Cosmological RHD simulations of early galaxy formation

Karl Joakim Rosdahl

Université Claude Bernard Lyon 1

UNIVERSITÉ DE LYON

N° d'ordre

Année 2012

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE PHYSIQUE ET ASTROPHYSIQUE

DIPLOME DE DOCTORAT

(arrêté du 7 août 2006)

soutenue publiquement le 15 Juin 2012

par

M. Karl Joakim ROSDAHL

TITRE :

Cosmological RHD simulations of early galaxy formation

Directeur de thèse :
M. Jérémy BLAIZOT

JURY : M. Andrea FERRARA
M. Benoit SEMELIN
M. Julien DEVRIENDT
M. Jean-François GONZALEZ
M. Bruno GUIDERDONI
M. Christophe PICHON
M. Romain TEYSSIER

Abstract

With the increasing sophistication and efficiency of cosmological hydrodynamics codes, it has become viable to include ionizing radiative transfer (RT) in cosmological simulations, either in post-processing or in full-blown radiation-hydrodynamics (RHD) simulations. In spite of the many hurdles involved, there has been much activity during the last decade or so on different strategies and implementations, because a number of interesting problems can be addressed with RT and RHD, e.g. how and when the Universe became reionized, how radiation from stars and active galactic nuclei plays a part in regulating structure formation on small and large scales, and what predictions and interpretations we can make of observed phenomena such as the Lyman-alpha forest and diffuse sources of radiation.

This coincides with the advent of the James Webb space telescope (JWST) and other state-of-the-art instruments which are about to give us an unprecedented glimpse into the end of the dark ages of the Universe, when the cosmos switched from a cold and neutral state to a hot and ionized one, due to the turn-on of ionizing radiative sources.

With a primary interest in the problem of radiative feedback in early structure formation, we have implemented an RHD version of the `Ramses` cosmological code we call `RamsesRT`, which is moment based and employs the local M1 Eddington tensor closure. This code allows us to study the effects of ionizing radiation on-the-fly in cosmological RHD simulations that take full advantage of the adaptive mesh refinement and parallelization strategies of `Ramses`. For self-consistent RHD we have also implemented a non-equilibrium chemistry of the atomic hydrogen and helium species that interact with the transported radiation.

I present in this thesis an extensive description of the `RamsesRT` implementation and numerous tests to validate it.

Thus far we have used the RHD implementation to study extended line emission from accretion streams, which are routinely predicted to exist at early redshift by cosmological simulations but have never been unambiguously verified by observations, and to investigate whether gravitational heating in those streams could be the dominant power source of so-called Lyman-alpha blobs, an observed phenomenon which has been much studied and debated during the last decade or two. Our conclusions from this investigation are that Lyman-alpha blobs can in principle be powered by gravitational heating, and furthermore that accretion streams are on the verge of being directly detectable for the first time with upcoming instruments.

My future intent is to use `RamsesRT` for high-resolution cosmological zoom simulations of early galaxy formation, up to the epoch of reionization, to study how radiative feedback affects the formation and evolution of those galaxies and to make observational predictions that can be tested with upcoming instruments such as the JWST.

d

RHD simulations cosmologiques de la formation des premieres galaxies

Résumé

Avec l'essor actuel de la sophistication et de l'efficacité des codes de cosmologie hydrodynamique, il est devenu possible d'inclure le transfert radiatif (RT) des photons ionisants dans les simulations cosmologiques, soit en post-traitement, soit en simulations couplées rayonnement+hydrodynamique (RHD). Malgré de nombreux obstacles, il y a eu cette derniére décennie beaucoup de recherches menées sur les différentes stratégies et implémentations, dû au fait qu'un nombre de problèmes intéressants peuvent être désormais abordés par la RT et RHD, par exemple comment et quand l'Univers s'est réionisé, comment l'émission radiative des étoiles et des noyaux actifs de galaxies se comportent pour réguler la formation des structures à des échelles petites et grandes, et quelles prédictions et interprétations nous pouvons faire des phénomènes observés, tels que la forêt Lyman-alpha et des sources diffuses de rayonnement.

Cela coïncide avec l'avènement du télescope spatial James Webb (JWST) et d'autres instruments de pointe qui sont sur le point de nous donner un aperçu sans précédent sur la fin des âges sombres de l'Univers, quand le cosmos est passé d'un état froid et neutre à un état chaud et ionisé, à la suite de l'apparition des sources radiatives.

Notre préoccupation principale étant les rétroactions radiatives des premieres structures, nous avons mis en place une version RHD du code cosmologique Ramses, que nous appelons RamsesRT, basée sur la méthode des moments. Ce code nous permet d'étudier les effets du rayonnement ionisant dans les simulations cosmologiques RHD qui tirent pleinement profit des stratégies de raffinement adaptif de grille et de parallélisation de Ramses. Pour rendre auto-cohérent le RHD nous avons également mis en œuvre une thermochimie hors-équilibre incluant des espèces de l'Hydrogène et de l'Hélium qui interagissent avec le rayonnement transporté.

Je présente dans cette thèse une description détaillée de RamsesRT et de nombreux tests contribuant à sa validation.

Jusqu'à présent nous avons utilisé RamsesRT pour étudier l'émission Lyman-alpha de courants d'accrétion, qui sont prédits à grand redshift par les simulations cosmologiques, mais n'ont jamais été clairement identifiés par les observations. Nous avons également étudié le chauffage gravitationnel dans ces courants pour déterminer si ce dernier pouvait être la source motrice principale des Lyman-alpha blobs, un phénomène observé qui a été beaucoup étudié et débattu au cours de la dernière décennie. Cet étudie nous permet de conclure que les Lyman-alpha blobs peuvent, en principe, être alimentés par le chauffage gravitationnel, et que d'autre part, les courants d'accrétion sont sur le point d'être directement détectables avec des instruments à venir.

Mes intentions futures sont d'utiliser RamsesRT dans les simulations cosmologiques à haute résolution, de la formation des premiéres galaxies jusqu'à l'époque de la réionisation, et ainsi étudier comment la rétroaction radiative affecte la formation et l'évolution de ces galaxies et de faire des prévisions d'observation qui peuvent être testées avec des instruments sophistiqués tels que le JWST.

Discipline
Physique et Astrophysique

Key words
radiative transfer – radiation-hydrodynamics – methods: numerical – galaxies: formation, evolution – cosmology: dark ages, reionization – diffuse radiation – large-scale structure of the Universe.

Mots-clés
transfert radiatif – rayonnement+hydrodynamique – méthodes: numérique – galaxies: formation, évolution – cosmologie: âges sombres, réionisation – rayonnement diffus – structure à grande échelle de l'Univers.

Intitule et adresse du laboratoire
CRAL - Centre de Recherche Astrophysique de Lyon
Observatoire de Lyon
9, avenue Charles Andre
69561 Saint Genis-Laval cedex

Acknowledgements

It's been a great three years and I have gotten to know many wonderful people. First of all, kisses to the ELIXIR kids who have shared this experience and some wonderful times in various contries. It's been a pleasure knowing you and I'm sure this is not the end of it. Many thanks to Stephan Charlot for organizing the whole thing.

Computational astrophysics is obviously not an occupation best fitted to maintaining a good physical condition or a healthy complexion. So a big thanks to my running partners for motivation and comradery in my attempts to keep in shape: Bernhard, Emeline, Florent, and Sam.

I am ever grateful to Jeremy Blaizot, for picking me and sticking with me. I'm extremely lucky in having a supervisor so open and enthusiastic, never obtrusive yet always available to tackle new problems (except that one time he hid under the table). Thanks, in equal measure, to Karma and Klaus Galsgaard for bringing our paths together.

Thanks to Romain Teyssier for providing me with permission, advice, and lots of ideas for messing around with his code. To Dominique Aubert for helping out with moment based RT, and to Stephanie Courty, Julien Devriendt and Yohan Dubois for navigating me through `Ramses` and making it all such a pleasant experience.

Thanks to all the nice people who have made me feel very welcome at the observatoire, despite my lack of linguistic abilities. Especially to my friends and fellow Galpac team members: Anne, Johan, Leo, Mamta, Manu, Maxim, Pierre, Sam, Steph, Thibault, Xavier. Special thanks to Bernhard Dorner for helping out with matters of layout and aesthetics for this thesis.

Last but not least, love and billions and billions of kisses to my family, Melo, Snorri and Benni, for bearing with me and even believing in me through all this. The Universe is super important and all, but you guys are much more important.

This work was funded by the Marie Curie Initial Training Network ELIXIR of the European Commission under contract PITN-GA-2008-214227.

h

Contents

List of Tables

n

List of Figures

1 Introduction

1.1 What do we know about the Universe?

Our picture of the Universe has evolved enormously in recent times. Only a hundred years ago, the existence of external galaxies was not known, and the Universe was thought to be an infinite, eternal and static configuration of massive bodies.

The basis for the cosmological model of the time was the somewhat speculative *cosmological principle*, which states that the Universe is homogeneous and isotropic, i.e. that all locations and directions are much the same. The principle seems to have originated in the philosophical idea that there is nothing special about our epoch or location in the Universe, perhaps in defiance to the view which had prevailed so long that we are the center of all. In fact, not only does the cosmological principle remove us from the center of the Universe, – it removes *any* center from the Universe. Despite the philosophical origin, the principle looks today to be well founded and is supported by a number of observational evidence [1]. The idea of a static and eternal Universe, however, is not, and indeed physicists of the early 20th century were becoming painfully aware that such a Universe is unstable and prone to gravitational collapse.

