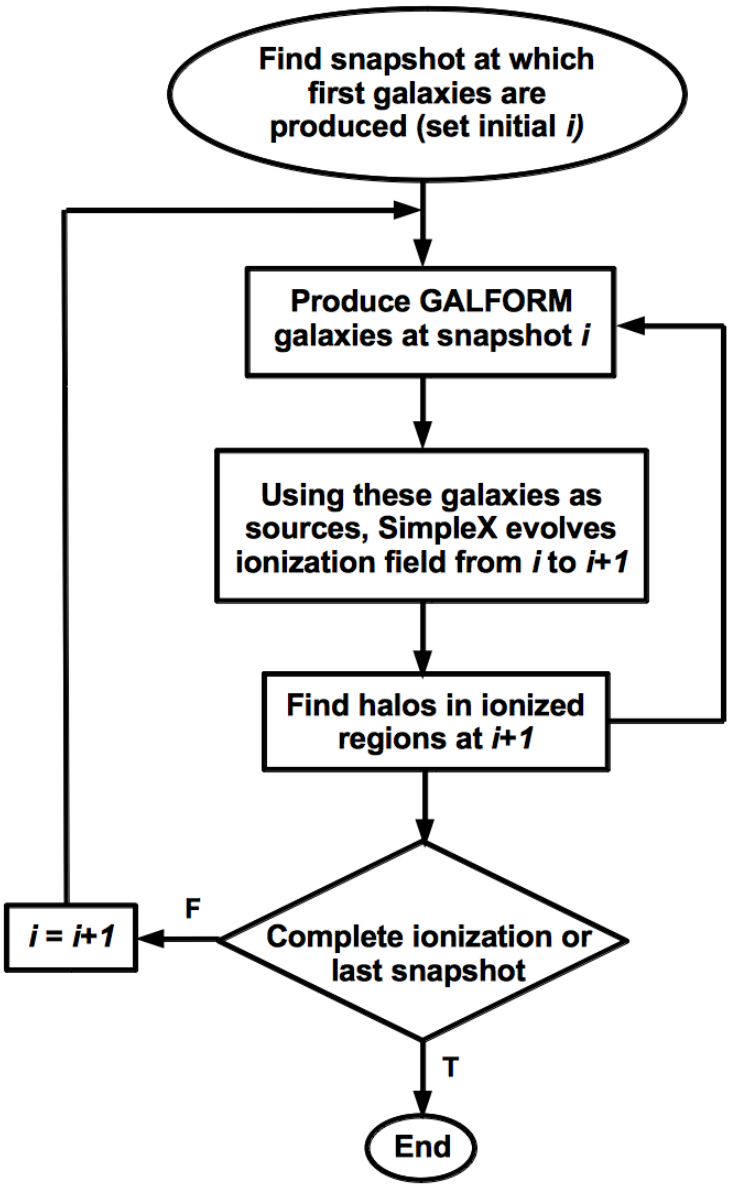

Figure 5.1: A diagram explaining the GALFORM and SimpleX coupling algorithm. The right-hand side arrow connecting the third step to the first signifies that the ionization state information from the end of step i is used when estimating galaxies at step $i + 1$.

we determine the redshift (and its corresponding snapshot) on the merger trees where the first galaxies appear. In practise, we do this by running GALFORM alone for each model once, until the first galaxies appear in the output. This provides us with the

starting snapshot for the subsequent runs using the same GALFORM model and N-body trees. Next, for every following snapshot i, we take the following steps:

1. Run the GALFORM model and output the galaxy population at the redshift of snapshot i.

2. Extract positions and LC luminosities of these galaxies to use as sources for the RT calculation.

3. Using an intermediate code (based around a parallelized octree search routine), setup the initial conditions for RT at snapshot i:

   - Load the positions and masses of N-body particles representing the density field of snapshot i (this can be either all particles from the original N-body run or a sub-selection to which we assign leftover masses[5], as discussed in 3.3. These particles provide nuclei and masses of the Voronoi grid cells representing the density field.

   - Copy the ionization state in each N-body particle from the previous snapshot SimpleX output. We use the same particles of each snapshot, identified by their IDs stored in the Gadget output. Using these IDs, we copy the ionized fraction from the previous snapshot's SimpleX output to the current snapshot initial conditions. Note that this procedure may result in the number of ionized atoms not being conserved between snapshots, due to the change of mass assigned to individual particle i.e. Voronoi cell. This issue has no effect on the evolution of the global ionized fraction and only marginally affects the redshift at which the earliest halos are found to be ionized. We checked this by implementing an alternative scheme which explicitly conserves the number of ionized atoms at the expense of performance.

   - For each GALFORM galaxy, find the nearest N-body particle representing the density field and assign its LC luminosity to it. We remind the reader that this procedure is equivalent to finding in which Voronoi cell a given galaxy lies.

[5]When we use a sub-selection of particles as Voronoi nuclei, the mass of unused particles is assigned to the nearest selected ones and a new N-body file is produced with these modifications.

4. Using the initial conditions set-up in the previously described fashion, we construct the SimpleX Voronoi grid and evolve the ionization field from snapshot i to $i+1$. The density field is assumed to be static and the LC luminosity of sources is assumed to be constant over this period.

5. Using the SimpleX output file produced in step 4, find which halos are in the ionized regions at the redshift of snapshot $i+1$. This is done again by using the octree-based neighbour search. The halo is marked as ionized if it is found in a Voronoi cell with ionized fraction $x_{\rm HII} \geq 0.7$. This value is chosen fiducially and the exact choice from the range $\xi = 0.5 - 1$ does not affect our results[6]. The ionization state of individual halos is then added to the halo merger tree data and used by GALFORM to determine which halos are affected by photo-ionization feedback in the next step.

6. If $i+1$ is not the last snapshot and the box is not completely ionized, repeat the procedure starting by running GALFORM up to snapshot $i+1$.

The key for coupling the two codes is step 5, when the ionization state of each halo is determined. At the moment, this is only a boolean value which tells whether the halo is in the ionized region or not at that redshift. The end product is therefore the ionization history for every merger tree, a cartoon of which is shown in Fig. 5.2. Note that a halo that is found to be ionized at redshift i may become neutral at some later time, both due to recombinations reducing the host cell's ionized fraction and the halo itself moving to a less ionized neighbouring cell. This information is used to replace the single redshift of reionization, $z_{\rm cut}$, used in the GALFORM scheme and introduced in Section 4.2.7. Instead, we cut the cooling only in halos that are marked as ionized at a given redshift. This allows for the inhomogeneous nature of the reionization process to be taken into account in the formation and evolution of GALFORM galaxies.

The previously described procedure is automated with a Python script. The input to the scheme consists of the choice of GALFORM parameters and the N-body data to be used. The outputs are galaxy populations and ionization fields for each snapshot, as well as the reionization history of each halo found in the merger trees.

[6]The results remain unaffected in the large boxes, where the I-front is not resolved by more than one cell as is the case in the runs we discuss in this chapter (see Section 5.4.1)

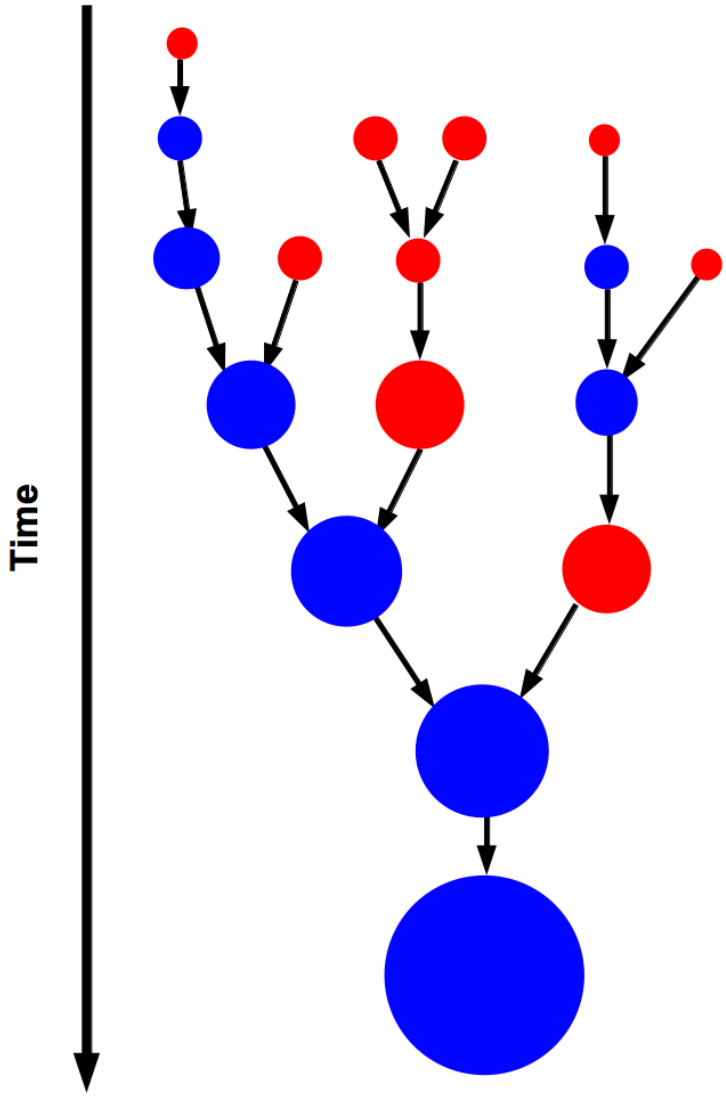

Figure 5.2: A cartoon showing the ionization history of a single merger tree. All halos (circles) on the tree are either in an ionized (blue) or neutral (red circles) region. The halo can be in a neutral region, even though all of its progenitors are found to be in an ionized region (see e.g. right hand side branch). The size of the circle indicates halo mass. As reionization progresses, more halos are found in ionized regions, finally all being ionized after the final overlap.

5.3 Numerical convergence of the coupled scheme

We have previously discussed potential numerical issues for both major components of the scheme we are presenting in this Chapter. In this section we will discuss differences arising in the final coupled scheme due to the choice of the RT grid resolution. In Section 5.3.1 we will revisit the use of the local clumping factor as a tool for eliminating the effect of the RT grid on the evolution of the mean ionized fraction in our simulation boxes, which is related to the speed of I-fronts during a reionization run. Section 5.3.2 examines how the redshift of reionization of individual halos depends on the RT grid resolution. This dependence is a consequence of the way we choose when a halo is seen as ionized which we presented in the previous section. This may in turn have an effect on the efficiency of source suppression.

5.3.1 Local clumping factor in reionization runs

In this section, we return to the use of a locally computed clumping factor in order to include all available information about the density field from a given N-body simulation into the RT grid used. The RT grids we use in practise have a much lower resolution (in terms of the number of particles/Voronoi nuclei used) than the N-body simulations from which we extract the density field information. In Section 3.4.2 we introduced the clumping factor computed in individual RT cells and how it affects the global evolution of ionized fraction on static density fields. Now, we show how important it is in a reionization run using the method we described in Section 5.2.

We performed several coupled code runs based on the L20N512 N-body simulation (see Table 4.1). This is the same simulation which supplied the snapshot used in Chapter 3 tests, Figs. 3.16 and 3.17. There, we used only a single output at $z = 5$ as the static density field and used an arbitrarily chosen simulation time of $\Delta t = 1$ Myr. The runs we are presenting here include the evolving density field (i.e. the ionization field is evolved using more than a single snapshot as explained in 5.2) and the simulation time is given by the time corresponding to the snapshot redshifts for our chosen cosmology[7]. We use ≈ 50 snapshots between $z \approx 25 - 5$, motivated by our findings in Section 4.5.2. The

[7] Again, we use the same cosmology as the Millennium simulation (Springel et al., 2005b): $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_b = 0.045$, $h = 0.73$, $n_s = 1.0$ and $\sigma_8 = 0.9$.

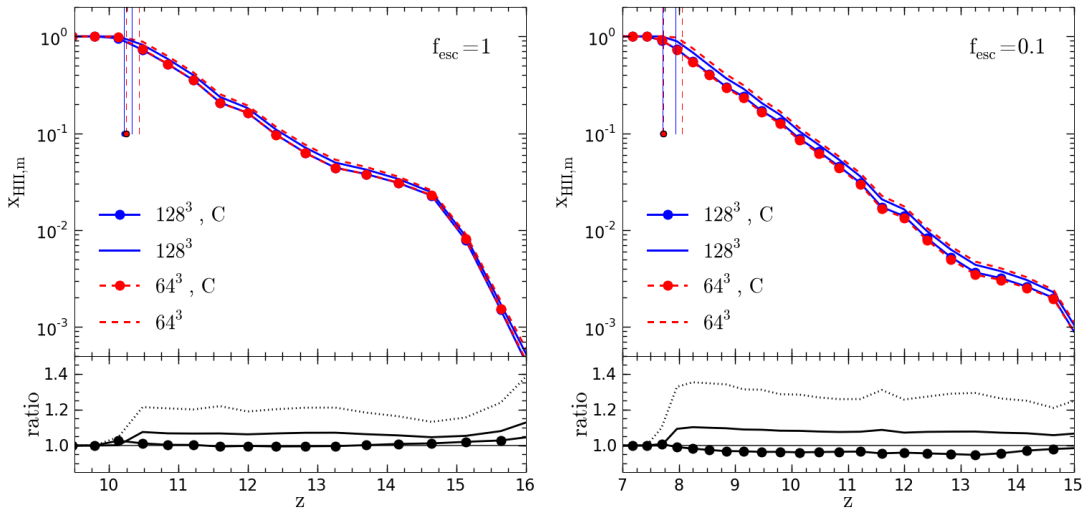

Figure 5.3: The mean mass-weighted ionized fraction as a function of redshift with respect to the RT grid resolution (number of Voronoi cells) and the use of the local clumping factor. We assume $f_{\rm esc} = 1$ (left panels) or 0.1 (right panels). The RT grids compared have 64^3 (red dashed lines) and 128^3 cells (blue solid lines). The runs which include local clumping factor are marked with circles (and C in legend). Bottom panels give ratios of $x_{\rm HII,m}$ on 64^3 and 128^3 grids for: runs with C (solid line with circles), without C (solid line) and 64^3 without and 128^3 with C (dotted line). See text for details and discussion.

L20N512 simulation has approximately the same mass resolution as L50N1024 (see Table 4.1) that we will use for the discussion in later parts of this chapter. The latter run is 1/8th volume and the same mass resolution as the Millennium-II run (Boylan-Kolchin et al., 2009) that we will use in our future work. We expect that the effects of the RT resolution we find in L20N512 runs to be a good representation of the behaviour in larger boxes just mentioned. The source model assumed is the default Baugh05. Finally, in order to isolate the RT grid resolution effects, source suppression is not included.

Figure 5.3 shows the evolution of mean mass-weighted ionized fraction in runs with varying RT grid resolution and use of the local clumping factor. We compare the runs that assume that every photon produced by galaxies escapes to the IGM ($f_{\rm esc} = 1$) and a more realistic escape fraction of $f_{\rm esc} = 0.1$ (as assumed in Chapter 4). The main difference

between the two values is the duration of the reionization process, which correlates with the relative importance of recombinations. The differences between runs are subtle, but measurable and are best seen in a direct comparison between different runs. The bottom panels give the ratios of two different resolution runs, $\mathrm{ratio}(z) = x_{\mathrm{HII},64^3}(z)/x_{\mathrm{HII},128^3}(z)$: lines with circles compare runs with local C, solid lines runs without local C and dotted lines give the full difference due to the representation of the density field by comparing 64^3 run without with a 128^3 one with local C. For $f_{\mathrm{esc}} = 1$ (left panels), different grid resolution runs that employ the local clumping factor have practically identical evolution of $x_{\mathrm{HII,m}}$, resulting in the same redshift of reionization (which we define here as redshift at which $x_{\mathrm{HII,m}} = 0.9$, marked by vertical lines). Without C, the 64^3 resolution grid has about 10% faster I-fronts than the 128^3 one, which results in $\Delta z \approx 0.1$ earlier reionization as we defined it here (compare blue solid and red dashed vertical lines). When a more realistic $f_{\mathrm{esc}} = 0.1$ is used (right panel), the differences between different RT resolution runs are only slightly larger. First, the runs with local C show a modest difference, with 64^3 run being $\approx 5\%$ slower than the 128^3 one. Without C, the ratio goes in the opposite direction meaning that the low resolution run is a few percent faster relative to high resolution one than when $f_{\mathrm{esc}} = 1$. The relative difference in the redshift of reionization between pairs of different RT grid resolution runs also remains unaffected. Note also the clear extension of reionization when assuming lower f_{esc}. For $f_{\mathrm{esc}} = 1$, $x_{\mathrm{HII,m}}$ goes from 10^{-3} to 1 from $z \sim 16 - 10$, which for the assumed cosmology corresponds to $\Delta t \approx 230$ Myr. By comparison, for $f_{\mathrm{esc}} = 0.1$, the redshift range is $z \sim 15 - 8$ leading to $\Delta t \approx 380$ Myr, i.e. a $\approx 60\%$ longer duration of reionization.

In this instance, extending the process of reionization (by using a lower f_{esc}) did not clearly expose the RT grid issues. This result points to the relatively small importance of recombinations in the current setup. To test this assumption we modify the runs by including all N-body particles in the estimate of both the cell densities and local clumping factors (i.e. setting $\Delta_{\mathrm{thresh}} = \infty$, see Section 3.4.2). This results in a significant increase in the number of recombinations in the high density regions where dark matter halos reside. Figure 5.4 shows how this change affects the previously discussed runs. Again, the use of local C clearly improves convergence in the evolution of $x_{\mathrm{HII,m}}$ with respect to RT grid resolution. While the run with $f_{\mathrm{esc}} = 1$ (left panels) shows more or less the same ratios as seen in Fig. 5.3, the run with $f_{\mathrm{esc}} = 0.1$ (right panels) shows more significant

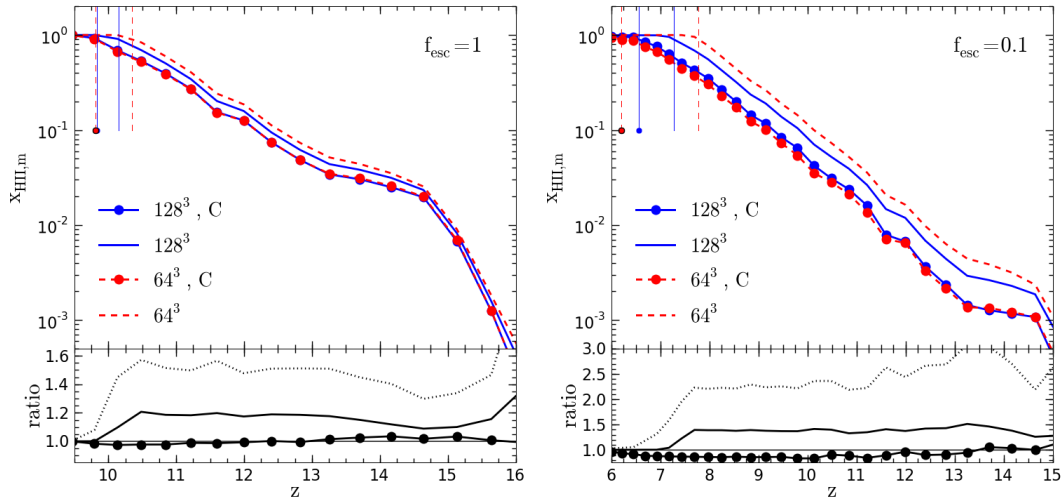

Figure 5.4: Same as Fig. 5.3, but with all particles used to compute local clumping factors ($\Delta_{\rm thresh} = \infty$, see Section 3.4.2). The much stronger recombination rates expose the effect of RT resolution more clearly. Even in this extreme case, the use of local C significantly improves convergence (compare ratios of runs with (line with circles) and without (solid line) C in bottom right panel).

differences. Again, the use of local C results in somewhat overestimated recombination rates on the 64^3 grid, as evidenced by $\approx 15\%$ lower fractions when compared to the 128^3 run (line with circles in the bottom panel). This results in a more noticeable difference in redshift of reionization between the two runs, $\Delta z \approx 0.25$. While not perfect, this relation is still better than when the clumping factor is not used: the low resolution run has $\approx 50\%$ higher fractions resulting in $\Delta z \approx 0.5$ earlier reionization than in the higher resolution case.

To back up our previous claim that the representation of the density field becomes more important for more extended reionization scenarios, we must quantify the relative importance of recombinations. This can be done by comparing the number of LC photons produced per hydrogen atom up to some redshift (a value we used extensively in Chapter 4), with the redshift of reionization obtained from a full RT reionization run. As we stated in Chapter 4, if no photon sinks were present, the IGM would be fully ionized when one photon per H atom reaches the IGM. In the runs we are currently discussing,

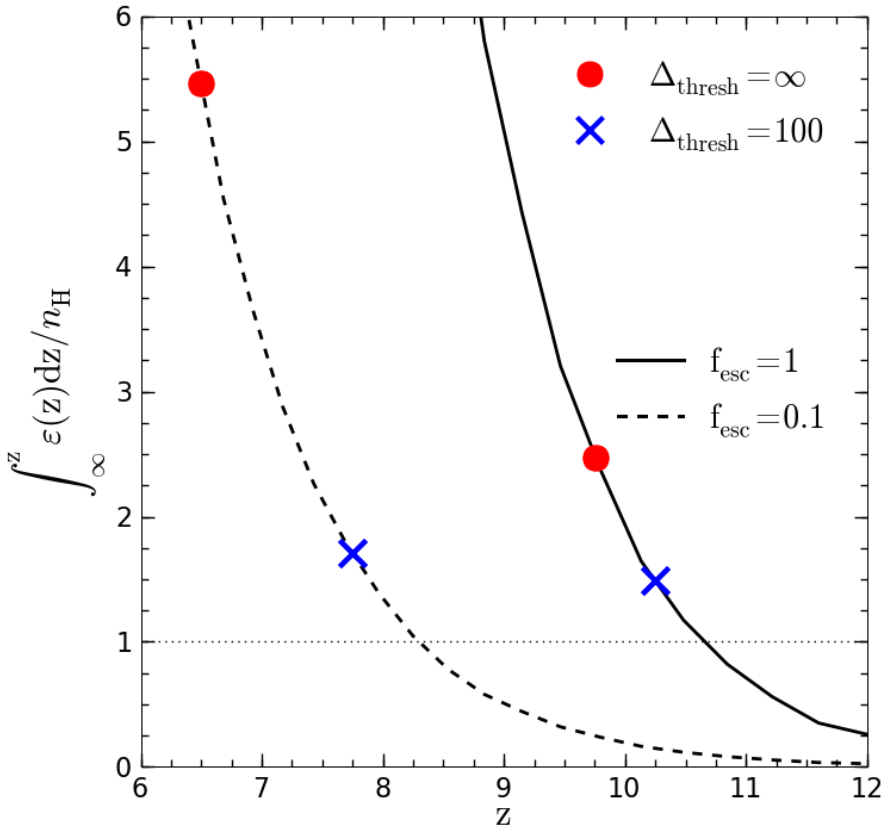

Figure 5.5: The importance of recombinations with respect to the escape fraction and density field representation. Lines give the number of LC photons per H atom emitted into the IGM by redshift z using the Baugh05 model in the L20N512 box for $f_{\rm esc} = 1$ (solid line) and 0.1 (dashed line). The symbols mark the number of LC photons produced by the redshift of reionization ($x_{\rm HII,m} = 0.9$) of the 128^3 runs with local C presented in Figs. 5.3 and 5.4 (blue lines with circles in those figures): red circles mark runs that include all N-body particles in the density field ($\Delta_{\rm thresh} = \infty$) and blue crosses the ones that exclude halo particles ($\Delta_{\rm thresh} = 100$). See text for discussion.

recombinations are the only photon sinks. As all the runs have the same emissivities (since we do not include source suppression), we can gauge the importance of recombinations for a given run by seeing how many LC photons have been produced by the redshift of reionization for that run. This is what we present in Fig. 5.5. Here, we show

how many ionizing photons were produced by GALFORM galaxies by the redshift of reionization of 128^3 runs with local C presented in Figs. 5.3 and 5.4 (blue lines with circles). Our default setup, which excludes halo particles by assuming $\Delta_{\rm thresh} = 100$, (blue crosses in Fig. 5.5) requires not even a whole extra photon beyond the first to finish reionization: if $f_{\rm esc} = 1$ (solid line) reionization is completed when ≈ 1.5 photons per H atom are produced, while for $f_{\rm esc} = 0.1$, ≈ 1.7 photons are needed. As nothing else except $f_{\rm esc}$ is different between the two runs, the importance of recombinations increases with a more extended time of reionization. This is seen much more clearly when including halo particles in the density field ($\Delta_{\rm thresh} = \infty$, red circles): for $f_{\rm esc} = 1$, ≈ 2.5 photons per atom are needed while for $f_{\rm esc} = 0.1$ the ratio must be ≈ 5.5 to complete reionization. Obviously, the effect of the IGM "clumpiness" on the number of recombinations during reionization is enhanced by the duration of the reionization era. This result may seem counter-intuitive as the mean recombination rate in the Universe is expected to decrease with time[8]. One would therefore expect that the number of photons per H atom needed to complete reionization to stay, in the worst case, constant. The opposite behaviour, as seen in Fig. 5.5, is caused by the lower speed of I-fronts in more extended reionization scenarios: slower I-fronts remain longer in the high density regions where the recombinations are most effective therefore increasing their importance. This feature cannot be captured in analytical models that obtain the recombination rate from a universal mean value (e.g. Benson et al., 2006), as RT calculations are necessary to compute the correct I-front speeds in the cosmological density field.

We have shown that the use of the local clumping factor introduced in Section 3.4.2 provides a simple method for reducing the dependence of $x_{\rm HII,m}$ evolution on the grid cell size. The next section will deal with another issue for our reionization modeling scheme related to the RT grid resolution: the redshift (or time) at which halos are found to be in ionized regions.

[8]The recombination rate is $\propto C\langle n_{\rm H}\rangle^2$, where C is the clumping factor of the Universe and $\langle n_{\rm H}\rangle$ the mean density of hydrogen. The mean density depends on redshift as $\propto (1+z)^3$ i.e. it decreases with decreasing z. On the other hand, C increases with decreasing redshift (due to the growth of structure), but not as strongly: $C(z) \propto (1+z)^{-1.5}$ for the global value of C in Fig. 3.10. As the recombination rate depends on the square of the density, it should clearly decrease with decreasing redshift.

5.3.2 When are halos found in ionized regions?

As we stated in Section 5.2, a halo is considered to be "ionized" (found to be in an ionized region for applying photo-ionization feedback rules of Section 4.2.7) when it is found to be in a Voronoi RT cell with $x_{\rm HII} \geq 0.7$ at the redshift corresponding to a snapshot output (sources are static during the RT step that evolves the ionization field between snapshots i and $i+1$). The results obtained with this criterion clearly depend on the RT grid resolution, i.e. the volumes of individual cells. As an example, consider a galaxy-hosting halo found somewhere in a Voronoi cell of some volume V and number density $n_{\rm H}$, both of which we assume to be constant. Let us further assume that there is no incoming LC flux, i.e. that the galaxies in the said halo ionize the cell by themselves. This is indeed what is happening to the earliest galaxies as we will discuss in more detail shortly. If the galaxy has a constant LC luminosity L, then it will be marked as ionized after $\Delta t = 0.7 n_{\rm H} V/L$, i.e. the time from when the source is first turned on to when it ionizes 70% of its host cell. Now assume the spatial resolution of the RT grid is increased by splitting each grid cell into 8 equal volume cells, $V' = V/8$. If we take the mass to be uniformly distributed within volume V then the finer resolution cell will have the same number density, $n_{\rm H}{}' = n_{\rm H}$. The luminosity of the galaxy hosted by the halo is independent from the RT grid and remains L. On the new grid, the time it would take for the galaxy to sufficiently ionize its host cell to be marked as ionized is $\Delta t' = 0.7 n_{\rm H} V'/L = \Delta t/8$, i.e. it will be marked as ionized 8 times faster. The argument is identical if one assumes that the cell is ionized by external flux or a combination of internal and external sources. This issue is unavoidable unless one employs some kind of a sub-cell prescription to ensure that the time Δt remains constant with respect to RT grid resolution changes. We are therefore interested in measuring how big of an effect the change of RT grid resolution may have on the redshifts of reionization of individual halos and whether this will have any effect on photo-ionization feedback.

We compare two runs presented in the previous section: the L20N512 N-body simulation with halo particles removed ($\Delta_{\rm thresh} = 100$), with the local clumping factor included and $f_{\rm esc} = 1$, default Baugh05 GALFORM model and with RT grid resolution of 64^3 and 128^3 grid cells, respectively. These two runs have already been presented in the left panel of Fig. 5.3 (red and blue lines with circles, respectively). There we have shown that the evolution of $x_{\rm HII,m}$ in those two runs is in excellent agreement, which guarantees

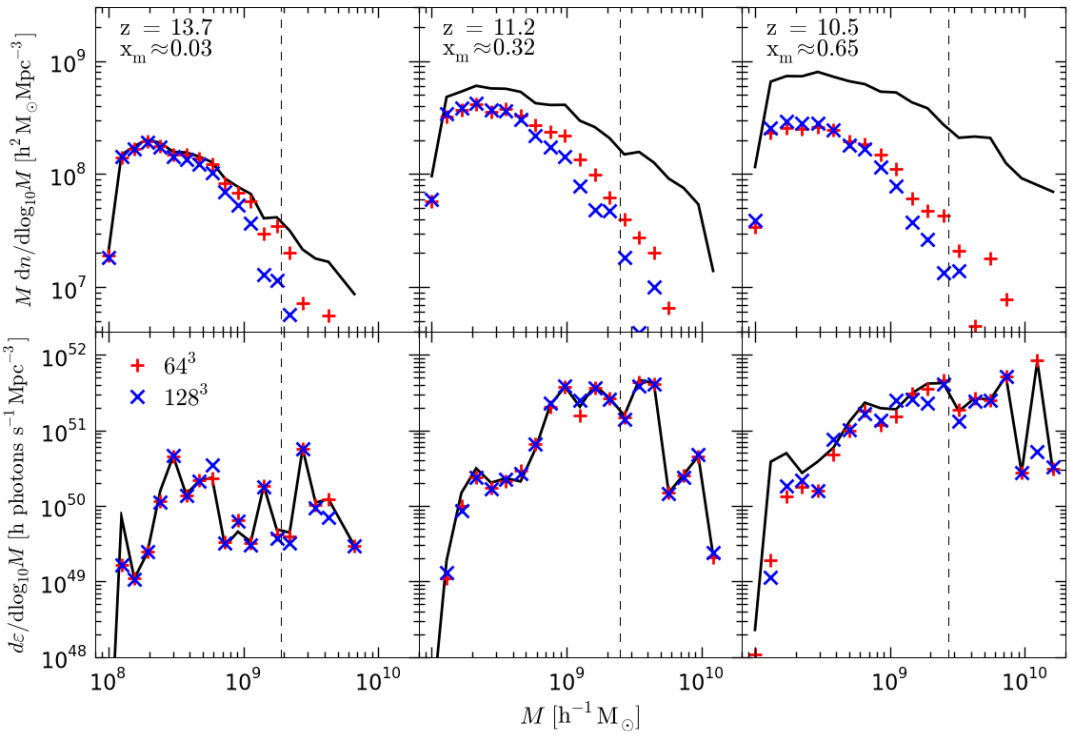

Figure 5.6: *Top panels:* Halo mass function of neutral halos at z =13.7 (left), 11.2 (centre) and 10.5 (right). The redshifts were chosen to correspond to different stages of the reionization process, where the mean mass-weighted ionized fraction is $x_{\rm HII,m} \approx 0.03$ (left), 0.32 (centre) and 0.65 (right). Black solid line marks the MF of all halos, while the coloured symbols give the MF of halos marked as neutral using 64^3 (red pluses) and 128^3 (blue crosses) RT grids. *Bottom panels:* The LC emissivity as a function of halo mass at the same redshifts as the top panels. The solid line gives the value of the run without source suppression, while coloured symbols give the values with source suppression on the 64^3 and 128^3 RT grids (colour and symbol style the same as in top panels). See text for discussion.

that any differences in ionization times between the runs we find are not caused by the difference in the global ionized fraction.

Figure 5.6 shows when halos of different masses become subject to photo-ionizing feedback. The top panels plot the halo mass function (MF) of *neutral* halos, the ones not found to be in ionized regions. This MF is compared to the MF of *all* halos, either marked

as neutral or ionized. The bottom panels show the LC emissivity that corresponds to those halos and how it is affected by the photo-ionization feedback applied to ionized halos ($V_{\rm H} \leq V_{\rm cut}$ as discussed in Section 4.2.7). The vertical dashed lines mark the halo mass corresponding to $V_{\rm cut}$ at the given redshift. At earlier redshifts when only a small fraction of the simulation box is ionized ($x_{\rm m} \approx 0.03$, left panels), we see that only some of the most massive halos are located in ionized regions. These are the halos whose progenitors hosted the first galaxies at $z \gtrsim 20$, that produced the first HII regions thus ionizing their own surroundings. Smaller, more recently formed halos remain unaffected by ionizing radiation, either due to not forming their own HII regions or because HII regions around more massive halos are not large enough to significantly ionize their neighbours. The emissivity at this redshift (bottom left panel) is quite noisy due to the small number of halos found in the box. At these redshifts, the MF evolves quickly, e.g. the number density of halos grows tenfold between z =13.7 and 11.2 (corresponding to $\Delta t \approx 100$ Myr). As the mean ionized fraction grows (centre and right top panels), some of the smallest halos are found in ionized regions, though the emphasis is still on more massive halos. By $z = 10.5$, practically all halos with mass $M \gtrsim 3 \times 10^9 {\rm M}_\odot/h$ are marked as ionized, while about 1 in 10 halos with $M \lesssim 10^9 {\rm M}_\odot/h$ remains neutral. The difference between the two RT grid resolutions is mostly noticeable in the high mass halo range, where the higher resolution grid results in more halos being marked as ionized. This is a good example of the problem that we described at the beginning of this section, as the finer RT grid resolution allows halos to be found to be ionized earlier than if the lower resolution one were used. Still, we note that the largest discrepancy between the two RT grid resolution runs is in the highest mass range, where $V_{\rm H} > V_{\rm cut}$. Those halos are unaffected by photo-ionizing feedback, thus we expect the emissivity from that halo mass range to remain unaffected irrespective of how early they were ionized. This is indeed seen in the bottom panels of Fig. 5.6, which shows the LC emissivity to be almost completely unaffected by source suppression due to the ionized regions. The only effect is seen at $z = 10.5$ (bottom right panel), where there is evidence for some suppression of the smallest halos, $M \lesssim 10^8 {\rm M}_\odot/h$. Note that the percieved suppression in the high mass, in the $M \approx 10^{10} {\rm M}_\odot/h$ bin in the bottom right panel of Fig. 5.6, is caused by the small number of halos in that range, i.e. a single halo can significantly skew the result. This relatively weak effect of photo-ionizing feedback is a feature introduced by

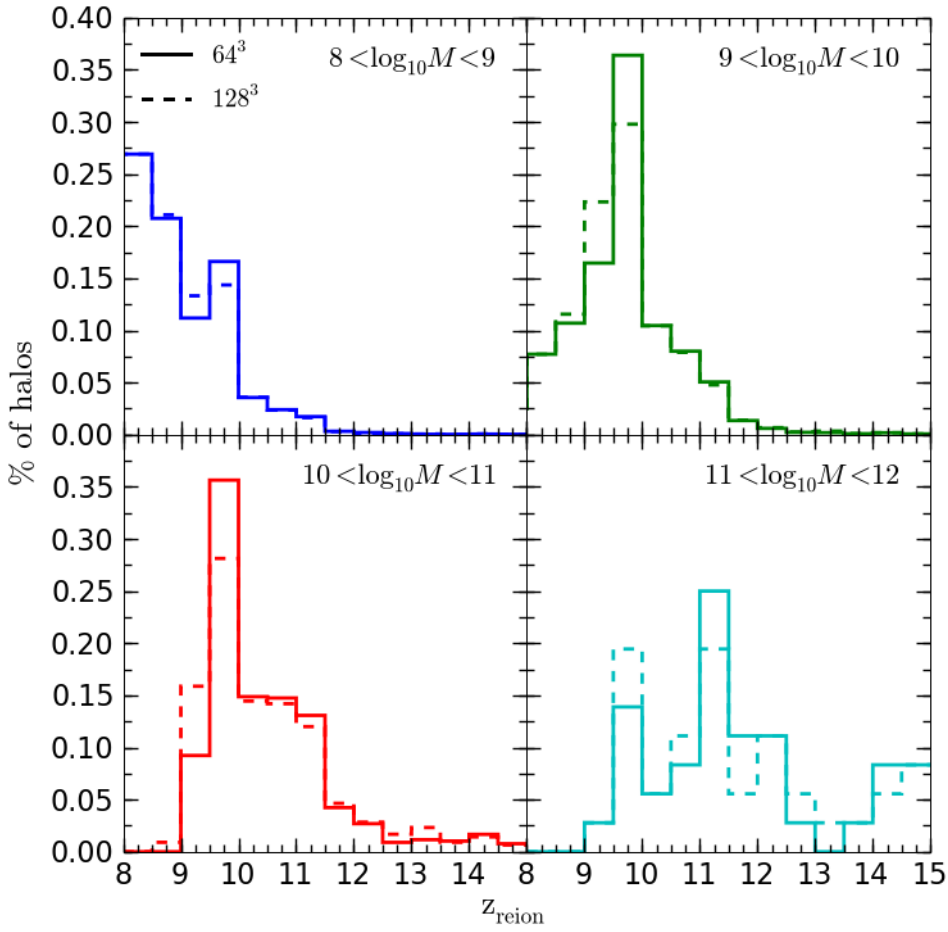

Figure 5.7: The distributions of redshifts of reionization, $z_{\rm reion}$, for halos of different masses at $z = 5$ as a function of RT grid resolution. For each halo, we define $z_{\rm reion}$ as the redshift at which the main branch first becomes ionized. The considered halo masses increase from the top to bottom and left to right (see mass ranges in panels; mass is quoted in $\rm M_\odot/\mathit{h}$ units). The choice of RT grid resolution (solid lines for 64^3 and dashed lines for 128^3) does not affect the scatter or peaks in $z_{\rm reion}$ i.e. does not change the general conclusion about individual halo reionization histories.

the delayed suppression as it is implemented in GALFORM, which we discuss in more detail in Section 5.4.3.

A different way of looking at individual halo reionization histories is to ask the ques-

tion when were the halos selected at a given time reionized as a function of their assembly histories and environment. The main question is whether halos of a given mass were reionized by their own, internal sources or by external ones (e.g. Weinmann et al., 2007; Alvarez et al., 2009). This, in turn, may have important consequences on present day observables, such as the predicted number of satellites of our own galaxy (e.g. Busha et al., 2009; Munoz et al., 2009). Our N-body simulations end at $z = 5$, so we discuss these issues for halos selected at that redshift. Figure 5.7 shows the distribution of reionization redshifts, $z_{\rm reion}$, of different mass halos found at $z = 5$ (the merger tree "roots") with respect to the resolution of the RT grid used to propagate the ionization field. We define $z_{\rm reion}$ as the redshift at which the main branch[9] of the merger tree *first* becomes ionized (we do not take into account that halos may become neutral again at some later time, as previously mentioned). Alternative definitions are also possible, e.g. $z_{\rm reion}$ may be the redshift at which more than 50% of the mass in all progenitors is found to be ionized. The exact choice of the $z_{\rm reion}$ definition does not affect the conclusions of the discussion. The first thing to notice is the position of the $z_{\rm reion}$ peak for different mass ranges. Remember that, for these runs, reionization is globally complete ($x_{\rm HII,m} \approx 1$) for $z \approx 10$ (Fig. 5.3). Practically all halos that at $z = 5$ have mass between 10^8 and $10^9 {\rm M}_\odot/h$ have been marked as ionized after the end of reionization. This is because the low mass halos at $z = 5$ have been formed (i.e. have the earliest progenitors) after $z \approx 10$. Clearly, the halo formation time bias is a factor in the discussion of $z_{\rm reion}$. Regardless, two features are clear: halos more massive at the final time have earlier $z_{\rm reion}$ on average, and the scatter in $z_{\rm reion}$, while large for all mass ranges, also grows with the mass of the final halo. More than 50% of halos more massive than $10^{11} {\rm M}_\odot/h$ (bottom right panel) have been ionized before $x_{\rm HII,m} = 0.5$, with $\approx 10\%$ of halos being ionized before $x_{\rm HII,m} = 0.01$ ($z \gtrsim 15$), which strongly implies that these halos were reionized by internal sources. We will discuss the physical implications of this result in our future work - for now, we are interested in how the $z_{\rm reion}$ distribution is affected by the resolution of the RT grid used to follow the ionizing radiation. As Fig. 5.7 demonstrates, it has a small effect on both

[9]The main branch of the merger tree is defined by following the main progenitor of the root halo to its formation redshift. The main progenitor of a halo is usually the most massive direct progenitor (halo "A"). The exception is made in the case of a less massive direct progenitor ("B") that in turn has a direct progenitor ("C") that is more massive than "A". Then "B" is selected as the main progenitor and followed further (see e.g. Genel et al., 2009).

the position of the $z_{\rm reion}$ peaks and the scatter. As these two are the most important values for the statistical description of halo reionization redshifts, we can conclude that our future discussion of this issue will not be strongly affected by the resolution of the RT grids.

For an even closer look, we directly compare $z_{\rm reion}$ for each halo at $z = 5$ found using 64^3 and 128^3 RT grids in Fig. 5.8. We find that the majority of the final redshift halos are found to be ionized at the same redshift on both RT grids, as evidenced by the median value (thick dashed line) and the small 50% range of values (red error bars). For the minority of halos that do have different $z_{\rm reion}$, they tend to be ionized earlier on higher resolution grids, as we expected from the earlier description of the problem. Note that the effect of the grid resolution was largest for the most massive halos in Fig. 5.7, which are the ones that are reionized earliest. Figure 5.8 allows us to conclude that the choice of RT resolution does not severely affect the reionization histories of halos.