In the early 1920s, as the existence of external galaxies was being established, Alexander Friedmann constructed out of general relativity – and the cosmological principle – an equation stating how the Universe must expand or contract due to the energy (or equivalently mass) embedded in it, and how the curvature of space – spherical, flat, or saddle-shaped – is determined by its energy content.

Friedmann's equation initially went unnoticed, but a few years on, Edwin Hubble measured the velocity-distance relation for a sample of galaxies and showed that the Universe

[1] This is of course not apparent in our immediate surroundings, but on scales of about 100 Mpc or more, the Universe does indeed have a homogeneous structure.

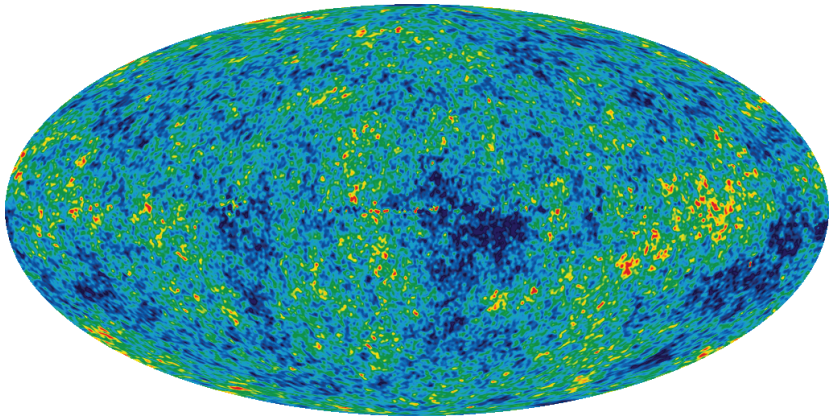


Figure 1.1: *Temperature fluctuations in the CMB, from the seven year WMAP data. Credit: NASA / WMAP Science Team.*

isn't in a static configuration, but is expanding (Hubble, 1929)[2]. The rate of expansion, known as the Hubble constant, can be derived by dividing the velocity by the distance, the best current estimate being

$$H_0 = 70 \pm 7 \text{ km/s/Mpc},$$

i.e. objects ℓ megaparsec apart, at any location in the current Universe, recede each other by about 70 ℓ km/s. If we assume the expansion to be constant in time and space, the age of the Universe can be extrapolated backwards to an origin, a 'Big Bang', by $1/H_0 \sim 14$ Gyr.

The discovery of expansion was a game-changer, and in its wake a number of cosmological models competed to explain it. The contest was settled in favour of the Big Bang model in 1964 when Penzias and Wilson accidentally discovered the cosmic microwave background (CMB) radiation (Penzias and Wilson, 1965).

The CMB is thermal radiation that fills our Universe. It is a perfect blackbody with a temperature of 2.7 K and comes very nearly uniformly from all directions, with spatial temperature fluctuations of only about one in 10^5 (see Fig. 1.1). The least contrived explanation for its existence is that it is radiation emitted from the surface of last scattering, about 400 thousand years after the big bang, when the Universe was $a \sim 1/1100$ of its current size[3]. At this point, the primordial fluid of photons, protons, neutrons and electrons had expanded and cooled enough that atoms formed, allowing the photons to decouple from the fluid and flow freely. Apart from directly implying its Big Bang origin, the CMB contains a wealth of information about the properties of Universe, which is to date being actively mined with increasingly detailed observations, the latest being the 7 year data from the Wilkinson microwave anisotropy probe (WMAP; Komatsu et al., 2011).

One piece of information constrains the curvature of the Universe: Spatial anisotropy

[2] Lemaître (1927) actually made the same observation two years ahead of Hubble, but made the mistake of announcing it in some obscure and unpronounceable language.

[3] The size of the universe is expressed in terms of the expansion factor a, which is defined to be equal to one at the current time. Distances, or equivalently epochs of the Universe are also commonly expressed in terms of how much light emitted at that epoch has been redshifted by the cosmological expansion when we observe it, i.e. we observe a wavelength shift in the light of $z = 1/a - 1$. The CMB is thus observed a redshift of $z \sim 1100$.

measurements of the CMB reveal the strongest temperature fluctuations to exist on scales of $\sim 1^\circ$ seen from us. The fluctuations represent mass/energy concentrations at the time of last scattering, which according to the theory should be strongest in gravitational potentials with sizes matching the Hubble distance – the largest distance at which two points can communicate. It so happens that this distance at the time of last scattering is also $\sim 1^\circ$, with the caveat that the Universe has no curvature on the large scale. Thus CMB measurements demonstrate that the Universe is flat, or very close to being so. With this knowledge of flatness, the energy density of the Universe is trivially derived from the Friedmann equation.

Furthermore, with the Big Bang theory set in place, the baryonic content of the Universe can be precisely estimated from *primordial nucleosynthesis*, a time-line of chemical reactions up to about ten minutes after the Big Bang, from which the relative abundances of the primordial elements can be derived. This reveals a neutron-to-proton ratio of 0.15, a helium mass fraction of $Y = 0.27$, a hydrogen mass fraction of $X \approx 1 - Y$, and traces of heavier elements, Li and Be. In other words, the baryonic component of the primordial Universe is composed of 27% (by mass) helium, 73% hydrogen and just a tiny little bit of lithium and beryllium. These ratios depend on the baryon-to-photon fraction of the Universe: Detailed observations of elemental mass ratios in primordial gas clouds, along with the photon energy density of the Universe, which is well known from the CMB, thus allows one to determine the baryon number density in the Universe

$$n_b \approx \frac{2.3\ 10^{-7}}{a^3}\ \mathrm{cm}^{-3},$$

where a is the cosmological expansion factor. From this the *current* baryon mass fraction in the Universe can be estimated to be

$$\Omega_{b,0} \approx 0.04,$$

i.e. baryons comprise only about 4% of the total mass in the universe. The baryon mass fraction can be independently derived from the magnitudes of the temperature fluctuations in the CMB, and astonishingly this estimate arrives at exactly the same mass fraction!

What about the other 96%, then? Observations of galactic rotation curves and velocity dispersions in galaxy clusters consistently indicate that only a small fraction of the mass in the galactic environments is accounted for by luminous baryons, i.e. stars. The missing mass, which must dominate the gravitational potentials to explain these observations has become known as *dark matter* (DM). There is more evidence which support its existence: Observations of gravitational lensing indicate plenty of unobserved mass, and cosmological simulations are not able to create a Universe with galaxies by including only the baryon mass; the gravitational potentials are too weak to form galaxies in a Hubble time.

The latest dramatic development in Cosmology was the discovery, made from precise measurements of distant type Ia supernovae, that the Universe is not only expanding, but accelerating in the expansion (Riess et al., 1998; Perlmutter et al., 1999). This requires the existence of a positive pressure or repulsive force in the Universe, which has been coined the cosmological constant, Λ, also often referred to as *dark energy*.

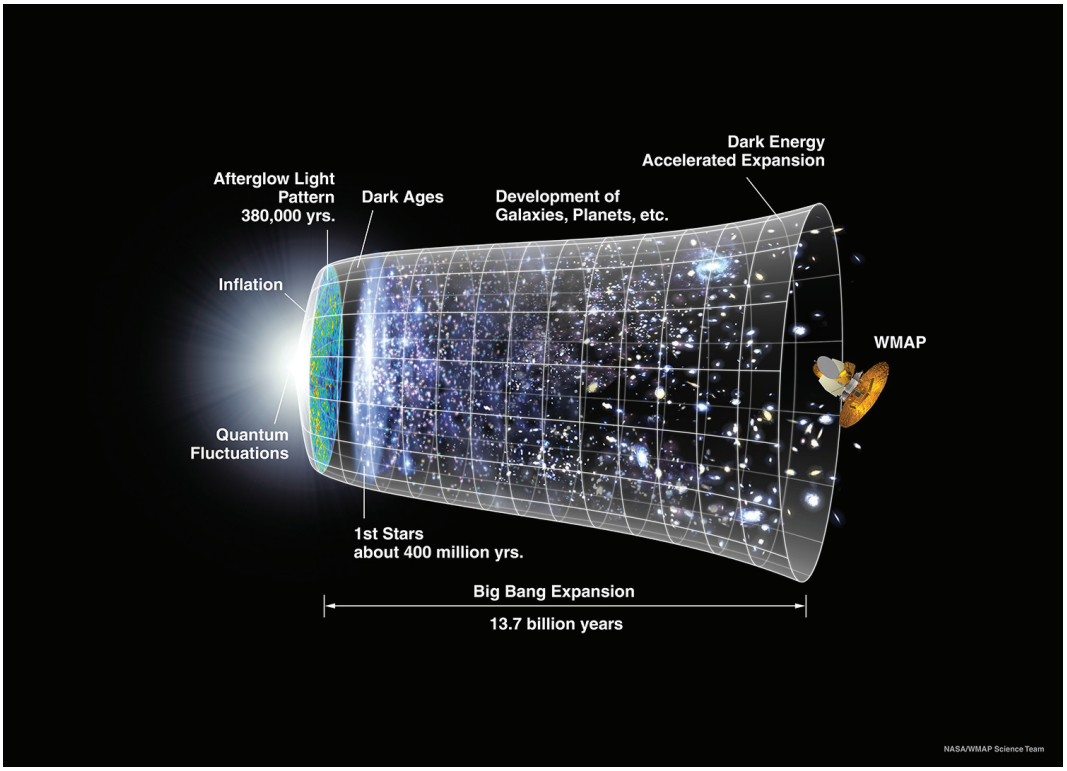


Figure 1.2: *Timeline of the Universe, with the WMAP satellite looking through almost the whole history so far at the CMB. Credit: NASA / WMAP Science Team.*

The ΛCDM model of the Universe

With the joint data of said supernovae measurements and state-of-the-art CMB observations with WMAP (Komatsu et al., 2011), the relative content of each of these components has been constrained, in a model referred to as the Λ-cold-dark-matter (ΛCDM) model, *cold* referring to the dark matter moving at non-relativistic speeds. In the model, the current values of the density parameters (fractions) for Λ, dark matter and baryons are, respectively,

$$\Omega_{\Lambda,0} = 0.73, \quad \Omega_{DM,0} = 0.23, \quad \Omega_{b,0} = 0.04,$$

plus tiny contributions from neutrinos and photons.

These ratios are not constant in time as the different components dilute differently with the cosmological expansion: The baryons and dark matter as a^{-3}, photons as a^{-4} and Λ does not dilute at all. Thus there was an earlier period where matter dominated the density and a very early period where it was dominated by photons.

So here is a summary of the history of the Universe, also shown in Fig. 1.2: 13.7 billion years ago, there happened for some unexplained reason a Big Bang where space and time came into existence, followed by a very quick and speculative period of inflation, where the size of the Universe grew by something like factor of 10^{78} (-ish) in volume. It was initially populated with a one-fluid of photons and primordial particles that combined into a hierarchy of composite particles as it diluted along with the expansion of space and cooled adiabatically. About 400 thousand years later, atoms were formed, transitioning the fluid

from an ionized to neutral state, at which point the photons decoupled from the fluid and could freely propagate through space (hence the CMB). Now began a period called the *dark ages*, when the only radiation was that of the CMB, rapidly diluting as the Universe further expanded. During this period, the small scale fluctuations in the dark matter density field grew, creating gravitational wells into which the gas could fall – now free of the counter-acting pressure from the photons – and form the seeds of the first galaxies, which grew in size via continued accretion of gas and mergers, and presumably started forming the first generation of stars. A vital ingredient in all this is radiative cooling, which allowed the condensing gas, from the galactic scales down to the stellar interiors, to relieve the adiabatic pressure buildup by emitting photons. At some point, the Universe became filled again with fresh photons, and as a consequence transitioned back to an ionized state. The most obvious evidence that this *reionization* happened is the fact that the current universe that surrounds us is to a very large degree ionized. After reionization, the galaxies have kept growing in size via mergers and accretion up to the current epoch.

1.2 The frontier – reionization and end of the dark ages

We have today an arsenal of ground-based and orbiting telescopes which tell us a lot about the evolution of the Universe since the epoch of reionization, providing many validations of the ΛCDM model of the Universe and allowing us to understand many cosmological and astrophysical phenomena. But here lies the frontier in astrophysics: Around redshift 6, darkness overtakes, and much of what happened between there and last scattering is shrouded in mystery, for the simple reasons that it is extremely distant and very redshifted, and the sources of radiation are fewer and dimmer than in the current Universe. We thus have little information about what went on during the dark ages; how the early galaxies formed and evolved, how the first stars formed and what their properties were, and how the photons were produced that reionized the Universe, that may or may not have come from these first stars – all of which are factors that determine the history of reionization.

Time of reionization

There exist some (and somewhat contradictory) observational constraints on the time of reionization:

We can look back in time and see that the Universe has been mostly ionized since at least $z \sim 6$, which is roughly 1 Gyr after the Big Bang. Scarce observations of extremely distant AGN sources show so-called Gunn-Peterson troughs, total absorption of radiation at the Lyman-alpha (Lyα) frequency over a large and continuous range of redshifted frequencies, which indicate that reionization was complete around redshift 6 (Fan et al., 2006).

WMAP measurements of polarization in the CMB radiation, which is the product of electron (Thomson) scattering, however suggests reionization at $z \sim 10.5$ (Larson et al., 2011), and Thomson optical depths which are difficult to consolidate with the Fan et al. (2006) results (e.g. Aubert and Teyssier, 2010). Also, somewhat in contradiction, measurements of the temperature in the lower-redshift inter-galactic medium (IGM) and considerations of its temperature history suggest intermediate reionization redshifts at $z < 9$ (Theuns et al., 2002)

and $z \gtrsim 8$ (Raskutti et al., 2012).

Reionization thus likely happened gradually over a large redshift range. Furlanetto and Briggs (2004) even suggest phases of reionization by different sources, with in-between periods of relapse.

The sources of reionization

Since the epoch of reionization appears to coincide with the appearance of galaxies, it is natural to think that the Universe was reionized by the first stars, but it is by no means obvious that the first stars radiated sufficiently to reionize the Universe as early as $z \sim 10$, as the average star formation rate in the Universe is found to decrease too quickly with redshift (Bunker et al., 2010). The efficiency of star-formation at this epoch is not well known and neither is the escape fraction of ionizing photons: Observations of low-redshift galaxies (and our own) give escape fractions of $f_{esc} \lesssim 2 - 10\%$ (see references in Razoumov and Sommer-Larsen, 2010), while $f_{esc} \gtrsim 10\%$ is needed for reionization consistent with WMAP data ($z \sim 10$), assuming the star formation efficiency is not dramatically different at early redshift. A number of simulation works have been undertaken to study how escape fractions vary with e.g. redshift, star formation rate and halo mass, with non-conforming results (Gnedin et al., 2008; Wise and Cen, 2009; Razoumov and Sommer-Larsen, 2010; Yajima et al., 2011a,b).

Most works suggest dwarf galaxies to be the likeliest predominant sources of the ionizing radiation (though see Gnedin et al., 2008), but very little observational evidence exists about those at such high redshifts, and even in the closer Universe it is hard to spot primordial dwarfs because of their low luminosity. Alternative models have also been devised for the sources of reionization. Quasars have been considered as a candidate: They dominate the UV background in the Universe at $z \gtrsim 3$, but they are generally held to be too rare at higher redshift (Richards et al., 2006) – it is though perhaps premature to completely rule out quasars as an important factor in reionization, as very high redshift quasars are starting to emerge in observations (Maiolino et al., 2012). Other possibilities that have been suggested are e.g. dark matter annihilation (Belikov and Hooper, 2009), radiation from primordial black holes (Gibilisco, 1996) and radiation from shocks in gas accreting onto galaxies (Dopita et al., 2011), though these more exotic scenarios lack observational support and are thus not taken very seriously (Kuhlen and Faucher-Giguere, 2012).

1.3 An upcoming peek into the dark ages

Where there is a frontier there are also pioneers, and surely a new generation of instruments is emerging which will give an unprecedented peek into the reionization epoch and the dark ages.

The James Webb Space Telescope (JWST, Gardner et al., 2006) is the successor to the Hubble Space Telescope, designed to observe the assembly of stars and galaxies at high redshift. It is optimized to observe in infrared the radiation emitted in visible and UV at $z \gtrsim 6$. It has the objectives of providing us with information about star-formation rates, gas content, even morphologies of galaxies emerging from the reionization epoch, of identifying

the first luminous sources in the Universe, and determining the ionization history. The observations of these early galaxies will hopefully provide missing pieces in the puzzle of the hierarchical assembly of galaxies and clusters, and explain much about the statistics and morphologies of current day structures. The JWST is currently set to launch in 2018.

LOFAR (Low Frequency Array) is a radio telescope that came online in 2010, based in the Netherlands and stretching out to other European countries. It is built to detect the formation of the first stars and galaxies via power-spectrum measurements of the 21 cm spin-flip line in hydrogen. This line in principle also allows for 3d mapping of ionization patches during the dark ages, via interaction with the CMB radiative field: Such mapping is out of the range of LOFAR, but the much bigger SKA (square kilometre array, Garrett et al., 2010), which should come online around 2020, will have such capabilities.

Observations and theory exist in a symbiotic relationship: The theory tells us what to look for, i.e. what, where and how to observe, and how then to interpret existing observations, while the observations strengthen or disqualify the theory and provide fodder for new theories. With next generation instruments coming up, the time is ripe for theoretical work to focus on these early epochs in the history of the Universe.

1.4 Cosmological simulations: A complementary tool for the exploration of the Universe

Due to the non-linearity involved with the formation of structures and the geometries of reionization, simulations are an invaluable tool for the theory. A number of cosmological codes exist that can follow the formation and evolution of structures in a ΛCDM framework. These codes typically model the gravitational interactions and hydrodynamics of dark matter and gas on large scales, with the initial conditions provided by statistical information from the CMB. They have been invaluable at explaining how structures large and small form in the Universe. Fig. 1.3 shows a comparison between the observed large-scale structure of our cosmological neighbourhood (blue), and output from a cosmological simulation, post-processed with semi-analytic models (red). The large scale structure in the simulation is purely the result of initial conditions and expansion history derived from WMAP data, and gravity. The simulation and cosmological simulations in general are very good at re-producing the webby structure of the real Universe, and in fact also match many observational properties, such as the mass function of galaxies derived from observations.