5.4 Results

In this section, we discuss some of the preliminary results we obtained with the coupled scheme. The discussion will be focused on a run that makes use of the L50N1024 N-body density field (see Table 4.1) with the source model assumed to be the default Baugh05, but without a global $z_{\rm cut}$. The gas is assumed to trace the dark matter and to be composed of hydrogen only. All the photons are at the Lyman limit frequency. The temperature of the ionized gas is automatically set to 10^4 K when estimating recombinations. The simulation box is periodic, so the photon packets that leave the box return on the other side as an approxiamtion of the external flux. The limitations of the density field make our results incomplete: first, the mass resolution is not sufficient to resolve all of the relevant ionizing radiation sources (see discussion in Section 4.5.1; see red line in Fig. 4.25) and, second, the density field does not provide the correct representation of the IGM, lacking the effects of the gas pressure and the presence of minihalos and Lyman-limit systems (see Section 3.4.1). While these limitations mean that we cannot claim to have a complete representation of the reionization process, we can still discuss general trends. In Section 5.4.1 we will take a closer look at the topology of reionization in our scheme and at how the ionized fraction evolves as a function of local overdensity. In

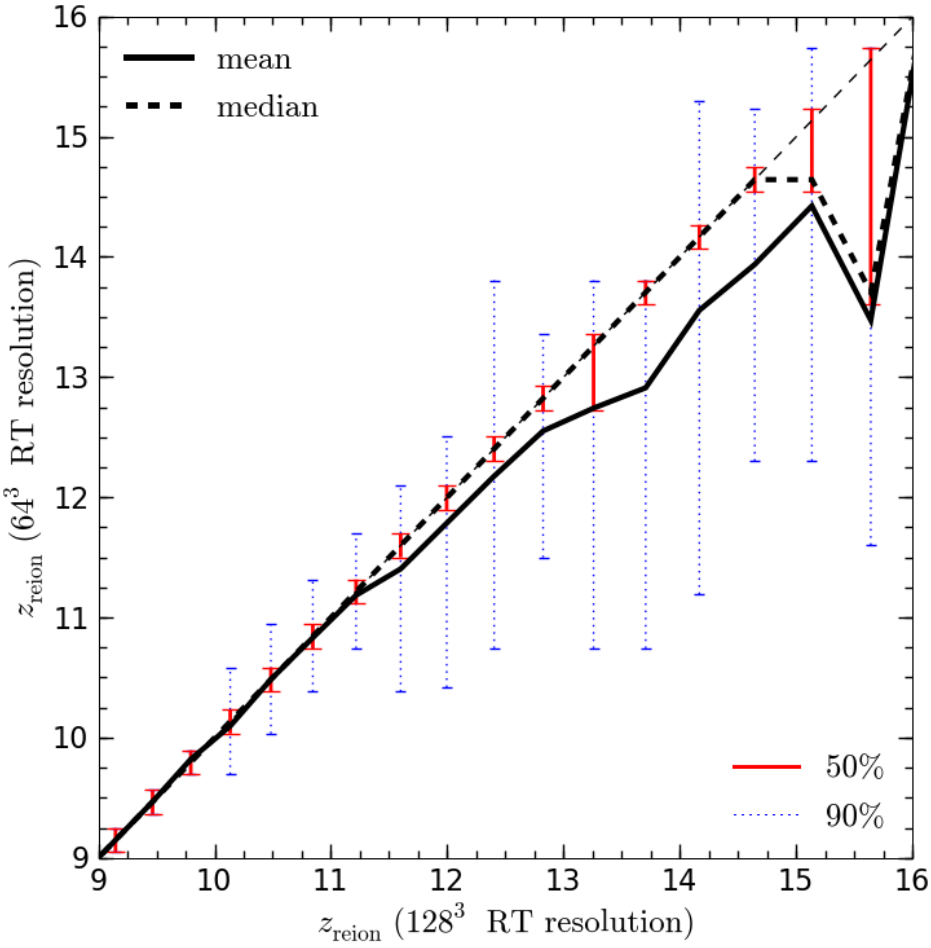

Figure 5.8: The value of $z_{\rm reion}$ found for each L20N512 halo at $z = 5$ as a function of RT grid resolution. If the RT grid had no effect on $z_{\rm reion}$, the value found on both 64^3 (y-axis) and 128^3 (x-axis) grids would be the same (thin dashed line). In practise, we find that $z_{\rm reion}$ for the majority of the halos is the same on both grid resolutions (median, thick dashed line and 50% range of values, red error bars). The halos that do have significantly different $z_{\rm reion}$, are more likely to be ionized earlier in the high resolution grid, especially at $z \gtrsim 14$ (mean, thick solid line and 90% range of values, dashed blue error bars).

Section 5.4.2, we discuss the distribution of HII region sizes in more quantitative terms and compare the results from the Baugh05 sources (which, we remind the reader, are characterised by large LC photon production during starbursts) with other prescriptions

previously used in the literature. Finally, in Section 5.4.3 we discuss the effect of source suppression on the course of the reionization process. Table 5.4 gives details about the different runs we employ during the following discussion and introduces shorthand labels for each.

5.4.1 Topology of reionization

We begin the discussion by showing the general progression of reionization in Fig. 5.9. The figure shows slices through the hydrogen density (grey-scale background) and ionization field (colour contours) in 4 steps of evolving global volume-weighted ionized fraction, as well as the positions of sources at the same redshifts (coloured circles). To give an idea of where the most luminous sources are, we mark the top 10% luminous halos with lime green circles, the rest are white. The density field on the uniform grid was constructed using nearest grid point interpolation from the N-body data. The ionization field is mapped directly from the Voronoi tessellation to a higher resolution uniform grid (256^3 cell in this case) by finding in which Voronoi cell a uniform grid cell resides and assigning it the ionized fraction of that Voronoi cell. In this way we preserve the exact volume distribution of the ionization field as represented by the Voronoi grid. Note that the chosen uniform grid resolution in this procedure is not sufficient to resolve the smallest Voronoi cells located in highest density regions, but such small scales are not important once the HII regions grow to include more than a single cell.

Figure 5.9 shows the general progression of reionization in all our runs. Early on ($x_{\rm HII,V} = 0.1$, top left panel), the ionization field consists of individual HII regions around the highest density regions that host the radiation sources. As the number of sources grows, so do the HII regions. Eventually, they begin to overlap (top right and bottom left panels), resulting in almost all of the simulation box being covered by a single HII region (bottom right panel) before it becomes completely ionized. Notice that the I-front thickness (given by the difference between red and blue contours) is almost always about one Voronoi cell (average Voronoi cell length scale is $\Delta l \approx 0.4\,{\rm Mpc/h}$), suggesting that the recombination rate correction[10] we introduced in Section 3.5 is generally applicable. The insides of HII regions are highly ionized ($x_{\rm HII} > 0.9$), although

[10]We remind the reader that the correction entails interpreting the ionized fraction of the cell as the fraction of the cell volume that is fully ionized.

Run	N-body*	$f_{\rm esc}$	Suppress	Details
D	L50N1024	1	no	default simulation setup
D1	L50N1024	1	no	same as D, except no local C in recombinations
D2	L50N1024	1	no	same as D, except densities and local C computed assuming $\Delta_{\rm thresh} = \infty$
ML1	L50N1024	1	no	source model as in Eq. (5.4), normalised to give same ϵ as Baugh05 at $z = 10.13$
ML1f	L50N1024	0.1	no	
ML2	L50N1024	1	no	source model as in Eq. (5.5), normalised to give same ϵ as Baugh05 at $z = 10.13$
DS	L50N1024	1	no	same as D, but no scatter in mass-luminosity relation; see text
S	L20N1024	1	no	default simulation setup
S1	L20N1024	1	yes	same as S, but with GALFORM suppression
S2	L20N1024	1	yes	same as S, but with extreme suppression (see text)
Sf	L20N1024	0.1	no	
S1f	L20N1024	0.1	yes	
S2f	L20N1024	0.1	yes	extreme suppression

Table 5.1: Coupled SimpleX-GALFORM reionization runs discussed in Chapter 5. First column gives a shorthand label for the run, second the label of the N-body run used to construct the RT density field and halo merger trees (see Table 4.1 for details of N-body runs), third gives the assumed escape fraction of LC photons, fourth states whether source suppression is considered and last column gives distinguishing details of each run. All runs use 128^3 Voronoi cells for RT, constructed over a random subset of N-body particles (see Section 3.3). All runs assume the default Baugh05 model for source LC luminosities (without global $z_{\rm cut}$), except where stated otherwise.

* All density fields are constructed with assumed $\Delta_{\rm thresh} = 100$ and the local clumping factor is used when evaluating recombinations (see Section 3.4.2) unless stated otherwise in the simulation details.

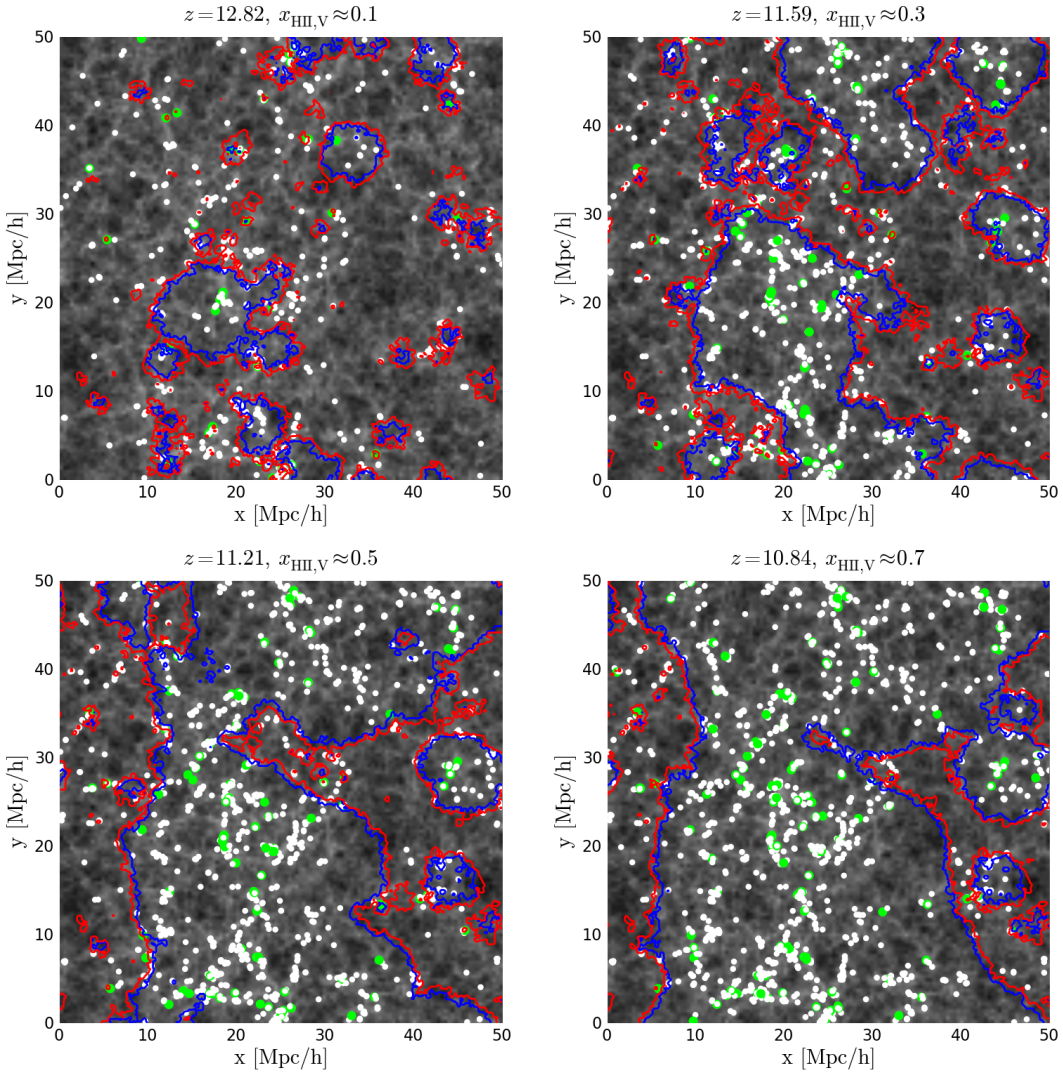

Figure 5.9: Topology of the HII regions during reionization in run D at 4 stages of reionization ($x_{\rm HII,V} \approx 0.1$, 0.3, 0.5 and 0.7, top to bottom left to right). The grey-scale background is a slice through the middle of the gas overdensity field ($\log_{10}\Delta$) with thickness $\Delta l = 0.4\,{\rm Mpc/h}$. Coloured contours mark the boundaries of regions with volume-weighted ionized fractions larger than 0.9 (blue) and 0.1 (red). Coloured circles mark the positions of sources in the slice, lime green for the top 10% most luminous sources and white for the other sources. See text for discussion.

there are a few "islands" of lower ionization in some high density regions. The ionized fraction in the highly ionized regions (the trace neutral gas, $x_{\rm HI} \lesssim 10^{-3}$) is also generally uniform, which is a feature of the SimpleX implementation we use, specifically of the

numerical diffusion caused by the violation of radiation isotropy mentioned in Section 3.2.2. If the isotropy is enforced as in the improvement introduced by Paardekooper et al. (2009), more structure in the ionization field within HII regions can be seen (i.e. somewhat higher neutral fractions in filaments), but this in no way affects the positions of the I-fronts themselves (i.e. the transition over $x_{\rm HII} = 0.5$ remains in the same Voronoi cells, even though their neutral fractions may change slightly).

The topology of reionization can be characterised by the ratio of the mass and volume-weighted ionized fractions, $x_{\rm HII,m}$ and $x_{\rm HII,V}$, within the simulation volume. To see why, let us first remember how we define these fractions:

$$x_{\rm HII,m} \equiv \sum_i x_{\rm HII,i} m_i / \sum_i m_i \tag{5.1}$$

$$x_{\rm HII,V} \equiv \sum_i x_{\rm HII,i} V_i / \sum_i V_i, \tag{5.2}$$

where index i indicates individual Voronoi cells with volume V_i, mass m_i and ionized fraction $x_{\rm HII,i}$. The ratio $x_{\rm HII,m}/x_{\rm HII,V}$ is then (e.g. Iliev et al., 2006a):

$$\begin{aligned} x_{\rm HII,m}/x_{\rm HII,V} &= \frac{\sum_i x_{\rm HII,i} m_i / \sum_i m_i}{\sum_i x_{\rm HII,i} V_i / \sum_i V_i} \\ &= \frac{M_{\rm ionized}/M_{\rm total}}{V_{\rm ionized}/V_{\rm total}} \\ &= \frac{1}{\langle\rho\rangle} \frac{M_{\rm ionized}}{V_{\rm ionized}}, \end{aligned} \tag{5.3}$$

where $M_{\rm total}$ and $V_{\rm total}$ are the total mass and the volume of the simulation box, while $M_{\rm ionized}$ and $V_{\rm ionized}$ are the mass and volume of all ionized gas in the box. Clearly, the ratio $x_{\rm HII,m}/x_{\rm HII,V}$ gives the mean density of ionized gas in units of the mean total gas density $\langle\rho\rangle \equiv M_{\rm total}/V_{\rm total}$. Consequently, when $x_{\rm HII,m}/x_{\rm HII,V} > 1$, the overdense regions ($\Delta > 1$) are preferentially ionized before underdense ones ($\Delta < 1$), pointing to the inside-out nature of the reionization process. The opposite is true when $x_{\rm HII,m}/x_{\rm HII,V} < 1$, suggesting outside-in reionization.

Figure 5.10 shows the evolution of $x_{\rm HII,V}$ with redshift (left panel) and the corresponding evolution of $x_{\rm HII,m}/x_{\rm HII,V}$. The issue of the reionization topology is rooted in the role of recombinations in the evolution of the ionization field, so we compare our default run D (blue solid lines) with two variations: D1 (black dash-dotted lines) is the same as D, but does not include the local clumping factor (Section 3.4.2) in the recombination estimates and D2 (red dashed lines) uses all particles ($\Delta_{\rm thresh} = \infty$) in the

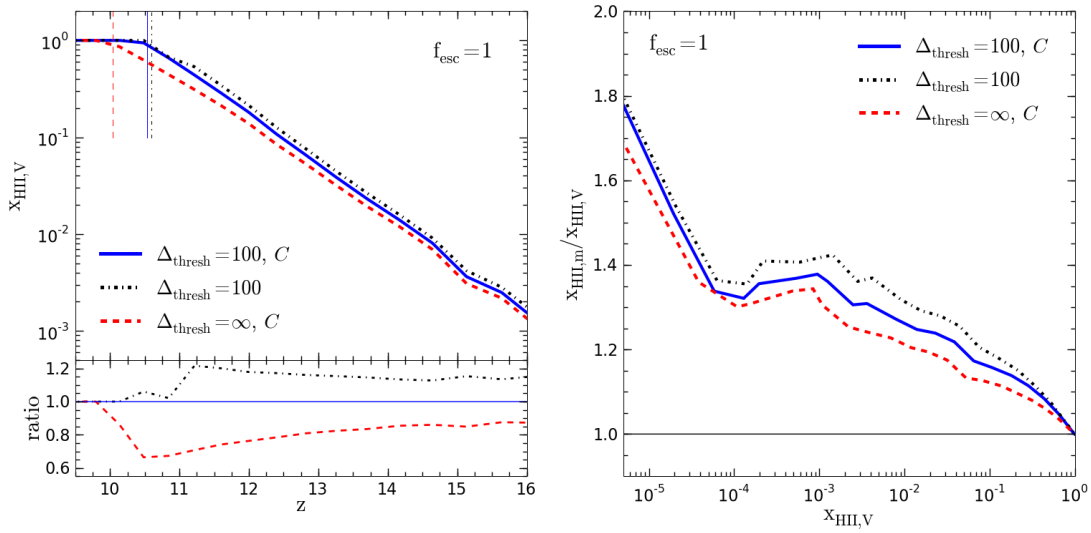

Figure 5.10: *Left:* The evolution of the volume-weighted ionized fraction $x_{\rm HII,V}$ in D (blue solid line), D1(black dash-dotted line) and D2 (red dashed line) with redshift. Bottom panel shows ratios of $x_{\rm HII,V}$ of D1/D (black dash-dotted) and D2/D (red dashed line). *Right:* The ratio $x_{\rm HII,m}/x_{\rm HII,V}$ for the same runs, compared at equal $x_{\rm HII,V}$. The mass-weighted ionized fraction is always higher than the volume-weighted one, even for the D2 run which has a significantly higher recombination rates (as seen by the significant delay in the redshift of reionization, $\Delta z \approx 0.5$, in the left panel). We find that reionization proceeds strictly inside-out.

computation of cell densities and local clumping factors. Both D1 and D2 use the same sources and Voronoi nuclei (so the cell *volumes* are exactly the same) as D. The evolution of $x_{\rm HII,V}$ with redshift (left panel) is similar for the two runs, D and D1, as a result of the relatively low values of the local clumping factors when the halo particles are excluded from this N-body run. Including all particles in the density field makes recombinations much stronger in the overdense regions that host halos. This in turn results in slower I-fronts and a delay in the redshift of reionization, as evidenced by the results of the D2 run ($\Delta z \approx 0.5$). Still, in all runs the mass-weighted ionized fraction remains higher than the volume-weighted one (right panel), suggesting a strictly inside-out reionization regardless of the way the density field is constructed. While not shown here, we find the

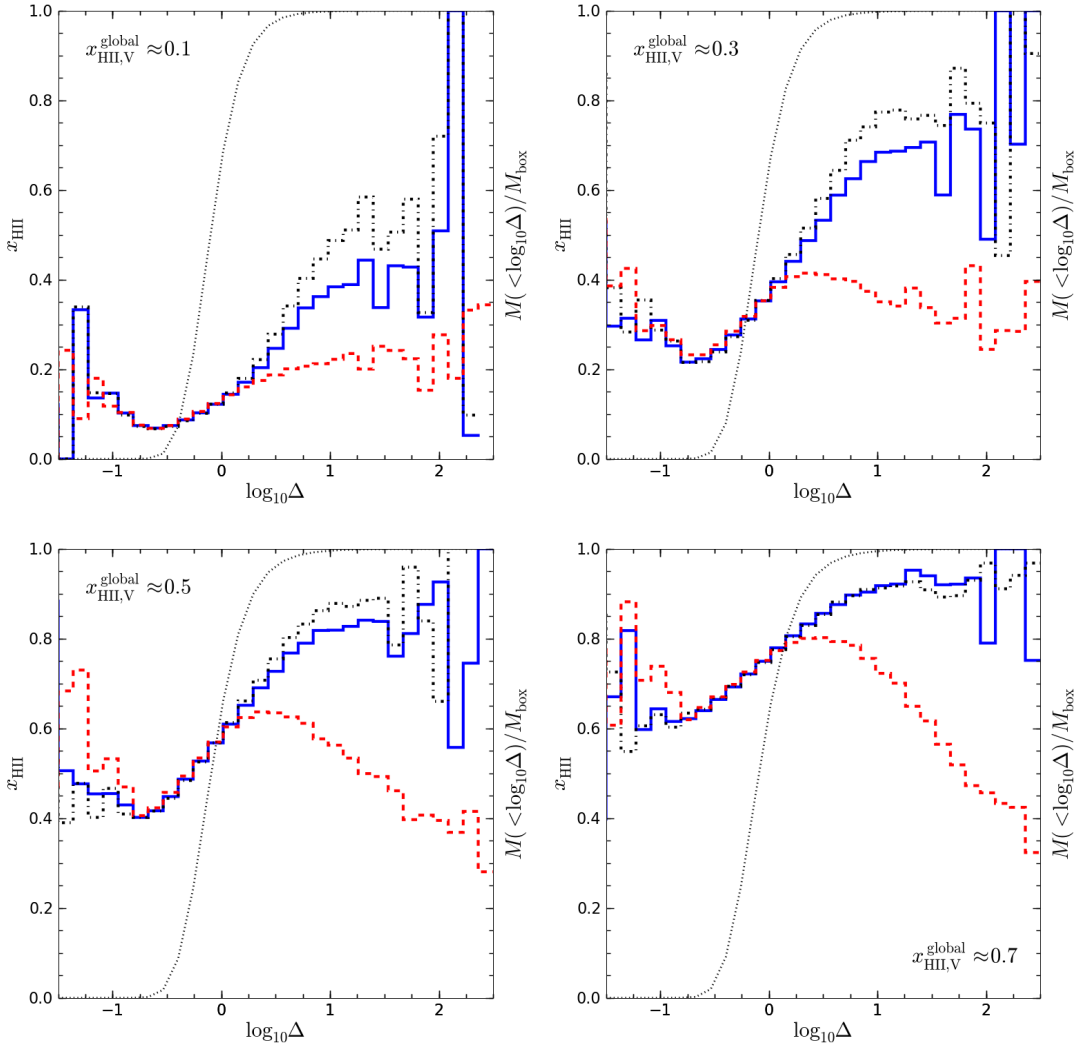

Figure 5.11: The ionized fraction as a function of cell overdensity for runs D (blue solid), D1 (black dash-dotted) and D2 (red dashed lines). Panels show the results at different values of the ionized fraction computed over the whole box, $x_{\rm HII,V}^{\rm global}$, corresponding to the slices shown in Fig. 5.9. The thin dotted line shows the fraction of total mass found in cells less overdense than $\log_{10}\Delta$. Different density fields result in very different values of $x_{\rm HII,V}$ in more overdense regions ($\log_{10}\Delta > 1$), but this does not affect the global mean because only a small fraction of mass ($\lesssim 5\%$) is found in those regions. See text for discussion.

same to be true if we extend reionization further by using $f_{\rm esc} = 0.1$.