There are also some ongoing issues of phenomena routinely produced in simulations but whose existence in the real Universe remains unknown: A topical example is the existence of cold accretion streams that primarily feed simulated galaxies with gas at semi-high redshifts. These simulated streams present a paradigm shift in how galaxies are thought to have obtained their gas reserves. Hints of such streams exist in observations, but direct evidence still eludes us.

RHD

With the surging interest in reionization and the first sources of light in the Universe, and likely also thanks to a steadily increasing computational power, cosmological simulation

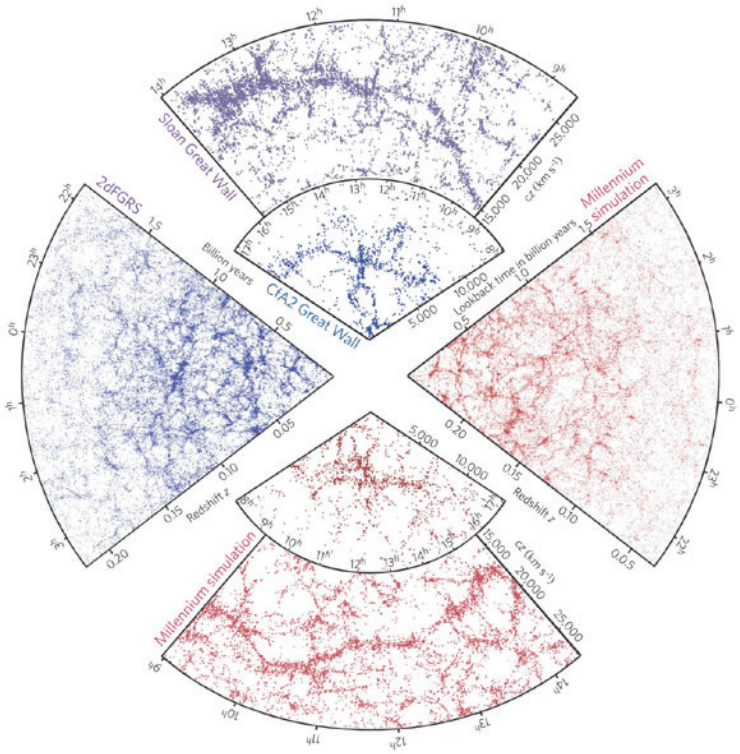


Figure 1.3: *Simulations versus observations of the large scale structure of the universe. The left and top wedges show results from sky-surveys of our neighbourhood, with each dot representing a galaxy. The bottom and right wedges show mock observations of the Milennium cosmological simulation at $z = 0$ with the same depths and fields of view as the real observations. Note how galaxies grow more diffuse as we reach further out in redshift in the left wedge due to the extinction of light at large distances. Yet this is only at a fraction of the redshift of the reionization epoch; there is very little to see out there with current instruments. Credit: Springel et al. (2006).*

codes have begun to include ionizing *radiative transfer* (RT) in the last decade or so. This is generally seen as a second-order component in most astrophysical processes, but important nontheless, and is obviously very important in the context of simulating reionization. Due to the challenges involved, most implementations have started out with the post-processing of ionizing radiation on otherwise frozen simulations, but a few have begun doing on-the-fly *radiation hydrodynamics* (RHD), which model the the interplay of radiation and gas.

It is highly desirable to follow self-consistently, with RHD simulations, the time-evolution and morphology of large-scale IGM reionization and at the same time the smaller scale formation of the presumed sources of reionization; how galaxy formation is regulated by the ionizing radiation being released, how much of the radiation escapes from the galaxies to ionize the IGM, how first generation stars are formed in a metal-free environment and how radiative and supernovae feedback from those stars affect the inter-galactic medium. The galaxies and the IGM are inter-connected via the ionizing radiation: The photons released from the galaxies affect the state of the surrounding gas via ionization and heating and may even prevent it from falling in or condensing into external gravitational potentials, especially small ones (e.g. Wise and Abel, 2008; Ocvirk and Aubert, 2011), which can significantly alter the ionization history.

The importance of RT and RHD is of course not limited to the epoch of reionization. Stars keep emitting ionizing radiation after this epoch and their radiative feedback likely has an effect on the post-reionization regulation of star-formation (e.g. Pawlik and Schaye, 2009; Hopkins et al., 2011b), the mass distribution of stellar populations (Krumholz et al., 2012) and even gas outflows (Hopkins et al., 2011a).

Localized and inhomogeneous ionizing radiation also affects the observational properties of extended gas via photoheating and photoionization. The emission and absorption properties of gas are extremely sensitive to the temperature and abundances of ion species. This applies generally to many emission mechanisms, but a highly relevant example is the emission/absorption of Lyα photons. Gas which is cold and free from ionizing radiation ($T \lesssim 10^4$ K) is Lyα absorbing, while warmer gas ($T \sim 10^4 - 10^5$ K) is Lyα emitting, due to a combination of cooling emission and fluorescence that may be induced by ionizing photons. The proximity of ionizing radiation can greatly affect these emission/absorption properties. Also, if the gas structure is large or dense enough, its interior may be shielded by the external radiation, which again affects the Lyα signature, and it may even 'shadow' some near-lying structures, affecting their radiative properties. Obviously, RT and RHD are needed to accurately capture these highly geometrical effects, to predict and explain the absorption/emission properties of gas clouds and filaments (e.g. Goerdt et al., 2010; Faucher-Giguère et al., 2010; Altay et al., 2011; Rosdahl and Blaizot, 2012; Fumagalli et al., 2011).

Radiative effects may also play a part in solving the overcooling problem in simulations, which is that simulated galaxies tend to be too compact and too rich in gas and stars compared to observations (e.g. Rasera and Teyssier, 2006). Prevention of cooling by high-energy metal-ionizing photons has been suggested to be the fix (Cantalupo, 2010).

Challenges

Radiation-hydrodynamics are complex and costly in simulations, especially those involving the interplay of the small-scale radiative sources and the large scale IGM. The inclusion of coupled radiative transfer in hydrodynamical codes in general is challenging mainly because of the high dimensionality of radiative transfer (space-, angular-, and frequency dimensions) and the inherent difference between the timescales of radiative transfer and non-relativistic hydrodynamics. To simulate the interaction of the small and large scales so relevant up to the epoch of reionization makes things even worse. One wants to simulate cosmic variance scales, i.e. on the order of 100 co-moving Mpc, the condensation of matter on these scales towards the formation of galaxy groups on Mpc scales and individual galaxies on kpc scales, the formations of stellar nurseries in those galaxies on pc scales, the formation of stars on sub-pc scales and then the effect of radiation from those stars back to the large scale IGM. This cycle involves size scale differences of something like 9-10 orders of magnitude – which is too much for the most advanced codes and computers today, actually even so without the inclusion of radiative transfer.

Due to these challenges, simulations typically focus on only a subset of these scales; either they consider reionization on large scales and apply sub-resolution recipes to determine stellar luminosities and UV escape fractions, or they ignore the cosmological context and focus on star formation and escape fractions in isolated galaxies or even isolated stellar

nurseries. [4]

A number of large scale 3d radiative transfer simulations of reionization have been carried out in recent years (e.g. Gnedin and Ostriker, 1997; Miralda-Escudé et al., 2000; Gnedin, 2000; Ciardi et al., 2003; Sokasian et al., 2004; Iliev et al., 2006b; Zahn et al., 2007; Croft and Altay, 2008; Baek et al., 2010; Petkova and Springel, 2010b; Aubert and Teyssier, 2010), though they must all to some degree use subgrid recipes for star formation rates, stellar luminosities and UV escape fractions, none of which are well constrained. The ionization history in these simulations thus largely depends on these input parameters and resolution – some in fact use the observational constraints of the ionization history to derive constraints on these free parameters (e.g. Sokasian et al., 2004; Croft and Altay, 2008; Baek et al., 2010; Petkova and Springel, 2010b; Aubert and Teyssier, 2010). Furthermore, most of these works have used a post-processing RT strategy instead of RHD, which neglects the effect the ionizing radiation has on the formation of luminous sources.

1.5 This PhD thesis: A tool for understanding the early evolution of galaxies

The primary driver behind this thesis is the desire to understand the birth of galaxies and stars during the dark ages, and how they link with their large scale environment. We have thus undertaken the project of implementing an RHD version of the widely used cosmological code `Ramses`, that we call `RamsesRT`, with the goal of running cosmological RHD simulations, optimized for *galactic* scale radiation-dynamics. `Ramses` is an adaptive mesh refinement (AMR) code, which greatly cuts costs by adaptively allowing the resolution to follow the formation of structures: The RHD implementation takes full advantage of the AMR strategy, allowing for high resolution simulations that can self consistently model the interplay of the reionizing Universe and the formation of the first galaxies.

Some of the goals we would like to tackle with this implementation are:

- Study radiative feedback effects in primordial galaxies. These galaxies are by definition young and small, and the first stars are thought to be gigantic and very bright due to the lack of metals. The ionizing radiation from these first stars is likely to have a dramatic effect on the galaxy evolution and may even stop short the further accumulation of gas. This is closely associated with the formation of molecules, needed to form the first stars, which is sensitive to the radiative field. Radiative feedback effects also appear to be relevant in lower-redshift galaxies, and likely have a considerable impact on the initial mass function of stellar populations.

- Investigate the escape of ionizing photons from early galaxies, how it affects the ionization history and external structure formation, e.g. the formation of satellite galaxies.

[4] An in-between alternative which is popular today is cosmological zoom simulations, where hand-picked sub-volumes of high resolution are simulated within a larger volume with coarser resolution, thus allowing one to have high resolution in a galaxy or galaxies of interest, while keeping the gravitational forces of its large scale environment.

- Study complementary scenarios of reionization: A proposed source of reionization are ionizing photons from accretion shocks onto galaxies (Dopita et al., 2011): The ability of our RHD implementation to use adaptive resolution and continuous sources of radiation make it very well suited to investigate this scenario. Also, the ionizing radiation from quasars in a cosmological context may be studied.

- Study the emissivity of galaxies and extended structures. Observable emission from gas is highly dependent on the ionization state of said gas. To predict it correctly, one thus needs to predict the ionization state correctly, for which RHD simulations are needed.

- Improve sub-resolution recipes: Of course we have not implemented a miracle code, and we are still nowhere near simulating simultaneously the 9-10 orders of magnitude in size-scales needed for fully consistent simulations of reionization. Sub-resolution strategies are still needed, and part of the objective is to improve those via small-scale simulations of stellar feedback (SNe, radiation, stellar winds).

There is a plethora of exciting issues that can be addressed with RHD implementations. Alas, most of the work done during this PhD is on the implementation itself rather than the science, and this thesis is largely a description of the RHD approach of `RamsesRT`. In Chapter 2 we introduce the basics of RHD theory and review the different methods and implementations that have been used in cosmological simulations, pure RT and RHD alike, with a focus on the so-called M1 method we use and the `Aton` code which is the inspiration behind ours. Chapter 3 is a methods chapter, containing a detailed description of the numerical implementation and how it is fit into the framework of `Ramses`. Tests and results for the `RamsesRT` implementation are then described in Chapter 4. Although my primary interest is the early evolution of galaxies during the dark ages up to the epoch of reionization, I have embarked on a project as a first application for `RamsesRT` which concerns the Universe at a more recent time. This is a study of the observational Lyα properties of extended structures at an intermediate redshift of $z = 3$, which appears to be a peak period of extended observed structures called Lyα blobs which may (or may not) be connected to the asymmetric cold (10^4 K) accretion of gas onto galactic halos, which is predicted by simulations to peak at similar redshifts – though they have never been unambiguously detected. Chapter 4 contains the first results of this work in the form of a paper which has been accepted for publication, with some added relevant background information. Of course, this is really just the first phase in work I hope to pursue for years to come: Chapter 6 describes the future steps I intend to take with the `RamsesRT` tool, and concludes this thesis.

2

On the theory of ionizing radiative transfer in cosmological simulations

We introduce in this chapter the theoretical basics of *radiation hydrodynamics* (RHD) and the main computational challenges involved, and review some existing solutions and implementations. This should lay the groundwork for the detailed description of our implementation in the next chapter. Our RHD implementation is largely inspired by the `Aton` RT code (Aubert and Teyssier, 2008, hereafter referred to as AT08): This discussion is thus streamlined towards the RT strategy used in `Aton`, though we will take detours to describe alternative strategies.

We shall first make clear the distinction between continuum and line radiative transfer: Our goal is to study the interplay of ionizing radiation, e.g. from stellar populations and AGN, and the interstellar/intergalactic gas. We consider *continuum* radiation, because the spectra of stars (and AGN) are smooth enough that emission and absorption processes are not sensitive to subtle rest-frame frequency shifts, be they due to local gas velocities or cosmological expansion.

On the other side is *line* transfer, i.e. the propagation of radiation over a narrow frequency range, usually corresponding to a central frequency that resonates with the gas particles. An important example is the propagation of Lyα photons. Here, one is interested in the complex frequency and direction shifts that take place via scattering on the gas particles, and gas velocities and subtle frequency shifts are vital components. Line transfer is mostly done to interpret observational spectra, e.g. from Lyα emitting/absorbing galaxies (e.g. Verhamme et al., 2006), and is usually run in post-processing under the assumption that the line radiation has a negligible effect on the gas dynamics (through this assumption is not neccessarily true; see Dijkstra and Loeb, 2009).

There is a bit of a grey line between those two regimes of continuum and line radiation, some codes are even able to do both (e.g. Baek et al., 2009; Pierleoni et al., 2009; Yajima et al., 2011b), but our implementation deals strictly with continuum radiation, as do most

RHD implementations, for the sake of speed and memory limitations. We do approximate multi-frequency, but only quite coarsely, such that simulated photons represent an average of photons over a relatively wider frequency range, and any subtle frequency shifts and velocity effects are ignored.

So while we don't have to deal with complicated resonances and scatterings in continuum radiative transfer, we do have the added complication that it should be *coupled* to the hydrodynamics via photon-gas interactions, rather than passive. We thus need to make it a part of a cosmological hydrodynamics code, in order to execute *radiation-hydrodynamics* (RHD) simulations. A handful of such codes exist on the market that can model three-dimensional structure formation in a cosmological context, based on the theory of gravitational hydrodynamics. The inclusion of ionizing radiative transfer in those codes has started to happen during the last decade or so, but the field is still young, mainly because RHD is computationally expensive.

2.1 Radiation-hydrodynamics in a cosmological context

First the bare basics: We will now present the equations of gravitational hydrodynamics and radiative transfer, and discuss the complexities involved in meshing these together into one system of gravitational RHD that can be used in cosmological simulations.

2.1.1 Cosmological hydrodynamics

Gravitational hydrodynamics in cosmology are governed by the Euler equations of fluid dynamics with the added source terms of gravity and thermochemistry, which encompasses radiative cooling and heating of the gas. They may be written to express the conservation of mass, momentum and energy:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{2.1}$$

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p = \rho \nabla \phi \tag{2.2}$$

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot (\mathcal{E} + p)\,\mathbf{u} = -\rho \mathbf{u} \cdot \nabla \phi + \Lambda(\rho, \varepsilon) \tag{2.3}$$

where t is time, ρ the gas density, $\mathbf{u}$ the bulk velocity, ϕ the gravitational potential, p the pressure, $\mathcal{E}$ the gas energy density, and Λ represents radiative cooling and heating via thermochemistry terms (resp. negative and positive), which are functions of the gas density, temperature and ionization state. Often, collisional ionization equilibrium (CIE) is assumed, which allows the ionization states to be calculated as surjective functions of the temperature and density and thus they don't need to be explicitly tracked in the code. $\mathcal{E}$ is divided into kinetic and thermal energy density (ε) components:

$$\mathcal{E} = \frac{1}{2}\rho u^2 + \varepsilon. \tag{2.4}$$

The system of Euler equations is typically closed with an equation of state which relates the pressure and energy,

$$p = (\gamma - 1)\varepsilon, \tag{2.5}$$

where γ is the ratio of specific heats. The Euler equations can easily be adapted to co-moving coordinates, to account for cosmological expansion, by a simple transformation of variables (see §3.3).

Vast scale differences are an ever-present problem in cosmological simulations. One is usually interested in resolving galactic scale processes, e.g. mergers, accretion, star-formation or stellar feedback, that are under the influence of cosmological processes that take place on much larger scales, spatially and temporally. This makes detailed cosmological simulations very impractical to execute on homogeneous grids, which is the traditional framework for solving the Euler equations.

Cosmological hydrodynamics codes generally come in two flavours, which employ different strategies for attacking the scale problem. These are *adaptive mesh refinement* (AMR) and *smoothed particle hydrodynamics* (SPH).

AMR

AMR codes are grid-based, with grid cells describing the values of ρ, $\mathbf{u}$ and $\mathcal{E}$. They are Eulerian in nature: Gas can flow between cells, but the cells themselves are static in space, i.e. they do not move around. However, the cell sizes are adaptively refined along with the formation of structures in the simulation, allowing the resolution to focus on these structures. Dark matter (DM) and stars are modelled as collisionless particles that freely flow around the volume under the influence of the gravitational potential, which in turn is built up from the mass contained in those particles and the gas cells.

AMR is a great advantage over homogeneous grids, both in the sense that it allows simulations to represent useful processes with a minimum number of resolution elements, and in adapting their resolution (space and time) locally such as to solve properly a set of equations at a minimum computational cost. Of course an AMR grid is also much more complicated to deal with than a homogeneous grid: Cells of different sizes should be evolved with different time-step lengths, and the grid structure cannot be represented by ordinary arrays as is the case with homogeneous grids.

The main cosmological AMR codes today are `Art` (Kravtsov et al., 1997), `Orion` (Klein, 1999), `Flash` (Fryxell et al., 2000), `Ramses` (Teyssier, 2002), and `Enzo` (O'Shea et al., 2004).

SPH

SPH codes drop the grid altogether and model all matter – DM, stars and gas alike – as particles. They are Lagrangian in nature, i.e. the particles move around the volume, which naturally allows the resolution – roughly speaking the length between particles – to focus on massive structures. The gas properties can be derived at any point in space via smoothing of the surrounding particles. The most widely used examples are the codes `Gasoline` (Wadsley et al., 2004) and `Gadget` (Springel, 2005).