A more detailed look at how the ionized fraction depends on overdensity is provided

in Fig. 5.11. Note that the data range in the D and D1 runs extends beyond $\Delta = 100$, even though that is the overdensity at which we truncated the N-body density field data. This is due to the difference between SPH and Voronoi tessellation density estimations. Still, the fraction of cells with $\log_{10}\Delta > 2$ is very small, which is also the source of the large scatter in the values of $x_{\rm HII}$ in those overdensity bins. The D2 run density data was not truncated before Voronoi grid construction, resulting in the highest overdensity being $\log_{10}\Delta \sim 3.5$, but here we only show the overdensity range that is resolved by all runs. The biggest difference between the density fields is due to the recombination rate (a combination of density and clumping factor) in the overdense regions ($\Delta > 1$), clearly seen in Fig. 5.11. Not using local clumping factors in D1 results in slightly more ionized overdense regions than in D, with the difference larger at low global ionized fractions (top left panel) than at higher ones (bottom left panel), and eventually disappearing when the ionized fraction approaches unity (e.g. bottom right panel). This is due to the increase with time in the number of sources in these regions, which results in LC intensities capable of overpowering recombinations. A rather different result is obtained when halo particles are kept in the density field, as in D2 run. Even at early stages of reionization ($x_{\rm HII,V}^{\rm global} \approx 0.1$, top left panel), the overdense regions are significantly less ionized in D2 than in D. As the global ionized fraction increases, the difference between the two increases further, especially in significantly overdense regions ($\log_{10}\Delta \gtrsim 1$). This behaviour is also due to the growth of structure that we invoked to explain the decreasing difference between D and D1. While the LC intensity rises in the overdense regions in D2 in the same way as in D and D1, the growth of structure also results in the increase of the recombination rate, which overpowers the former effect to keep the most overdense regions mostly neutral. Note that we expected the cells hosting halos to have extremely high recombination rates and used this fact as one of the arguments against keeping halos in the density field estimate (Section 3.4.1). We therefore consider the results of the D2 run to be caused by an incorrect density field representation, but present them here none the less to illustrate the effect the recombination rate has on the reionization topology.

Recombinations are responsible for the depression in $x_{\rm HII}$ at $\log_{10}\Delta \approx -0.5$ seen most prominently at earlier stages of reionization (top panels). Early on, the LC intensity reaching the voids is low and time-variable. For example, the voids are first ionized by

HII regions around starbursts which results in a sharp increase in the ionized fraction due to a short "spike" of LC intensity. Once the starburst has ended, the recombination rate is the only regulator of the ionized fraction, so its dependence on density produces a depression such as the one seen in Fig. 5.11. At later stages of reionization (bottom panels), the increase in the number of sources means that the low density regions are exposed to more constant and stronger LC intensities resulting in a less pronounced depression. These results are in general agreement with Choudhury et al. (2008), who developed a semi-analytic method to explicitly study the issue of reionization topology. The main distinction between our results and theirs is that they find that the low density regions are fully ionized before the mean density ones, i.e. they find a more pronounced depression located around $\Delta \approx 1$. The cause of this difference may be their approximate treatment of RT, which lacks the effects of shielding by higher density regions, as well as the their simplified treatment of radiation sources. We will perform a more detailed comparison between the two models in our future work.

Our finding that the reionization proceeds strictly inside-out is in agreement with most recent numerical simulations of large-scale reionization (e.g. Iliev et al., 2006a; McQuinn et al., 2007; Zahn et al., 2007). However, there is still no consensus on this issue as a number of semi-analytic (e.g. Miralda-Escude et al., 2000; Furlanetto et al., 2004) and numerical simulations (e.g. Finlator et al., 2009a) report finding final parts of reionization history proceeding in the outside-in fashion. The way the density field is represented differs greatly among these works, so it is still early in the development of the reionization studies to claim any final answer.

5.4.2 Typical sizes of HII regions during reionization

While the discussion presented in the previous section provides a significant insight into how the ionization field evolves, there is a need for a more detailed understanding of the sizes of HII regions during reionization and their dependence on the various parameters assumed in a model. The distribution of HII region sizes is an observable quantity of the reionization process. Direct observations will require observatories still in development, such as SKA, but the related 21cm power spectrum (e.g. Furlanetto et al., 2006; Mellema et al., 2006) may become available in the next few years from telescopes such as LOFAR and MWA. As our models are incomplete at this stage, we do not attempt to make pre-

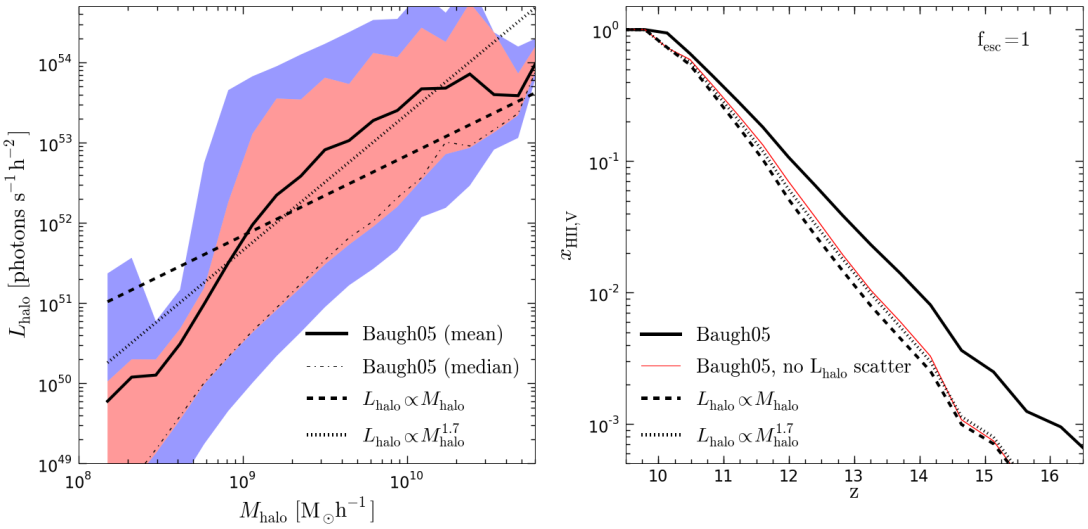

Figure 5.12: *Left:* The dependence of halo LC luminosity of on halo mass $M_{\rm halo}$ at $z = 10.13$ in D, ML1 (dashed line) and ML2 (dotted line) runs. The solid and dot-dashed lines mark the mean and the median value in the D run (default Baugh05 model). The red and blue shaded areas mark the 50% and 90% ranges, respectively. *Right:* Evolution of the mass-weighted mean ionized fraction with redshift in runs D (thick solid), ML1 (dashed), ML2 (dotted) and DS (thin red line). Even though the emissivities are equal between the runs (to within $\approx 1\%$), the evolution of $x_{\rm HII,m}$ still shows significant differences. The scatter in the $L_{\rm halo} - M_{\rm halo}$ relation in D is the most important factor. See text for discussion.

dictions for the future 21cm observations. Nevertheless, we are interested to see whether the detailed source model we employ here (default Baugh05) can be distinguished from simpler models used in the literature. To this end, our comparison will be with the work of McQuinn et al. (2007), who examined in detail the morphology of HII regions during hydrogen reionization with respect to different components of the reionization model.

We compare the results of our default simulation, D, with two other runs that assume simpler and less physically motivated source models. First, in ML1 run, we assume that the LC luminosity of each halo, $L_{\rm halo}$, is directly proportional to the halo mass, $M_{\rm halo}$:

$$L_{\rm halo} = A_1 \left(\frac{M_{\rm halo}}{{\rm M}_\odot/h} \right), \tag{5.4}$$

where A_1 is a constant with dimensions of LC luminosity. This relation corresponds to McQuinn et al. (2007) S1 simulation and is also used in Iliev et al. (2006a) and their subsequent work (we used their relation in Chapter 3, see Eq. 3.15). The second model, ML2 assumes a steeper slope of the mass-luminosity relation:

$$L_{\rm halo} = A_2 \left(\frac{M_{\rm halo}}{{\rm M}_\odot/h} \right)^{1.7}, \tag{5.5}$$

where A_2 is a different constant with the same dimensions as A_1. Ideally, the HII regions in different models should be compared at the same redshift and the same ionized fraction (e.g. Lidz et al., 2008) to minimise the effect of the underlying density field evolution on the results. To achieve this, we chose constants A_1 and A_2 so that the total emissivity in ML1 and ML2 is the same as the emissivity in D at $z = 10.13$. As the mean mass-luminosity relation in the default Baugh05 does not significantly evolve (see Fig. 4.2), this normalization should be appropriate for all redshifts. We compared the emissivities between the runs at all redshifts and found them to differ by no more than a percent. We could easily match the emissivity exactly by having time-dependent normalization coefficients, but chose to keep them constant to better match the models of McQuinn et al. (2007). The minimum mass for atomic cooling in a halo, $M_{\rm halo}(T_{\rm vir} = 10^4{\rm K})$, is below the mass resolution of the L50N1024 run (see Fig. 4.25), so we do not apply any low-mass cutoffs.

The left panel of Fig. 5.12 shows the mass-luminosity relations in the D, ML1 and ML2 runs. When compared to the mean value of $L_{\rm halo}$ vs. $M_{\rm halo}$ in D, both ML1 and ML2 have higher luminosities in the low mass range, $M \lesssim 10^9{\rm M}_\odot/h$, with ML1 also significantly higher than ML2. Consequently, ML2 has higher luminosities than ML1 in halos with $M \gtrsim \times 10^9{\rm M}_\odot/h$, but still lower than D until $M \gtrsim 10^{10}{\rm M}_\odot/h$. Of course, the Baugh05 values of $L_{\rm halo}$ exhibit a large scatter for a given $M_{\rm halo}$ (discussed in 4.3.3; shown by the shaded areas in Fig. 5.12), meaning that a small fraction of halos produces 1-2 orders of magnitude more LC photons than their counterparts in ML1 and ML2. The right panel gives a first glimpse of the importance of this scatter, by plotting the redshift evolution of the mass-weighted mean ionized fraction in the three models. Even though the total emissivities are almost exactly the same, the D run shows a large difference when compared to the ML1 and ML2. While all runs complete reionization at approximately the same time, the D run reaches lower ionized fractions much earlier than the

other two. This is entirely due to the speed with which I-fronts leave the highest density regions (itself a function of source luminosity, Eq. (2.23)), as we argued earlier. The same effect is seen, to a smaller extent, in the difference between the evolution in ML1 and ML2, with the latter being slightly faster. To make sure that the small number of highly luminous sources (i.e. scatter due to starbursts) is responsible for different evolution of $x_{\rm HII,m}$ in D, we performed another run, labelled DS, where all the halos of mass $M_{\rm halo}$ have an $L_{\rm halo}$ that corresponds to the mean of the Baugh05 model (solid line in the left panel), i.e. we removed the large scatter while preserving the mean emissivity and shape of the mass-luminosity relation of D. The evolution of $x_{\rm HII,m}$ in DS is shown with a thin red line in the right panel of Fig. 5.12 and it matches the evolution of ML1 and ML2 much more closely than D. We therefore expect the scatter in $L_{\rm halo}$ to be the chief distinguishing factor of the Baugh05 model in terms of the HII region morphology.

We compare the HII region morphologies visually in Fig. 5.13. The ionization fields are compared at 4 values of $x_{\rm HII,V} \approx 0.1, 0.3, 0.5$ and 0.7 presented in top to bottom rows, respectively. The mean ionized fractions are not exactly the same due to the limited number of outputs and different evolutions of $x_{HII,V}$ (right panel of Fig. 5.12), but the differences between all models are in the $\Delta x_{\rm HII,V} \pm 0.02$ range around the value of the D run. The ionization field maps were produced using the same procedure as for the contours in Fig. 5.9 described in the previous section. For increased clarity, we divide the ionization field into ionized (black) and neutral (white) regions. The boundary between the two is set at $x_{\rm HII} = 0.7$, the same ionized fraction above which we apply source suppression (see Section 5.2).

In all runs presented, the ionized bubbles generally trace the distribution of overdensity, which results in similar large scale features in all ionization fields. The main difference is the number and sizes of the HII regions emanating from the smallest halos. Halos with $M \lesssim 10^9 {\rm M}_\odot/h$ are most luminous in the ML1 run (Fig. 5.12), which translates to the highest number of small ionized bubbles in that run (left column). The mass-luminosity relation is steeper in ML2, which results in fewer small bubbles and more pronounced larger ones. This trend is continued in D, where the ionization field at lower $x_{\rm HII,V}$ is dominated by a handful of large HII regions, with even fewer small ones.

A more quantitative comparison of HII bubble sizes is presented in Fig. 5.14. To facilitate comparison, we adopted the McQuinn et al. (2007) technique for estimating

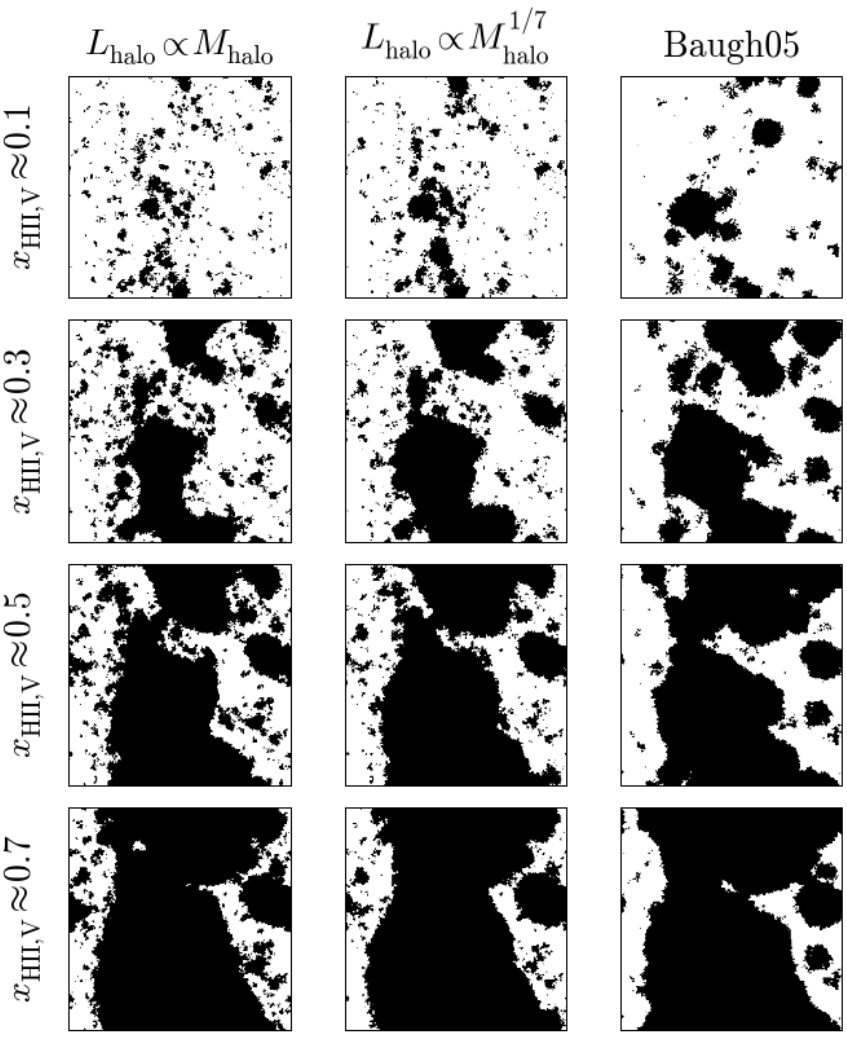

Figure 5.13: Comparison of ionization fields from three simulations with different source models: ML1 (left column), ML2 (middle column) and D (right column). All panels show slices through the middle of the box, with side length 50 Mpc/h. Black regions are ionized ($x_{\rm HII} \geq 0.7$), while white are neutral ($x_{\rm HII} < 0.7$). The redshifts corresponding to each $x_{\rm HII,V}$ are the same as in Fig. 5.9. The distinguishing difference between the runs is the number of smallest HII regions, which is highest in ML1 (left column). On the other hand, such small HII regions are almost non-existent in D (right column).

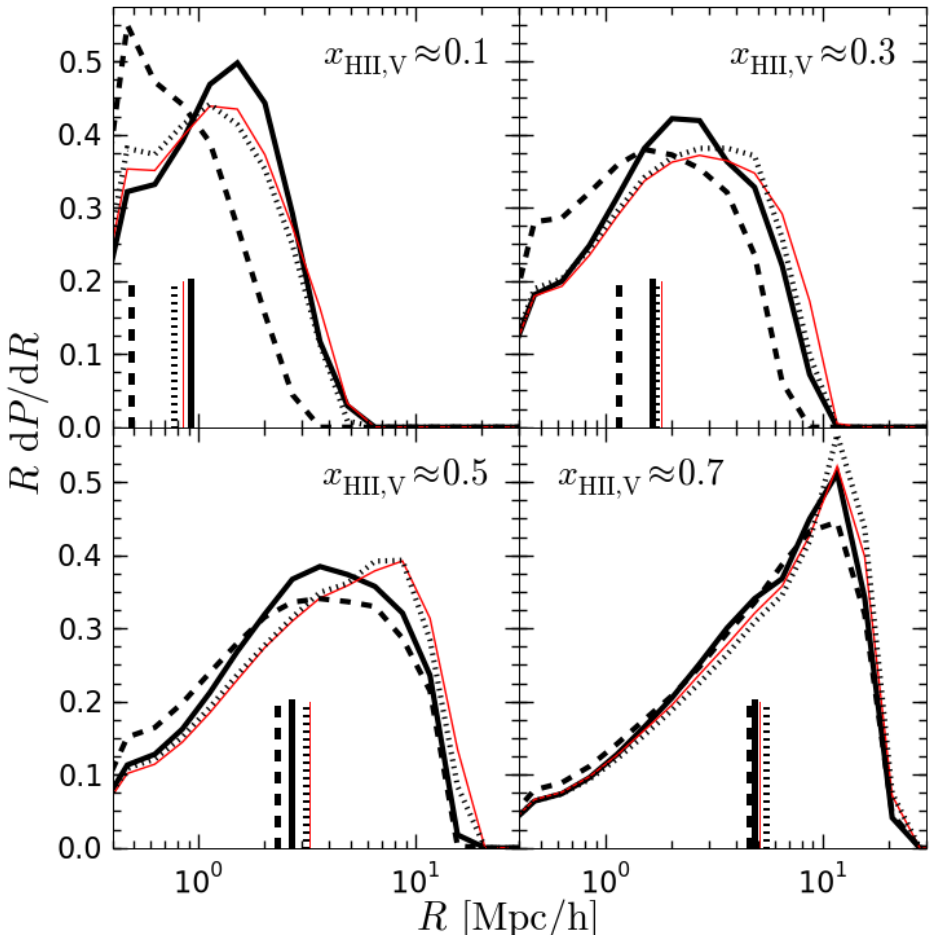

Figure 5.14: The bubble radius PDFs of runs D (thick solid lines), ML1 (dashed lines), ML2 (dotted lines) and DS (thin red lines) at $x_{\rm HII,V} \approx 0.1, 0.3, 0.5$ and 0.7 (top to bottom, left to right panels). See text for the definition of the bubble radius. Vertical lines give the value of the distribution median for each run, which may be used as a quantifier of the typical bubble size. Run ML1 clearly has smaller typical bubble size than the rest. Starbursts in D do not produce larger HII bubbles, but *more* of the large ones. Once scatter due to bursts is removed (DS run), the Baugh05 model bubble sizes closely match those of ML2.

bubble sizes, who in turn introduced it in order to compare results with semi-analytic models (Furlanetto et al., 2004). The procedure is as follows. For each pixel of the ionization field, we find the largest sphere centred on that pixel which is on average 90%

ionized. If such a sphere is found, the pixel is said to be in a bubble of that sphere's radius R. Otherwise, the pixel is assumed to be in a neutral region. The non-zero radii of all pixels are then tabulated in equal $\log_{10}R$ bins in order to estimate the probability density function (PDF). This definition of the bubble size allows for a comparison independent of their shapes or merging history. It was originally introduced for comparison with semi-analytic models that define the bubble size in very similar manner (Furlanetto et al., 2004). Note that alternative definitions exist, such as the friends-of-friends algorith employed by Iliev et al. (2006a). The bubble size PDF provides information similar to that provided by the ionized region power spectra. We chose to discuss only the former method to ease the comparison with McQuinn et al. (2007).