2.1.2 The equation of radiative transfer

Let $I_\nu(\mathbf{x}, \mathbf{n}, t)$ denote the radiation specific intensity at location $\mathbf{x}$ and time t, such that

$$I_\nu \, d\nu \, d\Omega \, dA \, dt$$

is the energy of photons with frequency over the range $d\nu$ around ν propagating through the area dA in a solid angle $d\Omega$ around the direction $\mathbf{n}$.

The classical equation of radiative transfer (e.g. Mihalas and Mihalas, 1984) describes the local change in I_ν as a function of propagation, absorption and emission,

$$\boxed{\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{n} \cdot \nabla I_\nu = -\kappa_\nu I_\nu + \eta_\nu,} \tag{2.6}$$

where c is the speed of light, $\kappa_\nu(\mathbf{x}, \mathbf{n}, t)$ is an absorption coefficient and $\eta_\nu(\mathbf{x}, \mathbf{n}, t)$ a source function.

A co-moving coordinate version of Eq. 2.6 is derived in Appendix A, which includes cosmological terms accounting for the volume dilution of radiative energy and the stretching of wavelength, which together make the radiative energy density dilute as a^{-4}, where a is the cosmological scale factor. This radiation dilution is usually neglected in RT/RHD implementations, but it is worth storing behind the ear as we partially include it in ours (see §3.3).

2.1.3 The RHD equations

The system of RHD equations consists of the Euler equations (2.1)-(2.3), their equation of state closure (2.5), and the RT equation (2.6). The hydrodynamics and radiative transfer couple *only* through the thermochemistry $\Lambda(\rho, \varepsilon)$, i.e. via terms of photoionization, photoheating, and photon-emitting recombinations, all of which affect the thermal energy ε, the ionization fractions of the gas, and the cooling rate. Thus, in RHD (and also in RT), the solving of thermochemistry involves not only updating the temperature but also the local radiation field via emission and absorption and preferably also the *non-equilibrium* ionization state of the gas, e.g. the ionization fractions of hydrogen, helium and even some relevant species of metals or molecules. In the case of *momentum transfer*, i.e. the direct transfer of momentum from the photons to the gas, an additional coupling term appears on the RHS of (2.2). We do not implement momentum transfer so far, but will likely do so in the future.

2.1.4 Main challenges of numerical RT and RHD

There are two main challenges involved in implementing radiative transfer and coupling it to hydrodynamics:

Dimensions: The RT equation (2.6), while not much to look at, is highly expensive computationally, because it involves seven dimensions: Three spatial, two angular, frequency and time. In comparison, the Euler equations of hydrodynamics involve four dimensions, space and time. To take an example (from Steinacker et al., 2002), a resolution of 100 bins (e.g. grid elements or points) in each dimension requires structures that contain on the order of 10^{12} entries if the RT equation is to be discretized in full form, and a number of operations on

the order of 10^{14}, if the time resolution is also 100 bins. In comparison, solving the discretized Euler equations at the same resolution requires only on the order of 10^6 memory entries and 10^8 operations. Thus, in order to develop a pragmatic RT implementation, whether it is to be coupled to hydrodynamics or just for pure RT experiments, some simplifications must be made.

Timescales: There is typically a large difference between the characteristic timescales of hydrodynamics and RT. The hydrodynamical timescale Δt is usually constrained by some characteristic length scale Δx and flow speed u such that,

$$\Delta t < \frac{\Delta x}{u},$$

i.e. one must constrain a time-step such that certain characteristic distances are not crossed (e.g. cell widths in grid codes, smoothing lengths in SPH) or flows don't overlap. In the strictest limit, the RT timescale has a similar constraint, but here involving the speed of light. The maximum speed of matter (DM, stars, gas) attained in cosmological or non-relativistic astrophysical simulations is on the order of one-thousandth of the light speed, which implies that the time-step length in RHD (or pure RT) simulations is a thousand times smaller than in an HD counterpart, or conversely, the number of operations is a thousand times higher.

Many continuum RT – and some RHD – implementations have appeared in the last two decades, and they can be placed in two general categories, *ray-based schemes* and *moment methods*.

2.2 Ray-tracing schemes

Here the approximation is made that the radiation field is dominated by a limited number of sources. This allows us to greatly simplify the RT equation, with the local intensity becoming a function of the optical depth τ along *rays* from each source,

$$I_\nu(\mathbf{r}) = I_\nu(0)\exp[-\tau_\nu(\mathbf{r})], \tag{2.7}$$

where $\mathbf{r}$ is the position relative to the source and $\tau_\nu(\mathbf{r}) \equiv \sigma_\nu N_c(\mathbf{r})$, with σ_ν the cross section and $N_c(\mathbf{r})$ the column density of the photon-absorbing species.

The simplest solution is to cast rays, or **long characteristics** from each source to each cell (or volume element) and sum up the optical depth at each endpoint. With the optical depths in hand, I_ν is known everywhere and the rates of photoionization, heating and cooling can be calculated. While this strategy has the advantage of being simple and easy to parallelize (each calculation of $N_c(\mathbf{r})$ is independent), there is a lot of redundancy, since any cell which is close to a radiative source is traversed by many rays cast to further-lying cells, and is thus queried many times for its contribution to the optical depth. The parallelization is also not really so advantageous in the case of multiprocessor codes, since rays that travel over large lengths likely need to access cell states over many CPU nodes, calling for a lot of inter-node communication. Furthermore, the method is expensive: The computational cost scales linearly with the number of radiative sources, and each RT timestep has order $\mathcal{O}(n_{\text{sources}}\, n_{\text{cells}})$ operations, where n_{sources} is the number of radiative sources and n_{cells} is the

number of volume elements. Implementation examples include e.g. Susa (2006), Abel et al. (1999) and Cen (2002).

Short characteristics schemes overcome the redundancy problem by not casting separate rays for each cell. Instead, the calculation of optical depths in cells is propagated outwards from the source, and is in each cell based on the entering optical depths in the inner-lying cells. Calculation of the optical depth in a cell thus requires some sort of interpolation from the inner ones. There is no redundancy, as only a single ray segment is cast through each cell in one time-step. However, there is still a large number of operations and the problem has been made inherently serial, since the optical depths must be calculated in a sequence which follows the radiation ripple away from the source. Some examples are Nakamoto et al. (2001), Mellema et al. (2006), Whalen and Norman (2006) and Alvarez et al. (2006).

Adaptive ray tracing (e.g. Abel and Wandelt, 2002; Razoumov and Cardall, 2005; Wise and Abel, 2011) is a variant on short characteristics, where rays of photons are integrated outwards from the source, updating the ray at every step of the way via absorption. To minimize redundancy, only a handful of rays are cast from the source, but they are split into sub-rays to ensure that all cells are covered by them, and they can even be merged again if need be.

Cones are yet another variant on short characteristics, used in conjunction with SPH (Pawlik and Schaye, 2008, 2010) and the AREPO code (Petkova and Springel, 2010a): The angular dimension of the RT equation is discretized into tesselating cones that can collect radiation from multiple sources and thus ease the computational load and even allow for the inclusion of continuous sources, e.g. gas recombinations.

A **hybrid method** proposed by Rijkhorst et al. (2006) combines the long and short characteristics on patch-based grids (like AMR), to rid of most of the redundancy while keeping the parallel nature. Long characteristics are used inside patches, while short characteristics are used for the inter-patches calculations.

Monte-Carlo schemes do without splitting or merging of rays, but instead reduce the computational cost by sampling the radiation field, typically both in the angular and frequency dimensions, into photon packets that are emitted and traced away from the source. The cost can thus be adjusted with the number of packets emitted, but generally this number must be high in order to minimize the noise inherent to such a statistical method. Examples include Ciardi et al. (2001), Maselli et al. (2003), Altay et al. (2008), Baek et al. (2009), and Cantalupo and Porciani (2010). An advantage of the Monte-Carlo approach of tracking individual photon packets is that it naturally allows for keeping track of the scattering of photons. For line radiation transfer, where doppler/redshift effects in resonant photon scattering are important, Monte-Carlo schemes are the only feasible way to go – though in these cases, pure RT is usually sufficient (e.g. Cantalupo et al., 2005; Verhamme et al., 2006; Laursen and Sommer-Larsen, 2007; Pierleoni et al., 2009).

Ray-based schemes in general assume infinite light speed, i.e. rays are cast from source to destination in one moment in time. Many authors note that this only affects the initial speed of ionization fronts (I-fronts) around points sources (being faster than the light speed), but it may also result in an over-estimated I-front speed in underdense regions (see §4.3.5), and may thus give incorrect results in reionization experiments where voids are re-ionized too quickly. Some ray schemes (e.g. Wise and Abel, 2011; Pawlik and Schaye, 2008; Petkova

and Springel, 2010a) allow for finite light speed, but this adds to the complexity, memory requirement and computational load.

It is also unavoidable with the ray-based schemes that the computational load always increases with the number of radiative sources. This defect largely disappears with moment methods, though of course other defects appear instead.

2.3 Moment-based radiative transfer

In this alternative to ray-tracing schemes, the basic idea is to reduce the angular dimensions by taking angular moments of the RT equation (2.6). Intuitively this can be thought of as switching from a beam description to that of a field or a fluid, where the individual beams are replaced with a 'bulk' direction that represents an average of all the photons crossing a point/volume in space. This infers useful simplifications: Two dimensions are eliminated from the problem, and the equations take a form of conservation laws which is akin to the Euler equations of hydrodynamics[1] and are thus rather easily coupled to these equations and can even be solved with numerical methods designed for hydrodynamics, and in a framework adapted to hydrodynamics.

The main advantage is also the main drawback: The directionality is largely lost in the moment approximation and the radiation becomes somewhat diffusive, which is generally a good description of the optically thick limit, where the radiation scatters a lot, but not of the optically thin limit where the radiation is free-streaming. Radiation has a tendency to creep around corners with the moment method. Shadows are usually only coarsely approximated, if at all, though we will see e.g. in section §4.3.4 that sharp shadows *can* be maintained with idealized setups and a specific solver. The light speed is also an issue. Moment methods have explicit or implicit time-stepping constraints that basically limit the radiation from crossing more than one length-scale (cell size) in a time-step, so either one has to do many expensive time-steps to simulate a light crossing time in the free-streaming limit, or reduce the light speed.

2.3.1 Moments of the RT equation

The ith moment of a function $f(x)$ is the f-weighted average of x^i, i.e.

$$\int x^i f(x) dx.$$

Taking the zeroth and first angular moments of the RT equation (2.6) yields

$$\frac{1}{c}\frac{\partial}{\partial t}\oint I_\nu \, d\Omega + \nabla \oint \mathbf{n} \, I_\nu \, d\Omega = -\kappa_\nu \oint I_\nu \, d\Omega + \eta_\nu \oint d\Omega, \tag{2.8}$$

$$\frac{1}{c}\frac{\partial}{\partial t}\oint \mathbf{n} \, I_\nu \, d\Omega + \nabla \oint \mathbf{n} \otimes \mathbf{n} \, I_\nu \, d\Omega = -\kappa_\nu \oint \mathbf{n} \, I_\nu \, d\Omega, \tag{2.9}$$

where we assumed isotropic absorption and emission. These equations contain the first three moments of the specific intensity, which correspond to radiative energy density E_ν,

[1] The Euler equations can in fact be derived by taking velocity moments of the Boltzmann equation.

radiative flux $\mathbf{f}_\nu$, and radiative pressure $\mathbb{p}_\nu$, defined by

$$\begin{aligned} E_\nu &= \frac{1}{c} \oint I_\nu \, d\Omega && \text{(energy per volume and frequency)}, \\ \mathbf{f}_\nu &= \oint \mathbf{n} \, I_\nu \, d\Omega && \text{(energy flux per area and time and frequency)}, \\ \mathbb{p}_\nu &= \frac{1}{c} \oint \mathbf{n} \otimes \mathbf{n} \, I_\nu \, d\Omega && \text{(force per area and frequency)}, \end{aligned} \tag{2.10}$$

where $\otimes$ denotes an outer vector product, making $\mathbb{p}_\nu$ a 3×3 tensor. Inserting the definitions in (2.10) into the moment integrals (2.8) and (2.9) gives four coupled equations that describe the conservation of radiative energy and flux:

$$\frac{\partial E_\nu}{\partial t} + \nabla \cdot \mathbf{f}_\nu = -\kappa_\nu c E_\nu + S_\nu, \tag{2.11}$$

$$\frac{\partial \mathbf{f}_\nu}{\partial t} + c^2 \nabla \cdot \mathbb{p}_\nu = -\kappa_\nu c \mathbf{f}_\nu, \tag{2.12}$$

where the emission is now described by a scalar (isotropic) source term S_ν.

The angular dimensions have been suppressed from the RT equations, which now instead describe an average directionality of flow. In this framework, opposite flows of photons of the same frequency meeting in a point are not able to pass through each other, but instead sort of bounce off each other.

In moment based implementations, the frequency subscript is removed via integration over the frequency: In the *monogroup approximation* the integral is performed over the whole relevant frequency range (typically from the hydrogen ionization frequency and upwards), such that E and $\mathbf{f}$ represent photons of all ionizing frequencies. In the *multigroup approximation* the frequency range is split into a handful of bins, or photon groups, (rarely more than a few tens due to memory and computational limitations) and the equations (2.11) and (2.12) can be solved separately for each group. We can thus drop the frequency subscript and replace it with a photon group subscript $i = \{1, 2, 3, ..., M\}$, where M is the number of photon groups:

$$\frac{\partial E_i}{\partial t} + \nabla \cdot \mathbf{f}_i = -\bar{\kappa}_i c E_i + S_i, \tag{2.13}$$

$$\frac{\partial \mathbf{f}_i}{\partial t} + c^2 \nabla \cdot \mathbb{p}_i = -\bar{\kappa}_i c \mathbf{f}_i, \tag{2.14}$$

where $\bar{\kappa}_i$ is a spectrum weighted average absorption coefficient. Note that in order to legally make the separation of $\int \kappa_\nu E_\nu d\nu$ into $\bar{\kappa}_i E_i$, one must make the approximation that the spectral shape within each photon group is constant over space and time, as discussed in detail in §3.1. Some reference spectrum, e.g. that of a blackbody or a stellar energy distribution model (SED) is assumed for deriving $\bar{\kappa}_i$.

The equations already suggest the strategy and choice of variables for a numerical scheme, namely to superpose the radiative transfer on a grid of cells that describes gas hydrodynamics (gas density, velocity, and pressure), with each cell containing additional variables for E_i and $\mathbf{f}_i = (f_x, f_y, f_z)_i$.

2.3.2 Closing the moment equations

The moment equations (2.13) and (2.14) are, for a single frequency bin and in three dimensions, four equations with 13 variables – one in E, three in $\mathbf{f}$ and nine in $\mathbb{p}$. One thus needs some meaningful and physical closure to solve the equations.

Flux limited diffusion (FLD)

In the simplest form of moment-based RT implementations, so-called flux limited diffusion, we stop at the zeroth moment and only make use of the radiative energy conservation equation (2.13). The closure is then provided in the form of the local diffusion relation

$$\mathbf{f}_i = -\frac{c\lambda}{\kappa_i}\nabla E_i,$$

where λ is a parameter adjusted such that the diffusion speed is realistically restrained by the limit set by the gas opacity in the optically thick limit ($\lambda \to 1/3$) and by the speed of light in the optically thin limit ($\lambda \to \kappa_i E_i / \nabla E_i$). The photons are very diffusive with the FLD method and essentially just follow the energy gradient – which is realistic only if the medium is optically thick. Shadows are non-existent. The method has been used by e.g. Krumholz et al. (2007), Reynolds et al. (2009) and Commerçon et al. (2011), mainly for the purpose of studying the *force* feedback of infrared radiation onto dusty and optically thick gas, rather than photoionization of hydrogen and helium.

The Eddington tensor

When both moments, Eqs. (2.11) and (2.12), are used, a closure must instead be provided in the form of the radiative pressure tensor $\mathbb{p}_i$, which is usually expressed in terms of the Eddington tensor, $\mathbb{d}_i$, defined by

$$\mathbb{p}_i = \mathbb{d}_i E_i. \tag{2.15}$$

First note that $\mathbb{p}_i$ is symmetric by definition, i.e. $\mathbb{p}_i^{jk} = \mathbb{p}_i^{kj}$. From the specific intensity moments, Eqs. (2.10), and the fact that $\mathbf{n}_i$ is a unit vector, it can be seen that $\mathrm{Tr}[\mathbb{p}_i] = E_i$. Also, considering the second intensity moment, if the radiation intensity I_ν is isotropic, such as in the optically thick limit, then $\mathbb{p}_i = p_i \mathbf{I}$, where p_i is a pressure scalar and $\mathbf{I}$ is the identity matrix. Isotropic radiation thus corresponds to

$$\mathrm{Tr}[\mathbb{p}_i] = 3p_i,$$

giving

$$\mathbb{p}_i = \frac{E_i}{3}\mathbf{I}.$$

One can intuitively think of the pressure tensor as consisting of two components; an isotropic or diffusive component, corresponding to some multiple of $\mathbf{I}$, and a directional component corresponding to the rest of the tensor, i.e. when the isotropic component has been subtracted.

To close the system of equations (2.13 and 2.14), the Eddington tensor should be expressed in terms of the other variables, E and $\mathbf{f}$. Two approaches have been used in the literature for cosmological RT, the OTVET closure and the M1 closure.

The OTVET Eddington closure

Gnedin and Abel (2001) and later Petkova and Springel (2009) use the optically thin variable Eddington tensor formalism (OTVET), in which $\mathbb{d}$ is composed on-the-fly in every point in space from all the radiative sources in the simulation, assuming that the medium between source and destination is transparent (hence optically thin). This calculation is pretty fast, given the number of relevant radiative sources is not overburdening, and one can neglect these in-between gas cells. Finlator et al. (2009) take this further and include in the calculation the optical thickness between source and destination with a long characteristics method, which makes for an accurate but slow implementation.