The general evolution of the bubble radius PDF with $x_{\rm HII,V}$ shows the same progression of the reionization process we discussed previously. Early on in the process (e.g. $x_{\rm HII,V} \approx 0.1$, top left panel), there is a large number of small individual HII regions. The different source models result in a marked differences in the typical bubble sizes, represented in Fig. 5.14 with the median values (vertical lines, compare ML1 and ML2, dashed and dotted lines respectively). As reionization progresses, the typical bubble size increases and the distribution of sizes becomes more peaked. The latter effect is caused by the overlap of HII regions, where more and more small bubbles merge into larger ones. As we approach the final overlap ($x_{\rm HII,V} \approx 0.7$, bottom right panel), the different source models show closer agreement, which is expected when only a few large bubbles are present in the ionization field. A bigger simulation volume is needed to properly model the bubble sizes at this and higher $x_{\rm HII,V}$. However, even now we see that the ML1 simulation has slightly more small bubbles, which in turn results in a lower peak at typical bubble size. The bubble sizes in the DS run closely trace the ones in ML2, meaning that the non-power-law shape of the mean $L_{\rm halo}$ in the Baugh05 model (Fig. 5.12) does not play a very significant role. Finally, the results of the D run show a somewhat different behaviour: at $x_{\rm HII,V} \approx 0.1$ (top left panel) D run produces the largest bubbles, but also "more" of them (suggested by the width of the distribution) than in any of the other runs. As reionization progresses, the ML2 and DS runs start producing somewhat larger bubbles. On the other hand, the D run still exhibits the more strongly peaked PDF, but the typical bubble size (vertical line) is now smaller than in the previous two runs, yet still bigger than in the ML1 run (see top right and to a lesser extent bottom left panel).

This behaviour suggests that the typical bubble sizes in D may not dependent only on the growth of structure (i.e. the evolution of the halo mass function), as in the ML1 and ML2 runs, but also on some other properties. One possible explanation may be that the typical bubble size is given by the halo mass that produces most LC photons in merger-induced starbursts. We find the basis for that claim in the result we presented in Fig. 4.5, where we compared the default Baugh05 emissivity (black line) at $z \approx 10$ with the same model with no starbursts (red line). There, the increase in emissivity that the bursts introduce has a clear peak, almost two orders of magnitude for halos of $M_{\rm halo} \approx 10^9 {\rm M}_\odot/h$. We expect that the position of this peak is a combination of the number density of halos with that $M_{\rm halo}$ and the sufficiently common triggering of bursts by mergers occurring in such halos. Note that the growth of structure with redshift "flattens" the distribution of emissivity with a given mass, as seen in Fig 4.3, which removes to some extent a clearly distinguishable peak. Still, at the redshifts we are considering here, the emissivity peak is present and clearly not associated with the most massive halos present in the box. This effect may lead to the HII bubble morphology we observed in Fig. 5.14.

These features in morphology of the D run may be partially skewed by the inadequate mass resolution of L50N1024, which results in a decreased number of bursts in halos with $M \lesssim 10^9 {\rm M}_\odot/h$, evidenced by a lower scatter in $L_{\rm halo}$ in that mass range seen in the left panel of Fig. 5.12. To check this, we repeated the runs used in the previous discussion but now based them on the L20N1024 N-body density data. This run resolves all the halos down to the cooling limit and below (see blue line in Fig. 4.25), but in an approximately 15 times smaller simulation volume. The results of those runs, which we do not present here, generally show the same behaviour as the runs in L50N1024. Unfortunately, the box size in L20N1024 is too small to properly sample the bubble sizes since the ionization field is practically dominated by a single HII region at even the moderate ionized fractions, $x_{\rm HII,V} \lesssim 0.5$. Therefore, to properly examine the bubble morphology in the models with starbursts we will require a simulation box with side $\geq 100 {\rm Mpc/h}$, with the merger trees fully resolved down to the atomic cooling halo mass. The morphology results we presented here are therefore not conclusive and we will explore them in more detail in our future work.

Dependence of morphology on emissivity

Before concluding the current discussion on ionization field morphology, we must examine in more detail whether there is any dependence of bubble sizes on the absolute value of the LC emissivity. The runs which we employed in the previous section have been explicitly set to the same emissivity to avoid any such effects. When comparing certain assumptions in the source models, as we did in Chapter 4, the emissivity is necessarily affected, often more so than the shape of the mass-luminosity relation. Therefore, if we intend to compare morphologies of HII regions produced by different GALFORM models (e.g. changing the IMF in bursts in the Baugh05 model), we need to have an idea of how much the difference in emissivities affects the bubble size distribution.

To that end, in Fig. 5.15 we compare the results of the ML1 run with a run with exactly the same parameters except for a 10 times lower luminosity normalization: $A_1' = A_1/10$. We denote this run as ML1f, since the scaling down of the luminosities is equivalent to assuming an escape fraction of $f_{\rm esc} = 0.1$. The results, compared at a given $x_{\rm HII,V}$ value, clearly show that the simple change of total emissivity produces a large difference in the distribution of bubble sizes, with the ML1 run having a few Mpc/h larger characteristic HII region radii than ML1f. This is the same order of difference as we found between the source models i.e. between ML1 and ML2 in Fig. 5.14. This fact will clearly complicate attempts to observationally distinguish source models. We note that McQuinn et al. (2007) did not find this kind of dependence of bubble sizes on emissivity. They compared morphologies for runs with the same emissivities at different redshifts, as well as runs with differing emissivities at same redshifts (which is what we do here). The possible explanation for this apparent discrepancy may be less resolved density field in their runs: by default, their RT grid has a somewhat lower spatial resolution in the highest density regions than ours, lower mass resolution of the underlying N-body simulation[11] and no spatially dependent clumping factor to include the density field data not resolved by the RT grid. All these effects combine to make their density fields significantly smoother than the ones we employ here. Since the representation of the density field is the most probable cause of the different bubble sizes in ML1 and ML1f, this may

[11]They also use Gadget-2 to produce N-body runs with 1024^3 particles, but they use a somewhat larger simulation box: 65.6 Mpc/h, as opposed to our 50 Mpc/h. The different volumes result in our particle mass to be ≈ 2 times lower then theirs.

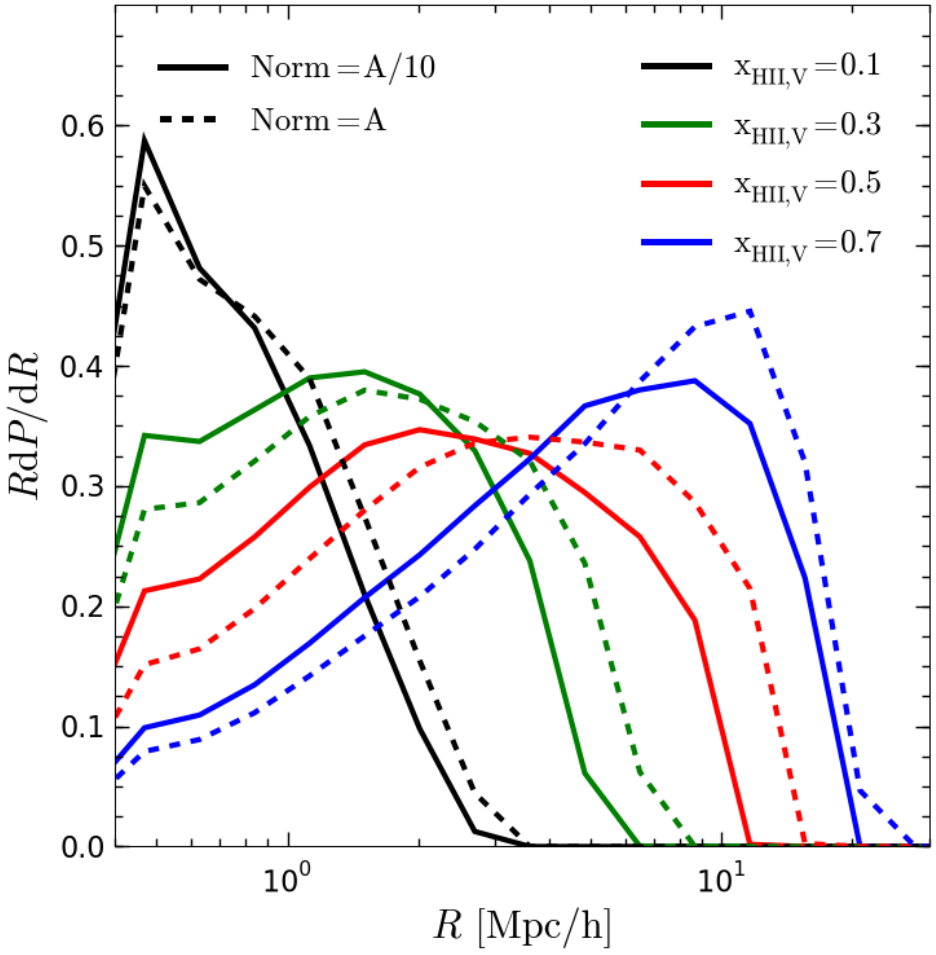

Figure 5.15: The comparison of bubble radius PDFs at different $x_{\rm HII,V}$ (values correspond to colors, given in the legend) in ML1 (dashed lines) and ML1f (solid lines) runs. The more extended reionization in ML1f due to lower emissivity results in significantly smaller bubble sizes. The importance of the absolute emissivity on bubble radii is comparable to the effect of the source model, shown in Fig. 5.14.

explain why McQuinn et al. did not obtain the same result. As our own density field is only a first approximation, we consider the result presented in Fig. 5.15 to be only preliminary. However, we do not expect the result to be negated by the use of a more detailed density field, as it will only introduce stronger spatially inhomogeneous photon sinks (e.g. minihalos), not the opposite.

5.4.3 Role of source suppression in the progression of reionization

The final effect we consider is the core feature of our coupled scheme: the suppression of LC luminosity from sources caused by photo-ionization feedback by an inhomogeneous ionization field. We have already introduced the method for including this effect in our scheme in Section 5.2. Here, we explore how the inclusion of source suppression affects the evolution of the mean ionized fraction. Previous large-scale numerical reionization simulations that include this effect are a few: Iliev et al. (2007) used a simple model where all halos with $M_{\rm halo} < 10^9 {\rm M}_{\odot}/h$ found in ionized regions are completely removed from the list of sources at each redshift step. They found that this feature can significantly extend the reionization process, since the most numerous small halos that begin the reionization process are subsequently suppressed and the final overlap is reached only when a sufficient number of more massive halos (which are not affected by the suppression) are formed. McQuinn et al. (2007) expanded slightly on Iliev et al. (2007) by testing the effect of time dependent suppression of the LC luminosity of affected halos, which is closer to the model we use here. They also found that source suppression can severely delay the epoch of reionization - $\Delta t \approx 200\,{\rm Myr}$ or $\Delta z \approx 2$, when assuming instantaneous source suppression, and $\Delta z \approx 1$ when the LC luminosity of suppressed sources decreases exponentially over $\tau_{SF} = 100$ Myr. They also found that the bubble size distribution remains unaffected by the source suppression. Note that both works assume a linear mass-luminosity relation, such as the one given by Eq. (5.4).

To examine the effect of source suppression, we performed a series of coupled SimpleX/GALFORM runs on the L20N1024 density field. We use a smaller box than in the previous discussion (L50N1024) in order to properly resolve all the small halos that are affected by the suppression. Simulation S denotes our reference simulation: it has exactly the same parameters as the D simulation discussed earlier, except for the different underlying N-body density and halo merger trees. The S run does not include suppression. On the other hand, S1 run differs from S by including GALFORM suppression i.e. by turning off gas cooling in halos with mass $M < M(z, V_{\rm cut})$ found to be in ionized regions (regions with $x_{HII} > 0.7$). For comparison with previous work, we include a run (denoted S2) with "extreme" suppression: all the halos with mass $M < M(z, V_{\rm cut})$ found in ionized regions are instantly removed from the list of sources (as in Iliev et al., 2007). In suppressed runs, we assume $V_{\rm cut} = 60\,{\rm km/s}$ (as in the default Baugh05 model). This,

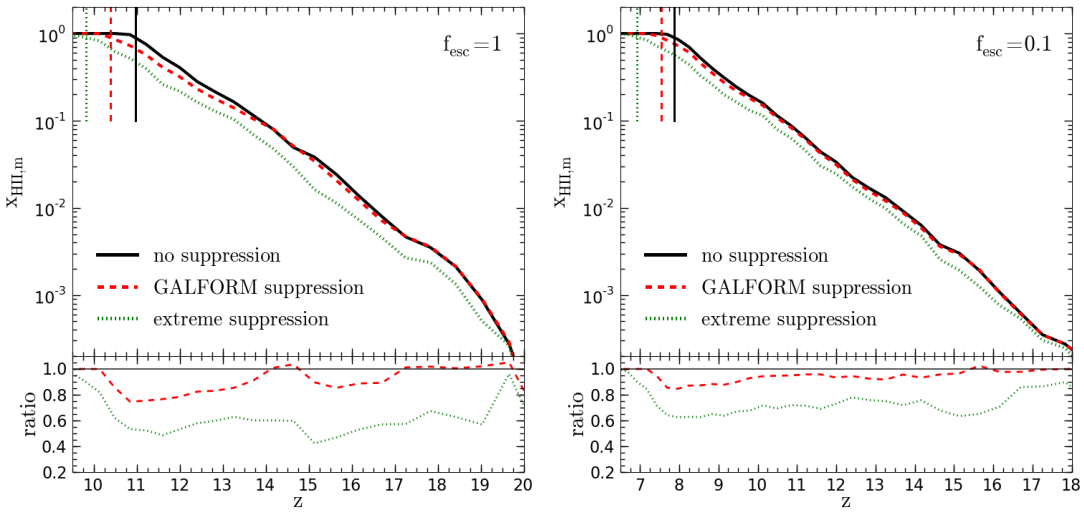

Figure 5.16: Evolution of $x_{\rm HII,m}$ with redshift in runs with variable source suppression prescriptions: S (solid black lines), S1 (dashed red lines) and S2 (dotted green line). Two values of $f_{\rm esc}$ are assumed: 1 (left panel) and 0.1 (right panel). The "extreme" suppression delays reionization $\approx$ 2 times more than the GALFORM prescription and affects the evolution of $x_{\rm HII,m}$ earlier on. Extending reionization by assuming a lower escape fraction (right panel) results in a smaller effect of suppression on $x_{\rm HII,m}$ as the reionization is competed by halos with circular velocities $V > V_{\rm cut}$ at lower redshifts.

for our choice of cosmological parameters, gives the following maximum mass affected by photo-ionizing feedback:

$$M(z, V_{\rm cut}) = 2.92 \times 10^9 {\rm M}_\odot / h \left(\frac{1+z}{11}\right)^{-3/2} \left(\frac{V_{\rm cut}}{60\,{\rm km/s}}\right)^3, \qquad (5.6)$$

where we assumed a matter-dominated Einstein-de Sitter Universe at high redshifts.

Figure 5.16 shows the evolution of $x_{\rm HII,m}$ in the S (solid), S1 (red dashed) and S2 (green dotted) runs (left panel). We compare these runs to ones with the same setup but $f_{\rm esc} = 0.1$ (labelled as Sf, S1f, S2f in Table 5.4; right panel of Fig. 5.16). At a first glance at the left panel, the extreme suppression has a much more pronounced signature: not only is the epoch of reionization delayed by $\Delta z \approx 1$ or $\Delta t \approx 77\,{\rm Myr}$ (a two times longer delay than the one produced by GALFORM suppression), but it also begins to affect the evolution of $x_{\rm HII,m}$ much earlier (at $z \approx 20$ as opposed to $z \approx 17$ for GALFORM sup-

pression). At these early stages of reionization (with low $x_{\rm HII,m}$), the sources are almost always self-ionized, so the latter effect is entirely caused by the previously discussed delayed nature of source suppression in GALFORM (e.g. right panel of Fig. 4.22). The same feature is responsible for the smaller delay in the epoch of reionization in the S1 run compared to S2. When the emissivity is decreased by assuming a lower escape fraction (right panel), the effect of suppression on the redshift evolution of $x_{\rm HII,m}$ becomes even less pronounced. Note, however, that in time units the epoch of reionization is actually some 50% longer when $f_{\rm esc} = 0.1$ (S2f) compared to $f_{\rm esc} = 1$ (S2). The importance of source suppression is decreased by the fact that, at lower redshifts, a larger fraction of LC photons is produced in halos with mass larger than the cutoff mass given in Eq. (5.6). When $f_{\rm esc} = 1$, $x_{\rm HII,m} \approx 0.5$ at $z \approx 12$, at which redshift more than 90% of LC photons are produced in halos with $M < M(z, V_{\rm cut})$. On the other hand, for $f_{\rm esc} = 0.1$ the same $x_{\rm HII,m}$ is reached at $z \approx 8.5$ when only 65% photons is produced in suppressible halos.

Note that the effect of source suppression we found here will only be decreased if the box size is increased, as this will increase the contribution of massive non-suppressible halos that are not found in this volume. This means that, in general, we find a much more modest effect of source suppression than Iliev et al. (2007) and McQuinn et al. (2007). The explanation for the difference lies in the assumed slope of the mass-luminosity relation. Both Iliev et al. and McQuinn et al. assumed a linear mass-luminosity relation, Eq. (5.4) in their models, which makes the low mass halos much more efficient in the production of LC photons than the relation predicted by Baugh05, as seen in Fig. 5.12. As a lower $f_{\rm esc}$ is more physically plausible, we expect the source suppression to have only a small effect on the progression of reionization in our future models as well. Also, the value $V_{\rm cut} = 60\,{\rm km/s}$ is almost certainly too high - a value of $V_{\rm cut} \approx 30\,{\rm km/s}$ is currently thought to be more appropriate based on gas dynamics simulations (e.g. Okamoto et al., 2008). Equation (5.6) shows that this value will result in 8 times lower maximum suppressible halo mass than we considered in this discussion, which will render the source suppression practically irrelevant for the reionization process. This, of course, does not mean that the inhomogeneous source suppression will not have any effect on lower redshift galaxy formation, which is a problem we will address in future work.

5.5 Conclusions

In this chapter we presented a scheme for simulating inhomogeneous large-scale reionization. The ionization field is evolved on precomputed N-body dark matter density fields using the SimpleX RT code, with the sources of ionizing radiation provided by the GALFORM semi-analytic galaxy formation model. At every discrete step of the density field's evolution (given by the number of N-body simulation outputs), the state of the ionization field is communicated to the source model to include the spatially inhomogeneous photo-ionization feedback. The resulting scheme provides a computationally efficient tool for investigating the role of reionization in the general picture of galaxy formation, without sacrificing accuracy in the treatment of radiative transfer effects. The most expensive reionization runs presented in this chapter took about 100 CPU hours to complete, which is equivalent to one day running on a single quad-core machine.

The first half of the chapter (Section 5.3) was dedicated to examining the possible effects the numerical parameters may have on the results obtained by the coupling scheme, building on similar discussions about both SimpleX and GALFORM individually, presented in Chapters 3 and 4 respectively. First, in Section 5.3.1, we showed that the differences in the evolution of the mean ionized fraction due to the RT grid choice are generally much smaller than the differences caused by the uncertainties introduced by the source model parameters which we explored in Chapter 4. The proper representation of the available N-body data with the local clumping factor also has a smaller effect in the reionization scenarios we explored here than in the single snapshot tests presented in Section 3.4.2. This is no surprise, as the single snapshot tests were set up to be the worst case scenario: we used the snapshot at $z = 5$ where matter clumping is much higher than at redshifts we are discussing here, e.g. reionization being complete by $z \approx 8$ (Fig. 5.3). However, even in these conditions the use of the local clumping factor improves convergence of results with respect to RT resolution. As it is inexpensive to calculate and apply, we cannot see any reason why not to use it.

The default run we examined was a 64^3 RT grid representing the L20N512 N-body simulation (see Table 4.1 for run parameters). We have shown that this 64^3 Voronoi adaptive grid is adequate for obtaining a converged evolution of $x_{\rm HII,m}$ with respect to the increase of RT grid resolution. For the L20N512, this number of grid cells results

in typical cell length scale[12] of $\Delta l \approx 0.05\,\mathrm{Mpc}/h$. The L20N512 run has approximately the same mass resolution as the L50N1024 which we use for the majority of Section 5.4 and Millennium-II (Boylan-Kolchin et al., 2009) which we will use in our future work. Therefore, we expect that the runs in those boxes will also give converged results if the RT resolution approximately matches the 64^3 cells in the 20 Mpc/hbox: 128^3 for 50 Mpc/h box side length of L50N1024 and 256^3 for 100 Mpc/h of Millennium-II.

In Section 5.3.2, we discussed the differences the RT grid resolution may introduce in finding the redshift of reionization for each halo, z_{reion}. We found that while the change in grid resolution can introduce significant scatter in z_{reion} for the majority of halos this value does not strongly depend on resolution in the ranges we considered here (see Fig. 5.8). Furthermore, as we discuss in Section 5.4.3, the source suppression has only a modest effect on the reionization progression so the issue of imprecision in the determination of z_{reion} will be equally unimportant. Note that it may have a more important effect on subsequent galaxy formation which we intend to examine in our future work.

In Section 5.4, we presented some preliminary results obtained with the coupled codes scheme. In Section 5.4.1, we found that reionization in our models proceeds strictly inside-out, from overdense to underdense regions, even in the more extreme density field construction scheme that include halos in the recombination rate. This finding agrees with more recent large-scale reionization numerical simulations (e.g. Iliev et al., 2006a; McQuinn et al., 2007; Iliev et al., 2007; Trac & Cen, 2007; Zahn et al., 2007). Figure 5.11 presented the distribution of the ionized fraction as a function of local overdensity. We found similar complex features caused by inhomogeneous recombination rates as found by Choudhury et al. (2008), who examined this issue using a semi-analytic model. In Section 5.4.2, we have shown that the large scatter in LC luminosities for halos of a given mass introduced by starbursts in the Baugh05 model (discussed in Section 4.3.3) produces specific features in the HII bubble size distribution that distinguish it from most models commonly used in the literature. We suggested that the origin of this feature may be that the source luminosity does not directly correlate with halo mass when

[12]The fact that the Voronoi grid cell volumes evolve together with the underlying N-body density makes determining an absolute spatial resolution limit difficult. The typical cell size we quote is the side of a cube with the volume V_{scale}, chosen so that $\approx 10\%$ of Voronoi cells have volumes $V < V_{\mathrm{scale}}$ at $z = 10$. In general, this volume corresponds to about 8 times smaller volume than a Cartesian uniform grid with the same number of cells would have.

starbursts are included. This claim will be examined more closely in following work. We also noted that, contrary to previous finding by McQuinn et al. (2007), a change in total LC emissivity may produce significant differences in the predicted HII region morphologies and thus complicate the comparison between models that do not predict the exact same evolution of ϵ. Finally, in Section 5.4.3, we have shown that, for our choice of the source model, the role of photo-ionization suppression of star formation on the progress of reionization is significantly smaller than what may be introduced by the source model parameters. This is a caused by a combination of the delayed suppression as implemented in GALFORM (suppressed halos only stop cooling, with no effect on the luminosities of already formed galaxies) and the steep mass-luminosity relation that predicts a smaller importance of low mass halos than what is assumed in other models in the literature.

We have not so far attempted any comparison with the currently available observational constraints. The reason is that our models are not complete at this stage: we are missing both an accurate representation of the IGM density field and the small mass halos that contribute a significant fraction of ionizing photons (as discussed in 4.5.1. However, even in these conditions, it is clear that the sources predicted by the Baugh05 model may indeed provide enough ionizing photons to reionize the Universe. The use of a lower escape fraction may yield a complete reionization at a slightly later redshift than the one derived from the CMB data assuming instanteneous reionization[13] (Komatsu et al., 2010). However, an argument can be made for assuming a higher escape fraction during bursts and, as bursts are the main producers of ionizing photons that would mean the ionization redshift will be much closer to what is obtained assuming $f_{\rm esc} = 1$. This claim will be revisited in future work. More on the plans for the future use of this scheme will be discussed in Chapter 6.

[13] A more accurate comparison between simulation results and those obtained from CMB observations is not on reionization redshift, but on the Thompson scattering optical depth, $\tau_{\rm el}$. We will use this criterion in future work.

Chapter 6 | *Conclusions*

6.1 Summary

The main aim of this thesis is to study the process of cosmological *reionization*, *i.e.* the transition from before there were many galaxies and the Universe was mostly dark and neutral, to the present state, in which galaxies are embedded in a highly ionised *intergalactic medium*. Observational constraints on how and when the Universe emerged from these 'Dark Ages' and what caused the transition, are currently limited. What was its effect on the galaxy population now and then? How does it affect the observability of source during the transition?

This field of research is about to be revolutionised over the next few years by the direct detection of reionization using radio telescopes such as LOFAR and SKA. Already now the *Hubble Space Telescope* detects galaxies at the end of the Dark Ages, the next generation *James Webb Telescope* will probe even deeper. This thesis develops and tests a set of numerical tools to study this epoch in detail and make testable predictions for the data to come. We focus in detail on two important aspects of reionization models, namely the nature of the sources of the ionizing photons, and how their ionizing radiation eventually fills the entire Universe.

Chapter 1 gives a brief overview of the basic premise of how galaxies and structure grow, based on the paradigm of hierarchical galaxy formation in a ΛCDM universe. This highly predictive cosmogony assumes that structures were seeded by quantum fluctuations, and, after a period of rapid expansion termed inflation, grew due to gravitational forces. Most of the matter is in an unknown form that does not interact with matter, called *cold dark matter* (CDM), and the expansion at late times is affected by a cosmological constant (Λ). Gas cools in dark matter *halos* to form galaxies.

Chapters 2 and **Chapter 3** discuss two numerical models for how ionizing photons propagate through a gas. Ionizing photons interact with gas in a variety of ways, affect-

ing not just its ionisation state but also heating the gas, affecting its dynamics. This is investigated in detail in **Chapter 2**, where a very accurate method for radiative transfer called *Hybrid characteristics* (HC) is discussed. The original HC implementation in an accurate hydrodynamics code (FLASH) is described, as are a number of important modifications made by us. In particular, we introduced a fully photon-conserving chemistry solver, which allows the radiative transfer calculation to be performed orders of magnitude faster. We contributed results for a set of radiative hydrodynamics tests to a code comparison project. Our improved implementation shows excellent agreement with the other schemes that took part in the project. These results have appeared in Iliev et al. (2009).

Chapter 3 describes another radiative transfer scheme, SimpleX which, unlike HC, is able to efficiently handle a very large number of sources. The SimpleX algorithm is based on the Voronoi tessellation of space. We improved the original scheme in a number of ways, in particular we discuss issues related to the sampling of the density field, the accurate treatment of recombinations using a *local clumping factor* and the treatment of cells with a large optical depth. This improves numerical convergence in the speed of ionisation fronts, which is important for accurate modelling the expansion of cosmological HII regions around sources during reionization. The parallel version of our code can propagate ionizing photons from millions of sources at high resolution in of order a day on a work station.

Chapter 4 describes GALFORM, the Durham semi-analytical model that describes the formation and evolution of galaxies. We analyse the properties of the GALFORM galaxies, and in particular investigate in detail their ionizing emissivities, and how these depend on various parameters. We find that the Bower et al. (2006) model does not produce a sufficient number of ionizing photons to reionize the Universe according to current constraints, yet the Baugh et al. (2005) (Baugh05) model does. We show that massive stars formed in abundance because of the assumed top-heavy stellar initial mass function (IMF) during starbursts in Baugh05, triggered by galaxy mergers, are the dominant source of ionizing photons. If the escape fraction of ionizing photons in bursts is not much less than 10 per cent, than the Baugh05 model can obtain reionization due to normal galaxies by redshift ≈ 9, which may be consistent with current constraints. In addition, the galaxy properties of Baugh05 are in excellent agreement with the ob-

served luminosity functions of galaxies at all redshifts up to where galaxies have been measured, $z \approx 10$. In this chapter we performed the first comparison of the Baugh05 model with recently observed Lyman-break galaxies at $z \gtrsim 8$. This makes the Baugh05 model a promising ingredient for a comprehensive model of how reionization proceeds. It predicts that most ionizing photons are produced during a burst in faint galaxies, consequently the currently detected galaxies at $z > 8$ are expected to contribute only a small fraction of the total ionizing emissivity. A paper describing the results in this chapter is in preparation (Raicevic et al. 2010).

Chapter 5 describes a scheme for simulating large-scale reionization by coupling the improved SimpleX radiative transfer code with the GALFORM galaxy formation model. This allows us to investigate how galaxy clustering leads to inhomogeneous reionization. We discuss the evolution of the sizes and morphology of ionised regions, as well as the effect of the suppression of galaxies overrun by an ionization front. The scheme presented is fast and accurate, making it an excellent tool for the numerical study of epoch of reionization. This allows a careful exploration of the presently poorly constrained parameters of the model. A paper describing the results in this chapter is in preparation (Raicevic et al. 2010).

6.2 Future work

The work presented in this thesis will be continued on two fronts, radiative hydrodynamics with FLASH-HC and large scale reionization studies with SimpleX-GALFORM.

6.2.1 Hybrid Characteristics and FLASH

Method improvements

Apart from finding and fixing the problems seen in the results of Test 7 (Section 2.6.3) we are planing to include two improvements to the HC method. First, in the procedure for finding which blocks are cut by each ray (described in Section 2.3.2), we will replace the global block map with a procedure that makes use of the FLASH AMR octree data structure. This will significantly decrease the memory consumption of the scheme, allowing the AMR grids to resolve an even higher dynamic range. Second, we will update the chemistry solver to include the evolution of species beyond hydrogen, based on the

scheme introduced by Anninos et al. (1997).

Metal enrichment by the first supernovae

Thus updated, we will use the FLASH-HC scheme to study the propagation of metals produces by Pop III supernovae. The initial conditions (of a first star forming in a minihalo at $z \sim 20$) will be imported from the Aquarius simulations, using the Gadget-to-FLASH converter described in Section 2.7. In the next step, the highly detailed HII region will be simulated using HC (similar to e.g. Whalen et al., 2004; Abel et al., 2007). After the life time of the star ($t \sim 10$ Myr), a supernova explosion will be set up. The AMR grid will allow tracking of the low metal fluxes propagating away from the SN blast with significantly higher precision than can be attained in SPH simulations of the same process (Scannapieco et al., 2005; Kobayashi et al., 2007; Greif et al., 2007). We will use the initial conditions of the Aquarius project, which have the advantage that they lead to the formation of a Milky-Way like galaxy. Therefore we will be able to address to what extent signatures of Pop III ejecta are expected in for example Milky Way satellites, or the Milky Way's bulge. This will be published as Creasy et al (2010).

6.2.2 Reionization with SimpleX and GALFORM

Method improvements

We intend to improve the scheme presented in this chapter in several ways, while preserving the low computational cost that is its key feature. First, the density field must be represented in more detail, as argued in Section 3.4.1. We will use the local clumping factors we introduced in Section 3.4.2 calibrated from higher resolution SPH simulations to include the effects of the gas pressure and temperature, as discussed by Pawlik et al. (2009). Second, we intend to include minihalos and Lyman-limit systems as dense, self-shielded absorbers (McQuinn et al., 2007). Finally, we will update the SimpleX chemistry solver to include helium ionizations as well as to compute the gas temperature (the energy equation introduced in Chapter 2, Eq. (2.19)).

We also intend to expand the basic SimpleX algorithm to describe the propagation of Lyman-α photons (Orsi et al, 2010).

Observability of reionization and its effects on galaxy formation

We will perform reionization simulations using the 'Millennium-II' simulation (MSII, Boylan-Kolchin et al., 2009), which evolves 2192^3 particles in a box of size $100\,\mathrm{Mpc}/h$. This will allow us to study in great detail the progression of reionization, including estimates of the 21-cm signal. Crucially, since this density field is evolved to the present day, we can investigate the effects of inhomogeneous reionization on the $z = 0$ luminosity function and other observables, for example the satellite luminosity function. This work will be published as Raicevic et al (2010). In addition we will be able to simulate the buildup and possible inhomogeneities in the post-reionization UV-background (e.g. Mesinger & Furlanetto, 2009). The combination of the superb statics provide by the 'Millennium-II' simulation, the realism of the sources provided by GALFORM, and the accurate treatment of radiative transfer with SimpleX, will allow us to make a significant contribution to studies of the Epoch of Reionization.

Bibliography

Abel T., Bryan G. L., Norman M. L., 2000, Astrophysical Journal, 540, 39

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

Abel T., Norman M. L., Madau P., 1999, Astrophysical Journal, 523, 66

Abel T., Wandelt B. D., 2002, Monthly Notices of the Royal Astronomical Society, 330, L53

Abel T., Wise J. H., Bryan G. L., 2007, The Astrophysical Journal, 659, L87

Agertz O., Moore B., Stadel J., Potter D., Miniati F., Read J., Mayer L., Gawryszczak A., Kravtsov A., ke Nordlund Pearce F., Quilis V., Rudd D., Springel V., Stone J., Tasker E., Teyssier R., Wadsley J., Walder R., 2007, Monthly Notices of the Royal Astronomical Society, 380, 963

Agertz O., Teyssier R., Moore B., 2009, Monthly Notices of the Royal Astronomical Society, 397, L64

Ahn K., Shapiro P. R., Iliev I. T., Mellema G., Pen U., , 2008, The Inhomogeneous Background of Hydrogen-Molecule Dissociating Radiation during Cosmic Reionization, http://adsabs.harvard.edu/abs/2008arXiv0807.2254A

Alexander D. M., Bauer F. E., Brandt W. N., Hornschemeier A. E., Vignali C., Garmire G. P., Schneider D. P., Chartas G., Gallagher S. C., 2003, Astronomical Journal, 125, 383

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

Altay G., Croft R. A. C., Pelupessy I., 2008, Monthly Notices of the Royal Astronomical Society, 386, 1931

Alvarez M. A., Bromm V., Shapiro P. R., 2005, astro-ph/0507684

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

Amanatides J., Woo A., 1987, Proc. Euro-graphics, 1, 3

Anninos P., Zhang Y., Abel T., Norman M. L., 1997, New Astronomy, 2, 209

Aubert D., Teyssier R., 2008, Monthly Notices of the Royal Astronomical Society, 387, 295

Axford W. I., 1961, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 253, 301

Babul A., Rees M. J., 1992, Monthly Notices of the Royal Astronomical Society, 255, 346

Baek S., Matteo P. D., Semelin B., Combes F., Revaz Y., 2009, Astronomy and Astrophysics, 495, 389

Bagla J. S., Prasad J., 2006, Monthly Notices of the Royal Astronomical Society, 370, 993

Barber C. B., Dobkin D. P., Huhdanpaa H., 1996, ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE, 22, 469

Barkana R., Loeb A., 1999, Astrophysical Journal, 523, 54

Barkana R., Loeb A., 2001, Physics Reports, 349, 125

Barkana R., Loeb A., 2002, Astrophysical Journal, 578, 1

Barkana R., Loeb A., 2004, Astrophysical Journal, 609, 474

Barnes J., Efstathiou G., 1987, Astrophysical Journal, 319, 575

Barnes J. E., 1998 Dynamics of galaxy interactions. p. 275

Barnes J. E., Hernquist L., 1992, Annual Review of Astronomy and Astrophysics, 30, 705

Bates D., Watts D., 1988, Nonlinear Regression and Its Applications. Wiley, New York

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Baugh C. M., Cole S., Frenk C. S., 1996, Monthly Notices of the Royal Astronomical Society, 282, L27

Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 1998, Astrophysical Journal, 498, 504

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, Monthly Notices of the Royal Astronomical Society, 356, 1191

Bechtold J., 2003 Quasar absorption lines. pp 131–184

Becker R. H., Fan X., White R. L., Strauss M. A., Narayanan V. K., Lupton R. H., Gunn J. E., Annis J., Bahcall N. A., Brinkmann J., Connolly A. J., Csabai I., Czarapata P. C., Doi M., Heckman T. M., Hennessy G. S., eljko Ivezi Knapp G. R., Lamb D. Q., McKay T. A., Munn J. A., Nash T., Nichol R., Pier J. R., Richards G. T., Schneider D. P., Stoughton C., Szalay A. S., Thakar A. R., York D. G., 2001, Astronomical Journal, 122, 2850

Beckwith S. V. W., Stiavelli M., Koekemoer A. M., Caldwell J. A. R., Ferguson H. C., Hook R., Lucas R. A., Bergeron L. E., Corbin M., Jogee S., Panagia N., Robberto M., Royle P., Somerville R. S., Sosey M., 2006, Astronomical Journal, 132, 1729

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, Astrophysical Journal, 599, 38

Benson A. J., Devereux N., 2009, Monthly Notices of the Royal Astronomical Society

Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002, Monthly Notices of the Royal Astronomical Society, 333, 177

Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, Monthly Notices of the Royal Astronomical Society, 333, 156

Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, Monthly Notices of the Royal Astronomical Society, 320, 261

Benson A. J., Sugiyama N., Nusser A., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society, 369, 1055

Berger M. J., Colella P., 1989, J. Comput. Phys., 82, 64

Berger M. J., Oliger J., 1984, J. Comput. Phys., 53, 484

Bergvall N., Zackrisson E., Andersson B., Arnberg D., Masegosa J., stlin G., 2006, Astronomy and Astrophysics, 448, 513

Bertoldi F., 1989, Astrophysical Journal, 346, 735

Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, Monthly Notices of the Royal Astronomical Society, 376, 215

Binney J., Tremaine S., 1987, Galactic dynamics

Blain A. W., Smail I., Ivison R. J., Kneib J., Frayer D. T., 2002, Physics Reports, 369, 111

Bland-Hawthorn J., Maloney P. R., 1999, Astrophysical Journal, 510, L33

Bodenheimer P., Laughlin G. P., Rzyczka M., Yorke H. W., 2007, Numerical Methods in Astrophysics: An Introduction

Bolton J., Meiksin A., White M., 2004, Monthly Notices of the Royal Astronomical Society, 348, L43

Bolton J. S., Haehnelt M. G., 2007, Monthly Notices of the Royal Astronomical Society, 382, 325

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, Astrophysical Journal, 379, 440

Bouwens R. J., Illingworth G. D., Bradley L. D., Ford H., Franx M., Zheng W., Broadhurst T., Coe D., Jee M. J., 2008, 0805.0593

Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, 0707.2080

Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2008, 0803.0548

Bouwens R. J., Illingworth G. D., Labbe I., Oesch P. A., Carollo M., Trenti M., van Dokkum P. G., Franx M., Stiavelli M., Gonzalez V., Magee D., 2009, 0912.4263

Bouwens R. J., Illingworth G. D., Oesch P. A., Stiavelli M., van Dokkum P., Trenti M., Magee D., Labbe I., Franx M., Carollo M., 2009, 0909.1803

Bouwens R. J., Illingworth G. D., Thompson R. I., Blakeslee J. P., Dickinson M. E., Broadhurst T. J., Eisenstein D. J., Fan X., Franx M., Meurer G., van Dokkum P., 2004, Astrophysical Journal, 606, L25

Bower R. G., 1991, Monthly Notices of the Royal Astronomical Society, 248, 332

Bower R. G., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2001, Monthly Notices of the Royal Astronomical Society, 325, 497

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society, 370, 645

Bower R. G., McCarthy I. G., Benson A. J., 2008, Monthly Notices of the Royal Astronomical Society, 390, 1399

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, 0903.3041

Bressan A., Chiosi C., Fagotto F., 1994, Astrophysical Journal Supplement Series, 94, 63

Bressan A., Granato G. L., Silva L., 1998, Astronomy and Astrophysics, 332, 135

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

Bromm V., Ferrara A., Coppi P. S., Larson R. B., 2001, Monthly Notices of the Royal Astronomical Society, 328, 969

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

Bromm V., Larson R. B., 2004, Annual Review of Astronomy and Astrophysics, 42, 79

Bruscoli M., Ferrara A., Fabbri R., Ciardi B., 2000, Monthly Notices of the Royal Astronomical Society, 318, 1068

Bruzual G., Charlot S., 1993, Astrophysical Journal, 405, 538

Bruzual G., Charlot S., 2003, Monthly Notices of the Royal Astronomical Society, 344, 1000

Bryan G. L., Machacek M., Anninos P., Norman M. L., 1999, The Astrophysical Journal, 517, 13

Bryan G. L., Norman M. L., 1997 Vol. 123, Simulating X-Ray clusters with adaptive mesh refinement. p. 363

Bryan G. L., Norman M. L., 1998, Astrophysical Journal, 495, 80

Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, Astrophysical Journal, 555, 240

Bunker A., Stanway E., Ellis R., Lacy M., McMahon R., Eyles L., Stark D., Chiu K., 2009, 0909.1565

Bunker A. J., Stanway E. R., Ellis R. S., McMahon R. G., 2004, Monthly Notices of the Royal Astronomical Society, 355, 374

Busha M. T., Alvarez M. A., Wechsler R. H., Abel T., Strigari L. E., , 2009, The Impact of Inhomogeneous Reionization on the Satellite Galaxy Population of the Milky Way, http://adsabs.harvard.edu/abs/2009arXiv0901.3553B

Carroll S. M., Press W. H., Turner E. L., 1992, Annual Review of Astronomy and Astrophysics, 30, 499

Cashwell E. D., Evrett C. J., 1959, A Practical Manual on the Monte Carlo Method for Random Walk Problems. Pergamon, New York

Castor J. I., 2004, Radiation Hydrodynamics

Catelan P., Theuns T., 1996, Monthly Notices of the Royal Astronomical Society, 282, 436

Cavaliere A., Fusco-Femiano R., 1976, Astronomy and Astrophysics, 49, 137

Cen R., 2005, Astrophysical Journal, 624, 485

Cen R., Miralda-Escude J., Ostriker J. P., Rauch M., 1994, The Astrophysical Journal, 437, L9

Chandrasekhar S., 1943, Astrophysical Journal, 98, 54

Chandrasekhar S., 1950, Radiative transfer.

Choudhury T. R., Haehnelt M. G., Regan J., 2008, 0806.1524

Christodoulou D. M., Shlosman I., Tohline J. E., 1995, Astrophysical Journal, 443, 551

Ciardi B., 2007 Epoch of reionization. p. 10

Ciardi B., Bianchi S., Ferrara A., 2002, Monthly Notices of the Royal Astronomical Society, 331, 463

Ciardi B., Ferrara A., 2005, Space Science Reviews, 116, 625

Ciardi B., Ferrara A., Governato F., Jenkins A., 2000, Monthly Notices of the Royal Astronomical Society, 314, 611

Ciardi B., Scannapieco E., Stoehr F., Ferrara A., Iliev I. T., Shapiro P. R., 2006, Monthly Notices of the Royal Astronomical Society, 366, 689

Ciardi B., Stoehr F., White S. D. M., 2003, Monthly Notices of the Royal Astronomical Society, 343, 1101

Clarke C., Oey M. S., 2002, Monthly Notices of the Royal Astronomical Society, 337, 1299

Cole S., 1991, Astrophysical Journal, 367, 45

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, Monthly Notices of the Royal Astronomical Society, 271, 781

Cole S., Lacey C., 1996, Monthly Notices of the Royal Astronomical Society, 281, 716

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, Monthly Notices of the Royal Astronomical Society, 319, 168

Colella P., Woodward P. R., 1984, Journal of Computational Physics, 54, 174

Coles P., Lucchin F., 2002, Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition

Colless M., Dalton G., Maddox S., Sutherland W., Norberg P., Cole S., Bland-Hawthorn J., Bridges T., Cannon R., Collins C., Couch W., Cross N., Deeley K., Propris R. D., Driver S. P., Efstathiou G., Ellis R. S., Frenk C. S., Glazebrook K., Jackson C., Lahav O., Lewis I., Lumsden S., Madgwick D., Peacock J. A., Peterson B. A., Price I., Seaborne M., Taylor K., 2001, Monthly Notices of the Royal Astronomical Society, 328, 1039

Combes F., 2000 Bulge formation. p. 413

Combes F., Debbasch F., Friedli D., Pfenniger D., 1990, Astronomy and Astrophysics, 233, 82

Couchman H. M. P., Rees M. J., 1986, Monthly Notices of the Royal Astronomical Society, 221, 53

Cowie L. L., Hu E. M., 1998, Astronomical Journal, 115, 1319

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, Astronomical Journal, 112, 839

Crain R. A., Theuns T., Vecchia C. D., Eke V. R., Frenk C. S., Jenkins A., Kay S. T., Peacock J. A., Pearce F. R., Schaye J., Springel V., Thomas P. A., White S. D. M., Wiersma R. P. C., 2009, 0906.4350

Croton D. J., Springel V., White S. D. M., Lucia G. D., Frenk C. S., Gao L., Jenkins A., Kauffmann G., Navarro J. F., Yoshida N., 2006, Monthly Notices of the Royal Astronomical Society, 365, 11

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

de Jong R. S., Lacey C., 2000, Astrophysical Journal, 545, 781

Dekel A., Silk J., 1986, Astrophysical Journal, 303, 39

Delliou M. L., Lacey C., Baugh C. M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S. L., 2005, Monthly Notices of the Royal Astronomical Society, 357, L11

Delliou M. L., Lacey C. G., Baugh C. M., Morris S. L., 2006, Monthly Notices of the Royal Astronomical Society, 365, 712

Dijkstra M., 2006, New Astronomy Review, 50, 204

Djorgovski S., Spinrad H., McCarthy P., Strauss M. A., 1985, Astrophysical Journal, 299, L1

Doroshkevich A. G., 1970, Astrofizika, 6, 581

Doroshkevich A. G., Zel'Dovich Y. B., Novikov I. D., 1967, Soviet Astronomy, 11, 233

Dunkley J., Spergel D. N., Komatsu E., Hinshaw G., Larson D., Nolta M. R., Odegard N., Page L., Bennett C. L., Gold B., Hill R. S., Jarosik N., Weiland J. L., Halpern M., Kogut A., Limon M., Meyer S. S., Tucker G. S., Wollack E., Wright E. L., 2009, Astrophysical Journal, 701, 1804

Efstathiou G., 1992, Monthly Notices of the Royal Astronomical Society, 256, 43P

Efstathiou G., 2000, Monthly Notices of the Royal Astronomical Society, 317, 697

Efstathiou G., Jones B. J. T., 1979, Monthly Notices of the Royal Astronomical Society, 186, 133

Efstathiou G., Lake G., Negroponte J., 1982, Monthly Notices of the Royal Astronomical Society, 199, 1069

Efstathiou G., Sutherland W. J., Maddox S. J., 1990, Nature, 348, 705

Einstein A., 1916, Annalen der Physik, 354, 769

Eke V. R., Cole S., Frenk C. S., 1996, Monthly Notices of the Royal Astronomical Society, 282, 263

Eke V. R., Cole S., Frenk C. S., Henry J. P., 1998, Monthly Notices of the Royal Astronomical Society, 298, 1145

Ellis R. S., 1997, Annual Review of Astronomy and Astrophysics, 35, 389

Fall S. M., Efstathiou G., 1980, Monthly Notices of the Royal Astronomical Society, 193, 189

Fan X., Carilli C. L., Keating B., 2006, Annual Review of Astronomy and Astrophysics, 44, 415

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

Fang T., Cen R., 2004, The Astrophysical Journal, 616, L87

Faucher-Giguere C., Lidz A., Hernquist L., Zaldarriaga M., 2008, Astrophysical Journal, 682, L9

Ferguson H. C., Dickinson M., Williams R., 2000, Annual Review of Astronomy and Astrophysics, 38, 667

Ferrara A., 1998, Astrophysical Journal, 499, L17

Ferrara A., Bianchi S., Cimatti A., Giovanardi C., 1999, Astrophysical Journal Supplement Series, 123, 437

Finlator K., Ozel F., Dave R., Oppenheimer B. D., 2009, 0908.1576

Finlator K., zel F., Dav R., 2009, Monthly Notices of the Royal Astronomical Society, 393, 1090

Font A. S., Bower R. G., McCarthy I. G., Benson A. J., Frenk C. S., Helly J. C., Lacey C. G., Baugh C. M., Cole S., 2008, Monthly Notices of the Royal Astronomical Society, 389, 1619

Ford H. C., Clampin M., Hartig G. F., Illingworth G. D., Sirianni M., Martel A. R., Meurer G. R., McCann W. J., Sullivan P. C., Bartko F., Benitez N., Blakeslee J., Bouwens R., Broadhurst T., Brown R. A., Burrows C. J., Campbell D., Cheng E. S., Feldman P. D., Franx M., Golimowski D. A., Gronwall C., Kimble R. A., Krist J. E., Lesser M. P., Magee D., Miley G., Postman M., Rafal M. D., Rosati P., Sparks W. B., Tran H. D., Tsvetanov Z. I., Volmer P., White R. L., Woodruff R. A., 2003 Vol. 4854, Overview of the advanced camera for surveys on-orbit performance. pp 81–94

Franco J., Tenorio-Tagle G., Bodenheimer P., 1990, Astrophysical Journal, 349, 126

Frank A., Mellema G., 1994, Astronomy and Astrophysics, 289, 937

Frenk C. S., White S. D. M., Bode P., Bond J. R., Bryan G. L., Cen R., Couchman H. M. P., Evrard A. E., Gnedin N., Jenkins A., Khokhlov A. M., Klypin A., Navarro J. F., Norman M. L., Ostriker J. P., Owen J. M., Pearce F. R., Pen U., Steinmetz M., Thomas P. A., Villumsen J. V., Wadsley J. W., Warren M. S., Xu G., Yepes G., 1999, Astrophysical Journal, 525, 554

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, Astrophysical Journal, 327, 507

Friedman A., 1922, Zeitschrift fr Physik A Hadrons and Nuclei, 10, 377

Frieman J. A., Turner M. S., Huterer D., 2008, Annual Review of Astronomy and Astrophysics, 46, 385

Fryxell B., Olson K., Ricker P., Timmes F. X., Zingale M., Lamb D. Q., MacNeice P., Rosner R., Truran J. W., Tufo H., 2000, Astrophysical Journal Supplement Series, 131, 273

Fujita A., Martin C. L., Low M. M., Abel T., 2003, Astrophysical Journal, 599, 50

Fukugita M., Kawasaki M., 1994, Monthly Notices of the Royal Astronomical Society, 269, 563

Furlanetto S., Oh S. P., Briggs F., 2006, astro-ph/0608032

Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, Astrophysical Journal, 613, 1

Gamow G., 1948, Physical Review, 74, 505

Genel S., Genzel R., Bouch N., Naab T., Sternberg A., 2009, Astrophysical Journal, 701, 2002

Giavalisco M., 2002, Annual Review of Astronomy and Astrophysics, 40, 579

Giavalisco M., Ferguson H. C., Koekemoer A. M., Dickinson M., Alexander D. M., Bauer F. E., Bergeron J., Biagetti C., Brandt W. N., Casertano S., Cesarsky C., Chatzichristou E., Conselice C., Cristiani S., Costa L. D., Dahlen T., de Mello D., Eisenhardt P., Erben T., Fall S. M., Fassnacht C., Fosbury R., Fruchter A., Gardner J. P., Grogin N., Hook R. N., Hornschemeier A. E., Idzi R., Jogee S., Kretchmer C., Laidler V., Lee K. S., Livio M., Lucas R., Madau P., Mobasher B., Moustakas L. A., Nonino M., Padovani P., Papovich C., Park Y., Ravindranath S., Renzini A., Richardson M., Riess A., Rosati P., Schirmer M., Schreier E., Somerville R. S., Spinrad H., Stern D., Stiavelli M., Strolger L., Urry C. M., Vandame B., Williams R., Wolf C., 2004, Astrophysical Journal, 600, L93

Gnedin N. Y., 1995, The Astrophysical Journal Supplement Series, 97, 231

Gnedin N. Y., 2000a, Astrophysical Journal, 535, 530

Gnedin N. Y., 2000b, Astrophysical Journal, 542, 535

Gnedin N. Y., 2008, Astrophysical Journal, 673, L1

Gnedin N. Y., Abel T., 2001, New Astronomy, 6, 437

Gnedin N. Y., Hui L., 1998, Monthly Notices of the Royal Astronomical Society, 296, 44

Gnedin N. Y., Ostriker J. P., 1997, Astrophysical Journal, 486, 581

Godunov S., 1959, Matematicheskii Sbornik, 47, 271

Gonzalez J. E. et al., 2010, in preparation

Gonzalez J. E., Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2009, Monthly Notices of the Royal Astronomical Society, 397, 1254

Governato F., Mayer L., Wadsley J., Gardner J. P., Willman B., Hayashi E., Quinn T., Stadel J., Lake G., 2004, Astrophysical Journal, 607, 688

Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, Astrophysical Journal, 542, 710

Greif T. H., Johnson J. L., Bromm V., Klessen R. S., 2007, The Astrophysical Journal, 670, 1

Gunn J. E., Gott J. R., 1972, Astrophysical Journal, 176, 1

Gunn J. E., Peterson B. A., 1965, Astrophysical Journal, 142, 1633

Guth A. H., 1981, Physical Review D, 23, 347

Haardt F., Madau P., 1996, Astrophysical Journal, 461, 20

Habib S., Pope A., Lukic Z., Daniel D., Fasel P., Desai N., Heitmann K., Hsu C., Ankeny L., Mark G., Bhattacharya S., Ahrens J., 2009, Journal of Physics: Conference Series, 180, 012019

Haiman Z., Holder G. P., 2003, Astrophysical Journal, 595, 1

Haiman Z., Loeb A., 1997, Astrophysical Journal, 483, 21

Haiman Z., Loeb A., 1998, Astrophysical Journal, 503, 505

Harary F., 1994, Graph Theory. Westview Press, Reading, MA

Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, Monthly Notices of the Royal Astronomical Society, 367, 1039

Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, Monthly Notices of the Royal Astronomical Society, 343, 75

Heckman T. M., 2000 Vol. 358, The role of starbursts in the formation of galaxies and active galactic nuclei. p. 2077

Heckman T. M., Sembach K. R., Meurer G. R., Leitherer C., Calzetti D., Martin C. L., 2001, Astrophysical Journal, 558, 56

Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., 2003, Monthly Notices of the Royal Astronomical Society, 338, 903

Hernquist L., Katz N., Weinberg D. H., MiraldaEscude J., 1996, ASTROPHYSICAL JOURNAL, 457, L51

Hernquist L., Mihos J. C., 1995, Astrophysical Journal, 448, 41

Hindmarsh A. C., 1980, ACM SIGNUM Newsletter, 15, 10

Hinshaw G., Weiland J. L., Hill R. S., Odegard N., Larson D., Bennett C. L., Dunkley J., Gold B., Greason M. R., Jarosik N., Komatsu E., Nolta M. R., Page L., Spergel D. N., Wollack E., Halpern M., Kogut A., Limon M., Meyer S. S., Tucker G. S., Wright E. L., 2009, Astrophysical Journal Supplement Series, 180, 225

Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles

Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles

Hoeft M., Yepes G., Gottlber S., Springel V., 2006, Monthly Notices of the Royal Astronomical Society, 371, 401

Hoyle F., 1951 The origin of the rotations of the galaxies. p. 195

Hoyle F., 1953, Astrophysical Journal, 118, 513

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

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

Hu W., White M., 1997, New Astronomy, 2, 323

Hubble E., Humason M. L., 1931, Astrophysical Journal, 74, 43

Hui L., Haiman Z., 2003, Astrophysical Journal, 596, 9

Hurwitz M., Jelinsky P., Dixon W. V. D., 1997, Astrophysical Journal, 481, L31

Iliev I. T., Ciardi B., Alvarez M. A., Maselli A., Ferrara A., Gnedin N. Y., Mellema G., Nakamoto T., Norman M. L., Razoumov A. O., Rijkhorst E., Ritzerveld J., Shapiro P. R.,

Susa H., Umemura M., Whalen D. J., 2006, Monthly Notices of the Royal Astronomical Society, 371, 1057

Iliev I. T., Mellema G., Pen U., Merz H., Shapiro P. R., Alvarez M. A., 2006, Monthly Notices of the Royal Astronomical Society, 369, 1625

Iliev I. T., Pen U., Bond J. R., Mellema G., Shapiro P. R., 2007, Astrophysical Journal, 660, 933

Iliev I. T., Whalen D., Mellema G., Ahn K., Baek S., Gnedin N. Y., Kravtsov A. V., Norman M., Raicevic M., Reynolds D. R., Sato D., Shapiro P. R., Semelin B., Smidt J., Susa H., Theuns T., Umemura M., 2009, 0905.2920

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

Iliev I. T., Shapiro P. R., Mellema G., Merz H., Pen U., , 2008, Simulating Cosmic Reionization, http://adsabs.harvard.edu/abs/2008arXiv0806.2887I

Iliev I. T., Shapiro P. R., Raga A. C., 2005, Monthly Notices of the Royal Astronomical Society, 361, 405

Inoue A. K., Iwata I., Deharveng J., 2006, Monthly Notices of the Royal Astronomical Society, 371, L1

Ivison R. J., Smail I., Borgne J. L., Blain A. W., Kneib J., Bezecourt J., Kerr T. H., Davies J. K., 1998, Monthly Notices of the Royal Astronomical Society, 298, 583

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, Monthly Notices of the Royal Astronomical Society, 321, 372

Jones E., Oliphant T., Peterson P., , 2001, SciPy: Open Source Scientific Tools for Python

Kahn F. D., 1954, Bulletin of the Astronomical Institutes of the Netherlands, 12, 187

Kashikawa N., Shimasaku K., Malkan M. A., Doi M., Matsuda Y., Ouchi M., Taniguchi Y., Ly C., Nagao T., Iye M., Motohara K., Murayama T., Murozono K., Nariai K., Ohta K., Okamura S., Sasaki T., Shioya Y., Umemura M., 2006, Astrophysical Journal, 648, 7

Kauffmann G., Guiderdoni B., White S. D. M., 1994, Monthly Notices of the Royal Astronomical Society, 267, 981

Kauffmann G., White S. D. M., 1993, Monthly Notices of the Royal Astronomical Society, 261, 921

Kauffmann G., White S. D. M., Guiderdoni B., 1993, Monthly Notices of the Royal Astronomical Society, 264, 201

Kennicutt R. C., 1983, Astrophysical Journal, 272, 54

Kitayama T., Tajiri Y., Umemura M., Susa H., Ikeuchi S., 2000, Monthly Notices of the Royal Astronomical Society, 315, L1

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

Kochanek C. S., Pahre M. A., Falco E. E., Huchra J. P., Mader J., Jarrett T. H., Chester T., Cutri R., Schneider S. E., 2001, Astrophysical Journal, 560, 566

Kohler K., Gnedin N. Y., 2005 Vol. 207, Large-Scale simulations of reionization. p. 1417

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

Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., Limon M., Page L., Spergel D. N., Halpern M., Hill R. S., Kogut A., Meyer S. S., Tucker G. S., Weiland J. L., Wollack E., Wright E. L., 2009, Astrophysical Journal Supplement Series, 180, 330

Komatsu E., Smith K. M., Dunkley J., Bennett C. L., Gold B., Hinshaw G., Jarosik N., Larson D., Nolta M. R., Page L., Spergel D. N., Halpern M., Hill R. S., Kogut A., Limon M., Meyer S. S., Odegard N., Tucker G. S., Weiland J. L., Wollack E., Wright E. L., 2010, 1001.4538

Koposov S., Belokurov V., Evans N. W., Hewett P. C., Irwin M. J., Gilmore G., Zucker D. B., Rix H., Fellhauer M., Bell E. F., Glushkova E. V., 2008, Astrophysical Journal, 686, 279

Kroupa P., 2001 Vol. 228, The local stellar initial mass function. p. 187

Lacey C., 2006 A Multi-Wavelength model of galaxy formation. p. 253

Lacey C., Cole S., 1993, Monthly Notices of the Royal Astronomical Society, 262, 627

Lacey C., Cole S., 1994, Monthly Notices of the Royal Astronomical Society, 271, 676

Lacey C., Silk J., 1991, Astrophysical Journal, 381, 14

Lacey C. G. et al., 2010, in preparation

Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., Orsi A., Silva L., Granato G. L., Bressan A., 2009, 0909.1567

Lacey C. G., Baugh C. M., Frenk C. S., Silva L., Granato G. L., Bressan A., 2008, Monthly Notices of the Royal Astronomical Society, 385, 1155

Landau L. D., Lifshitz E. M., 1959, Fluid mechanics

Lang K. R., 1999, Astrophysical formulae

Lanzetta K. M., McMahon R. G., Wolfe A. M., Turnshek D. A., Hazard C., Lu L., 1991, Astrophysical Journal Supplement Series, 77, 1

Larson R. B., 1974, Monthly Notices of the Royal Astronomical Society, 169, 229

Lehnert M. D., Bremer M., 2003, Astrophysical Journal, 593, 630

Leitherer C., Ferguson H. C., Heckman T. M., Lowenthal J. D., 1995, Astrophysical Journal, 454, L19

Lidz A., Zahn O., McQuinn M., Zaldarriaga M., Hernquist L., 2008, Astrophysical Journal, 680, 962

Lim A. J., Mellema G., 2003, Astronomy and Astrophysics, 405, 189

Linde A. D., 1990, Contemporary Concepts in Physics, 5

Loeb A., 2006, astro-ph/0603360

Loeb A., 2009, Journal of Cosmology and Astro-Particle Physics, 03, 022

MacNeice P., Olson K. M., Mobarry C., de Fainchtein R., Packer C., 2000, Computer Physics Communications, 126, 330

Madau P., 1995, Astrophysical Journal, 441, 18

Madau P., 1999 Vol. 470, Faint galaxies, extragalactic background light, and the reionization of the universe. pp 299–311

Madau P., 2000, astro-ph/0005106

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, Monthly Notices of the Royal Astronomical Society, 283, 1388

Madau P., Meiksin A., Rees M. J., 1997, Astrophysical Journal, 475, 429

Madau P., Pozzetti L., Dickinson M., 1998, Astrophysical Journal, 498, 106

Maraston C., 2005, Monthly Notices of the Royal Astronomical Society, 362, 799

Martin N. F., Ibata R. A., Chapman S. C., Irwin M., Lewis G. F., 2007, Monthly Notices of the Royal Astronomical Society, 380, 281

Maselli A., Ferrara A., Ciardi B., 2003, Monthly Notices of the Royal Astronomical Society, 345, 379

McKee C. F., Ostriker J. P., 1977, Astrophysical Journal, 218, 148

McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, Monthly Notices of the Royal Astronomical Society, 377, 1043

McQuinn M., Lidz A., Zaldarriaga M., Hernquist L., Hopkins P. F., Dutta S., Faucher-Gigure C., 2009, Astrophysical Journal, 694, 842

Meiksin A. A., 2007, 0711.3358

Mellema G., Iliev I. T., Pen U., Shapiro P. R., 2006, Monthly Notices of the Royal Astronomical Society, 372, 679

Mellema G., Raga A. C., Canto J., Lundqvist P., Balick B., Steffen W., Noriega-Crespo A., 1998, Astronomy and Astrophysics, 331, 335

Melnyk O. V., Elyiv A. A., Vavilova I. B., 2007 Vol. 235, 3-D voronoi's tessellation as a tool for identifying galaxy groups. p. 223

Mesinger A., Dijkstra M., 2008, Monthly Notices of the Royal Astronomical Society, 390, 1071

Mesinger A., Furlanetto S., 2009, 0906.3020

Mihalas D., Mihalas B. W., 1984, Foundations of radiation hydrodynamics

Miralda-Escude J., 2003, The Astrophysical Journal, 597, 66

Miralda-Escude J., Haehnelt M., Rees M. J., 2000, Astrophysical Journal, 530, 1

Miralda-Escude J., Rees M. J., 1994, Monthly Notices of the Royal Astronomical Society, 266, 343

Miyazaki S., Komiyama Y., Sekiguchi M., Okamura S., Doi M., Furusawa H., Hamabe M., Imi K., Kimura M., Nakata F., Okada N., Ouchi M., Shimasaku K., Yagi M., Yasuda N., 2002, Publications of the Astronomical Society of Japan, 54, 833

Mo H. J., Mao S., White S. D. M., 1998, Monthly Notices of the Royal Astronomical Society, 295, 319

Monaghan J. J., 1992, Annual Review of Astronomy and Astrophysics, 30, 543

Monaghan J. J., Lattanzio J. C., 1985, Astronomy and Astrophysics, 149, 135

Morton G., 1966, Ottawa, Canada: IBM Ltd.

Munoz J. A., Madau P., Loeb A., Diemand J., 2009, Monthly Notices of the Royal Astronomical Society, 400, 1593

Nagashima M., Gouda N., Sugiura N., 1999, Monthly Notices of the Royal Astronomical Society, 305, 449

Nagashima M., Lacey C. G., Baugh C. M., Frenk C. S., Cole S., 2005, Monthly Notices of the Royal Astronomical Society, 358, 1247

Nagashima M., Lacey C. G., Okamoto T., Baugh C. M., Frenk C. S., Cole S., 2005, Monthly Notices of the Royal Astronomical Society, 363, L31

Nakamura O., Fukugita M., Yasuda N., Loveday J., Brinkmann J., Schneider D. P., Shimasaku K., SubbaRao M., 2003, Astronomical Journal, 125, 1682

Navarro J. F., Frenk C. S., White S. D. M., 1995, Monthly Notices of the Royal Astronomical Society, 275, 720

Navarro J. F., Frenk C. S., White S. D. M., 1996, Astrophysical Journal, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, Astrophysical Journal, 490, 493

Nishi R., Tashiro M., 2000, Astrophysical Journal, 537, 50

Norberg P., Cole S., Baugh C. M., Frenk C. S., Baldry I., Bland-Hawthorn J., Bridges T., Cannon R., Colless M., Collins C., Couch W., Cross N. J. G., Dalton G., Propris R. D., Driver S. P., Efstathiou G., Ellis R. S., Glazebrook K., Jackson C., Lahav O., Lewis I., Lumsden S., Maddox S., Madgwick D., Peacock J. A., Peterson B. A., Sutherland W., Taylor K., Team) T. d., 2002, Monthly Notices of the Royal Astronomical Society, 336, 907

Norman M. L., 2000 Vol. 9, Introducing ZEUS-MP: a 3D, parallel, multiphysics code for astrophysical fluid dynamics. pp 66–71

Oesch P. A., Bouwens R. J., Illingworth G. D., Carollo C. M., Franx M., Labbe I., Magee D., Stiavelli M., Trenti M., van Dokkum P. G., 2009, 0909.1806

Oesch P. A., Stiavelli M., Carollo C. M., Bergeron L. E., Koekemoer A. M., Lucas R. A., Pavlovsky C. M., Trenti M., Lilly S. J., Beckwith S. V. W., Dahlen T., Ferguson H. C., Gardner J. P., Lacey C., Mobasher B., Panagia N., Rix H., 2007, Astrophysical Journal, 671, 1212

Oh S. P., Haiman Z., 2002, Astrophysical Journal, 569, 558

Okabe A., Boots B., Sugihara K., Chiu S. N., 2000, Spatial tessellations: Concepts and applications of Voronoi diagrams, 2nd edn. Probability and Statistics, Wiley, NYC

Okamoto T., Gao L., Theuns T., 2008, Monthly Notices of the Royal Astronomical Society, 390, 920

Okamoto T., Jenkins A., Eke V. R., Quilis V., Frenk C. S., 2003, Monthly Notices of the Royal Astronomical Society, 345, 429

Oke J. B., Gunn J. E., 1983, Astrophysical Journal, 266, 713

Omukai K., 2000, The Astrophysical Journal, 534, 809

Omukai K., Nishi R., 1999, Astrophysical Journal, 518, 64

Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, 0807.3447

Osterbrock D. E., 1989, Astrophysics of gaseous nebulae and active galactic nuclei

Ostriker J. P., Peebles P. J. E., 1973, Astrophysical Journal, 186, 467

Ouchi M., Mobasher B., Shimasaku K., Ferguson H. C., Fall S. M., Ono Y., Kashikawa N., Morokuma T., Nakajima K., Okamura S., Dickinson M., Giavalisco M., Ohta K., 2009, Astrophysical Journal, 706, 1136

Ouchi M., Shimasaku K., Akiyama M., Simpson C., Saito T., Ueda Y., Furusawa H., Sekiguchi K., Yamada T., Kodama T., Kashikawa N., Okamura S., Iye M., Takata T., Yoshida M., Yoshida M., 2008, Astrophysical Journal Supplement Series, 176, 301

Paardekooper J. P., Kruip C. J. H., Icke V., 2009, 0912.4273

Parkinson H., Cole S., Helly J., 2008, Monthly Notices of the Royal Astronomical Society, 383, 557

Partridge R. B., Peebles P. J. E., 1967a, Astrophysical Journal, 147, 868

Partridge R. B., Peebles P. J. E., 1967b, Astrophysical Journal, 148, 377

Pawlik A. H., Schaye J., 2008, Monthly Notices of the Royal Astronomical Society, 389, 651

Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, Monthly Notices of the Royal Astronomical Society, 394, 1812

Peacock J. A., 1999, Cosmological Physics

Pearce F. R., Jenkins A., Frenk C. S., Colberg J. M., White S. D. M., Thomas P. A., Couchman H. M. P., Peacock J. A., Efstathiou G., Consortium T. V., 1999, Astrophysical Journal, 521, L99

Peebles P. J. E., 1969, Astrophysical Journal, 155, 393

Peebles P. J. E., 1993, Principles of physical cosmology

Pelupessy F. I., Schaap W. E., van de Weygaert R., 2003, Astronomy and Astrophysics, 403, 389

Pen U., 1998, Astrophysical Journal Supplement Series, 115, 19

Perlmutter S., Aldering G., Goldhaber G., Knop R. A., Nugent P., Castro P. G., Deustua S., Fabbro S., Goobar A., Groom D. E., Hook I. M., Kim A. G., Kim M. Y., Lee J. C., Nunes N. J., Pain R., Pennypacker C. R., Quimby R., Lidman C., Ellis R. S., Irwin M., McMahon R. G., Ruiz-Lapuente P., Walton N., Schaefer B., Boyle B. J., Filippenko A. V., Matheson T., Fruchter A. S., Panagia N., Newberg H. J. M., Couch W. J., Project T. S. C., 1999, Astrophysical Journal, 517, 565

Petkova M., Springel V., 2009, Monthly Notices of the Royal Astronomical Society, 396, 1383

Power C., Knebe A., 2006, Monthly Notices of the Royal Astronomical Society, 370, 691

Pozzetti L., Cimatti A., Zamorani G., Daddi E., Menci N., Fontana A., Renzini A., Mignoli M., Poli F., Saracco P., Broadhurst T., Cristiani S., D'Odorico S., Giallongo E., Gilmozzi R., 2003, Astronomy and Astrophysics, 402, 837

Press W. H., Schechter P., 1974, Astrophysical Journal, 187, 425

Prochaska J. X., 1999, Astrophysical Journal, 511, L71

Quinn T., Katz N., Efstathiou G., 1996, Monthly Notices of the Royal Astronomical Society, 278, L49

Quirk J. J., 1994, International Journal for Numerical Methods in Fluids, 18, 555

Raga A. C., Navarro-Gonzlez R., Villagrn-Muniz M., 2000, Revista Mexicana de Astronomia y Astrofisica, 36, 67

Rauch M., 1998, Annual Review of Astronomy and Astrophysics, 36, 267

Razoumov A. O., Cardall C. Y., 2005, Monthly Notices of the Royal Astronomical Society, 362, 1413

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

Razoumov A. O., Sommer-Larsen J., 2007, Astrophysical Journal, 668, 674

Razoumov A. O., Sommer-Larsen J., 2009, 0903.2045

Reed D., Bower R., Frenk C., Jenkins A., Theuns T., 2006, astro-ph/0607150

Rees M. J., 1986, Monthly Notices of the Royal Astronomical Society, 222, 27P

Rees M. J., Ostriker J. P., 1977, Monthly Notices of the Royal Astronomical Society, 179, 541

Reynolds D. R., Hayes J. C., Paschos P., Norman M. L., 2009, 0901.1110

Ricotti M., 2002, Monthly Notices of the Royal Astronomical Society, 336, L33

Riess A. G., Filippenko A. V., Challis P., Clocchiatti A., Diercks A., Garnavich P. M., Gilliland R. L., Hogan C. J., Jha S., Kirshner R. P., Leibundgut B., Phillips M. M., Reiss D., Schmidt B. P., Schommer R. A., Smith R. C., Spyromilio J., Stubbs C., Suntzeff N. B., Tonry J., 1998, Astronomical Journal, 116, 1009

Rijkhorst E., , 2005, Numerical nebulae, http://adsabs.harvard.edu/abs/2005PhDT.....24R

Rijkhorst E., Plewa T., Dubey A., Mellema G., 2006, Astronomy and Astrophysics, 452, 907

Ritzerveld J., Icke V., Rijkhorst E., , 2003, Triangulating Radiation: Radiative Transfer on Unstructured Grids, http://adsabs.harvard.edu/abs/2003astro.ph.12301R

Ritzerveld N. G. H., , 2007, The simplicity of transport: triangulating the first light, http://adsabs.harvard.edu/abs/2007PhDT.....1R

Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics

Ryu D., Ostriker J. P., Kang H., Cen R., 1993, Astrophysical Journal, 414, 1

Salpeter E. E., 1955, Astrophysical Journal, 121, 161

Santos M. G., Cooray A., Haiman Z., Knox L., Ma C., 2003, Astrophysical Journal, 598, 756

Sargent W. L. W., Steidel C. C., Boksenberg A., 1989, Astrophysical Journal Supplement Series, 69, 703

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 552

Schaap W., 2007, PhD thesis, Rijksuniversiteit Groningen, The Netherlands

Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, Monthly Notices of the Royal Astronomical Society, 318, 817

Schaye J., Vecchia C. D., Booth C. M., Wiersma R. P. C., Theuns T., Haas M. R., Bertone S., Duffy A. R., McCarthy I. G., van de Voort F., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1536

Schechter P., 1976, Astrophysical Journal, 203, 297

Schmidt-Voigt M., Koeppen J., 1987, Astronomy and Astrophysics, 174, 211

Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, Astrophysical Journal, 571, 30

Scott D., Rees M. J., 1990, Monthly Notices of the Royal Astronomical Society, 247, 510

Semelin B., Combes F., Baek S., 2007, Astronomy and Astrophysics, 474, 365

Shapiro P. R., Giroux M. L., 1987, Astrophysical Journal, 321, L107

Shapiro P. R., Giroux M. L., Babul A., 1994, Astrophysical Journal, 427, 25

Shapiro P. R., Iliev I. T., Raga A. C., 2004, Monthly Notices of the Royal Astronomical Society, 348, 753

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

Sharp D., 1984, Physica D: Nonlinear Phenomena, 12, 3

Sheth R. K., Tormen G., 1999, Monthly Notices of the Royal Astronomical Society, 308, 119

Sheth R. K., Tormen G., 2002, Monthly Notices of the Royal Astronomical Society, 329, 61

Shin M., Trac H., Cen R., 2008, Astrophysical Journal, 681, 756

Silk J., 1977, Astrophysical Journal, 211, 638

Silva L., Granato G. L., Bressan A., Danese L., 1998, Astrophysical Journal, 509, 103

Sirko E., 2005, Astrophysical Journal, 634, 728

Smail I., Ivison R. J., Blain A. W., 1997, Astrophysical Journal, 490, L5

Sod G. A., 1978, Journal of Computational Physics, 27, 1

Sofue Y., Rubin V., 2001, Annual Review of Astronomy and Astrophysics, 39, 137

Sokasian A., Abel T., Hernquist L., Springel V., 2003, Monthly Notices of the Royal Astronomical Society, 344, 607

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

Somerville R. S., 2002, Astrophysical Journal, 572, L23

Somerville R. S., Kolatt T. S., 1999, Monthly Notices of the Royal Astronomical Society, 305, 1

Sparke L. S., Gallagher J. S., 2007, Galaxies in the Universe: An Introduction

Spergel D. N., Bean R., Dor O., Nolta M. R., Bennett C. L., Dunkley J., Hinshaw G., Jarosik N., Komatsu E., Page L., Peiris H. V., Verde L., Halpern M., Hill R. S., Kogut A., Limon M., Meyer S. S., Odegard N., Tucker G. S., Weiland J. L., Wollack E., Wright E. L., 2007, Astrophysical Journal Supplement Series, 170, 377

Spitzer L., 1978, Physical processes in the interstellar medium

Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

Springel V., 2009, 0901.4107

Springel V., White S. D. M., Jenkins A., Frenk C. S., Yoshida N., Gao L., Navarro J., Thacker R., Croton D., Helly J., Peacock J. A., Cole S., Thomas P., Couchman H., Evrard A., Colberg J., Pearce F., 2005, Nature, 435, 629

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Monthly Notices of the Royal Astronomical Society, 328, 726

Srbinovsky J. A., Wyithe J. S. B., 2007, Monthly Notices of the Royal Astronomical Society, 374, 627

Stark D. P., Bunker A. J., Ellis R. S., Eyles L. P., Lacy M., 2007, Astrophysical Journal, 659, 84

Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, Astrophysical Journal, 519, 1

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

Steidel C. C., Dickinson M., Persson S. E., 1994, Astrophysical Journal, 437, L75

Steidel C. C., Giavalisco M., Dickinson M., Adelberger K. L., 1996, Astronomical Journal, 112, 352

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

Steidel C. C., Hamilton D., 1992, Astronomical Journal, 104, 941

Steidel C. C., Hamilton D., 1993, Astronomical Journal, 105, 2017

Steidel C. C., Pettini M., Hamilton D., 1995, Astronomical Journal, 110, 2519

Stiavelli M., Scarlata C., Panagia N., Treu T., Bertin G., Bertola F., 2001, Astrophysical Journal, 561, L37

Stoyan D., Kendall W. S., Mecke J., 1996, Stochastic Geometry and Its Applications, 2nd edn. Wiley

Strelnitski V., 1997, Astrophysics and Space Science, 252, 279

Strömgren B., 1939, Astrophysical Journal, 89, 526

Sunyaev R. A., Chluba J., 2008 Vol. 395, The richness and beauty of the physics of cosmological recombination. p. 35

Susa H., 2006, Publications of the Astronomical Society of Japan, 58, 445

Susa H., Umemura M., 2004a, Astrophysical Journal, 610, L5

Susa H., Umemura M., 2004b, Astrophysical Journal, 600, 1

Sutherland R. S., Dopita M. A., 1993, Astrophysical Journal Supplement Series, 88, 253

Taniguchi Y., Ajiki M., Nagao T., Shioya Y., Murayama T., Kashikawa N., Kodaira K., Kaifu N., Ando H., Karoji H., Akiyama M., Aoki K., Doi M., Fujita S. S., Furusawa H., Hayashino T., Iwamuro F., Iye M., Kobayashi N., Kodama T., Komiyama Y., Matsuda Y., Miyazaki S., Mizumoto Y., Morokuma T., Motohara K., Nariai K., Ohta K., Ohyama Y., Okamura S., Ouchi M., Sasaki T., Sato Y., Sekiguchi K., Shimasaku K., Tamura H., Umemura M., Yamada T., Yasuda N., Yoshida M., 2005, Publications of the Astronomical Society of Japan, 57, 165

Tasker E. J., Brunino R., Mitchell N. L., Michielsen D., Hopton S., Pearce F. R., Bryan G. L., Theuns T., 2008, Monthly Notices of the Royal Astronomical Society, 390, 1267

Tegmark M., Silk J., Evrard A., 1993, Astrophysical Journal, 417, 54

Teyssier R., 2002, Astronomy and Astrophysics, 385, 337

Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, Monthly Notices of the Royal Astronomical Society, 301, 478

Thomas R. M., Zaroubi S., Ciardi B., Pawlik A. H., Labropoulos P., Jeli V., Bernardi G., Brentjens M. A., de Bruyn A. G., Harker G. J. A., Koopmans L. V. E., Mellema G., Pandey V. N., Schaye J., Yatawatta S., 2009, Monthly Notices of the Royal Astronomical Society, 393, 32

Thompson R. I., Illingworth G., Bouwens R., Dickinson M., Eisenstein D., Fan X., Franx M., Riess A., Rieke M. J., Schneider G., Stobie E., Toft S., van Dokkum P., 2005, Astronomical Journal, 130, 1

Thoul A. A., Weinberg D. H., 1996, Astrophysical Journal, 465, 608

Tinsley B. M., 1972, Astrophysical Journal, 178, 319

Tittley E. R., Meiksin A., 2007, Monthly Notices of the Royal Astronomical Society, 380, 1369

Toomre A., 1964, Astrophysical Journal, 139, 1217

Toomre A., 1977, Annual Review of Astronomy and Astrophysics, 15, 437

Toomre A., 1981 What amplifies the spirals. pp 111–136

Trac H., Cen R., 2007, Astrophysical Journal, 671, 1

Trac H., Cen R., Loeb A., 2008, Astrophysical Journal, 689, L81

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

Trac H., Pen U., 2004, New Astronomy, 9, 443

Tytler D., 1982, Nature, 298, 427

Valageas P., Silk J., 1999, Astronomy and Astrophysics, 347, 1

van de Weygaert R., , 2007, Voronoi Tessellations and the Cosmic Web: Spatial Patterns and Clustering across the Universe, http://adsabs.harvard.edu/abs/2007arXiv0707.2877V

Vishniac E. T., 1983, Astrophysical Journal, 274, 152

Walker I. R., Mihos J. C., Hernquist L., 1996, Astrophysical Journal, 460, 121

Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, Astrophysical Journal, 646, 881

Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, Astrophysical Journal, 399, 405

Warren M. S., Salmon J. K., 1993, in Proceedings of the 1993 ACM/IEEE conference on Supercomputing A parallel hashed Oct-Tree n-body algorithm. ACM, Portland, Oregon, United States, pp 12–21

Weinberg D. H., Hernquist L., Katz N., 1997, Astrophysical Journal, 477, 8

Weinmann S. M., Macci A. V., Iliev I. T., Mellema G., Moore B., 2007, Monthly Notices of the Royal Astronomical Society, 381, 367

Whalen D., Abel T., Norman M. L., 2004, The Astrophysical Journal, 610, 14

Whalen D., Norman M. L., 2008, Astrophysical Journal, 673, 664

White S. D. M., 1984, Astrophysical Journal, 286, 38

White S. D. M., Frenk C. S., 1991, Astrophysical Journal, 379, 52

White S. D. M., Rees M. J., 1978, Monthly Notices of the Royal Astronomical Society, 183, 341

Wiersma R. P. C., Schaye J., Smith B. D., 2009, Monthly Notices of the Royal Astronomical Society, 393, 99

Wilkins S. M., Bunker A. J., Ellis R. S., Stark D., Stanway E. R., Chiu K., Lorenzoni S., Jarvis M. J., , 2009, Probing $\sim L_*$ Lyman-break Galaxies at $z\approx 7$ in GOODS-South with WFC3 on HST, http://adsabs.harvard.edu/abs/2009arXiv0910.1098W

Williams R. E., Blacker B., Dickinson M., Dixon W. V. D., Ferguson H. C., Fruchter A. S., Giavalisco M., Gilliland R. L., Heyer I., Katsanis R., Levay Z., Lucas R. A., McElroy D. B., Petro L., Postman M., Adorf H., Hook R., 1996, Astronomical Journal, 112, 1335

Williams R. J. R., 1999, Monthly Notices of the Royal Astronomical Society, 310, 789

Wise J. H., Abel T., 2007, 0704.3629

Wise J. H., Cen R., 2009, Astrophysical Journal, 693, 984

Wolfe A. M., Gawiser E., Prochaska J. X., 2005, Annual Review of Astronomy and Astrophysics, 43, 861

Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D., 1986, Astrophysical Journal Supplement Series, 61, 249

Woodward P., Colella P., 1984, Journal of Computational Physics, 54, 115

Wyithe J. S. B., Loeb A., 2003, Astrophysical Journal, 586, 693

Xu G., 1997, Monthly Notices of the Royal Astronomical Society, 288, 903

Xue X. X., Rix H. W., Zhao G., Fiorentin P. R., Naab T., Steinmetz M., van den Bosch F. C., Beers T. C., Lee Y. S., Bell E. F., Rockosi C., Yanny B., Newberg H., Wilhelm R., Kang X., Smith M. C., Schneider D. P., 2008, Astrophysical Journal, 684, 1143

York D. G., Adelman J., Anderson J. E., Anderson S. F., Annis J., Bahcall N. A., Bakken J. A., Barkhouser R., Bastian S., Berman E., Boroski W. N., Bracker S., Briegel C., Briggs J. W., Brinkmann J., Brunner R., Burles S., Carey L., Carr M. A., Castander F. J., Chen

B., Colestock P. L., Connolly A. J., Crocker J. H., Csabai I., Czarapata P. C., Davis J. E., Doi M., Dombeck T., Eisenstein D., Ellman N., Elms B. R., Evans M. L., Fan X., Federwitz G. R., Fiscelli L., Friedman S., Frieman J. A., Fukugita M., Gillespie B., Gunn J. E., Gurbani V. K., de Haas E., Haldeman M., Harris F. H., Hayes J., Heckman T. M., Hennessy G. S., Hindsley R. B., Holm S., Holmgren D. J., hao Huang C., Hull C., Husby D., Ichikawa S., Ichikawa T., eljko Ivezi Kent S., Kim R. S. J., Kinney E., Klaene M., Kleinman A. N., Kleinman S., Knapp G. R., Korienek J., Kron R. G., Kunszt P. Z., Lamb D. Q., Lee B., Leger R. F., Limmongkol S., Lindenmeyer C., Long D. C., Loomis C., Loveday J., Lucinio R., Lupton R. H., MacKinnon B., Mannery E. J., Mantsch P. M., Margon B., McGehee P., McKay T. A., Meiksin A., Merelli A., Monet D. G., Munn J. A., Narayanan V. K., Nash T., Neilsen E., Neswold R., Newberg H. J., Nichol R. C., Nicinski T., Nonino M., Okada N., Okamura S., Ostriker J. P., Owen R., Pauls A. G., Peoples J., Peterson R. L., Petravick D., Pier J. R., Pope A., Pordes R., Prosapio A., Rechenmacher R., Quinn T. R., Richards G. T., Richmond M. W., Rivetta C. H., Rockosi C. M., Ruthmansdorfer K., Sandford D., Schlegel D. J., Schneider D. P., Sekiguchi M., Sergey G., Shimasaku K., Siegmund W. A., Smee S., Smith J. A., Snedden S., Stone R., Stoughton C., Strauss M. A., Stubbs C., SubbaRao M., Szalay A. S., Szapudi I., Szokoly G. P., Thakar A. R., Tremonti C., Tucker D. L., Uomoto A., Berk D. V., Vogeley M. S., Waddell P., i. Wang S., Watanabe M., Weinberg D. H., Yanny B., Yasuda N., 2000, Astronomical Journal, 120, 1579

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

Zahn O., Lidz A., McQuinn M., Dutta S., Hernquist L., Zaldarriaga M., Furlanetto S. R., 2007, Astrophysical Journal, 654, 12

Zaldarriaga M., 1997, Physical Review D, 55, 1822

Zaninetti L., 1991, Astronomy and Astrophysics, 246, 291

Zhang P., Pen U., Trac H., 2004, Monthly Notices of the Royal Astronomical Society, 355, 451

Zwicky F., 1933, Helvetica Physica Acta, 6, 110