The M1 Eddington closure

González et al. (2007) and later AT08 and Vaytet et al. (2010) – and now us – used a different Eddington closure formalism which has the advantage that it is completely local, i.e. one does not even have to consider the positions of the radiative sources, so only locally available quantities are needed for the calculation:

If one assumes the flux direction is an axis of symmetry of the local specific intensity (Levermore, 1984), the Eddington tensor may be written in the form

$$\mathbb{d} = \frac{1-\chi}{2}\mathbf{I} + \frac{3\chi-1}{2}\mathbf{n}\otimes\mathbf{n}, \tag{2.16}$$

where χ is the to-be-determined Eddington factor. Here one can recognize the diffusive and directional components to the Eddington tensor. One should always have $1/3 \leq \chi \leq 1$, where the lower limit corresponds to completely diffusive radiation and the upper limit to completely directional radiation. There is a variety of expressions for χ discussed in Levermore (1984). One very simple example is $\chi = 1/3$, which corresponds to the optically thick (isotropic) limit. The most meaningful (and yet simple) one is called the M1 model. It is obtained by applying a Lorentz transformation to a frame where the radiation is isotropic (and can also be derived by minimizing the radiative entropy; Dubroca and Feugeas, 1999). It gives

$$\chi = \frac{3+4f^2}{5+2\sqrt{4-3f^2}}, \tag{2.17}$$

where f is the reduced flux, defined by

$$f = \frac{|\mathbf{f}|}{cE}. \tag{2.18}$$

Note that by definition, one should always have $0 \leq f \leq 1$, so that the flux doesn't surpass the speed of light, and thus the constraint $1/3 \leq \chi \leq 1$ is always satisfied: The diffusive component disappears with $f = 1$ and the directional component with $f = 0$. Intermediate values of f represent a linear combination of directional and diffusive local radiation.

The M1 closure can establish and retain general directionality of photon flows, and can to some degree model shadows behind opaque obstacles. The M1 closure is very advantageous in the sense that it is purely local, i.e. it requires no information which lies outside the cell, which is not the case for the OTVET approximation.

As shown by Dubroca and Feugeas (1999), the M1 closure has the further advantage that it makes the system of RT equations (2.13 and 2.14) take the form of local hyperbolic conservations laws, where the characteristic wave speed c is the speed of light, and the constraint $f \leq 1$ is always ensured. Hyperbolic laws are well known and thoroughly investigated, and a plethora of numerical methods exist to deal with them (e.g. Toro, 1999). In fact, hydrodynamics are also a system of hyperbolic conservation laws, which implies we have the RT equations in a form which is well suited to lie alongside existing hydrodynamical solvers, e.g. in Ramses.

2.4 From Aton to RamsesRT

Aton (AT08) uses graphical processing units (GPUs) to post-process the transfer of single-frequency photons and their interaction with hydrogen gas. GPUs are very fast, but advantageous only if the volume is optimally structured, such that volume elements that are close in space are also close in memory, and if operations are easily parallelizable. It is ideal for post-processing RT on frozen simulation outputs, but hard to couple directly to play an active part in the simulations, though a sort of indirect coupling exists in the newest version where outputs are transferred back and forth between Ramses and Aton. Furthermore, due to the structuring constraint of GPUs, Aton RT must be performed on a homogeneous grid, so an AMR simulation output typically needs to be downgraded in resolution before being RT-postprocessed.

We have in our RamsesRT implementation used the same RT method as Aton does – the moment method with the M1 Eddington tensor closure. The basics of our radiative transport solver are in fact copied from an early version of Aton.

The biggest difference is that RamsesRT is built directly into the Ramses cosmological hydrodynamics code, allowing us to perform RHD simulations directly on the AMR grid, and without any transfer of data between different codes. Furthermore, we have expanded the implementation to include multigroup photons to approximate multifrequency, and we have added the interactions between photons and helium. We explicitly store and advect the ionization states of hydrogen and helium, and we have built into RamsesRT a new non-equilibrium thermochemistry functionality that evolves these states along with the temperature and the UV radiative field through chemical processes, photon absorption and emission. Finally, for realistic radiative feedback from stellar populations, we have implemented functions in RamsesRT that can read external SED models and derive from them luminosities and UV 'colors' of simulated stellar sources.

The sacrifice we make when inserting the RT implementation directly into Ramses is that we now use CPUs instead of GPUs, which makes a single radiative transfer step computationally slower, by something like a factor of one-hundred (Aubert and Teyssier, 2010). This is compensated for by employing the reduced speed of light approximation (RSLA), described in §3.2.4.

2.5 One more RHD implementation?

We have in this chapter listed a number of RT implementations, two of which even function in conjunction with `Ramses` (AT08, Commerçon et al., 2011), and one might ask whether another one is really needed?

To first answer for the `Aton` implementation, it is optimized for a different regime than `RamsesRT`. As discussed, `Aton` prefers to work with structured grids and requires transfer of outputs between the codes. It can work with large scale simulations, but it cannot deal well with adaptive refinement. This, plus the speed of `Aton`, makes it very good for studying large scale cosmological reionization, but not good for AMR simulations of individual halos/galaxies, e.g. cosmological zoom simulations, where the subject of interest is the effect of radiative feedback on the formation of structures and galaxy evolution, and escape fractions of ionizing radiation.

The Commerçon et al. (2011) implementation is on the opposite side of the spectrum. Being based on the flux limited diffusion method, it is optimized for RHD simulations of optically thick protostellar gas. It is a monogroup code that doesn't track the ionization state of the gas, but rather appears to focus on radiative force feedback. Furthermore, it uses an implicit solver, which makes it hard to adapt to other problems.

Out of the other codes we have listed, only three of those seem to have been used for published 3d cosmological RHD simulations with ionizing radiation. As far as we can see these are Gnedin and Abel (2001) (in `Art`), Wise and Abel (2008) (in `Enzo`), and Petkova and Springel (2009) (in `Gadget`). A few others that have been used for published astrophysical RHD simulations but without a co-evolving cosmology are Mellema et al. (2006), Susa (2006), Whalen and Norman (2006), and Baek et al. (2009). The rest apparently only do post-processing RT, aren't parallel or are otherwise not efficient enough. Many of these codes are also optimized for cosmological reionization rather than galaxy-scale feedback.

Thus there aren't so many cosmological RHD implementations out there, and there should be room for more. The main advantage of our implementation is that our method allows for an unlimited number of radiative sources and can even easily handle continuous sources, and is thus ideal for modelling e.g. the effects of radiative feedback in highly resolved simulations of galaxy formation, UV escape fractions, and the effects of self-shielding on the emission properties of gas and structure formation, e.g. in the context of galaxy formation in weak gravitational potentials.

3

Radiation-hydrodynamics with RamsesRT

In the previous chapter we derived the basic form of RT equations that we solve in our RamsesRT implementation, i.e. the four moment equations with the M1 closure for the Eddington tensor. Here we will deal with the numerical discretization of these equations and how it all fits into Ramses.

3.1 Multigroup H+He ionizing radiative transfer

We start with the moment equations of radiative transfer, Eqs. (2.11) and (2.12), that are closed with the M1 Eddington tensor, described by (2.16), (2.17) and (2.18). For numerical calculations, it is easier to work in terms of photon number density than energy density. We therefore divide the equations by the photon energy, $h\nu$, so they become

$$\frac{\partial N_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = -\sum_j^{\text{HI,HeI,HeII}} n_j \sigma_{\nu j} c N_\nu \quad + \dot{N}_\nu^\star + \dot{N}_\nu^{rec}, \tag{3.1}$$

$$\frac{\partial \mathbf{F}_\nu}{\partial t} + c^2 \nabla \cdot \mathbb{P}_\nu = -\sum_j^{\text{HI,HeI,HeII}} n_j \sigma_{\nu j} c \mathbf{F}_\nu, \tag{3.2}$$

$$\mathbb{P}_\nu = \mathbb{D}_\nu N_\nu, \tag{3.3}$$

where N_ν is photon number density (number per unit volume and frequency), $\mathbf{F}_\nu$ is photon number flux (number per unit area, time and frequency), $\mathbb{P}_\nu$ is a photon pressure tensor (with the same units as N_ν), and $\mathbb{D}_\nu$ is the Eddington tensor, given by the M1 closure. Here we have also broken the absorption coefficient, κ_ν, into constituent terms, $n_j \sigma_{\nu j}$, where n_j is number density of the photoabsorbing species j (=HI, HeI, HeII), and $\sigma_{\nu j}$ is the ionization cross section between ν-frequency photons and species j. Furthermore we have split the source term, S_ν, into (stellar) injection sources, $\dot{N}_\nu^\star$, and recombination radiation from gas, $\dot{N}_\nu^{rec}$.

Frequency discretization

We discretize the frequency distribution of UV photons into M photon groups, or *packages*, defined by

$$N_i = \int_{\nu_{i0}}^{\nu_{i1}} N_\nu \, d\nu, \qquad \mathbf{F}_i = \int_{\nu_{i0}}^{\nu_{i1}} \mathbf{F}_\nu \, d\nu, \tag{3.4}$$

where (ν_{i0}, ν_{i1}) is the frequency interval for package i. In the limit of one photon package, the frequency range is $(\nu_{i0}, \nu_{i1}) = (\nu_{\rm HI}, \infty)$; with $M > 1$ photon packages, the frequency intervals should typically be mutually exclusive and set up to cover the whole H-ionizing range:

$$[\nu_{00}, \nu_{01} : \nu_{10}, \nu_{11} : ... : \nu_{M0}, \nu_{M1}] = [\nu_{\rm HI}, \infty[.$$

Integrating the RT equations (3.1) and (3.2) over each frequency bin corresponding to the package frequencies yields M sets of four equations (plus the closures):

$$\boxed{\begin{aligned} \frac{\partial N_i}{\partial t} + \nabla \cdot \mathbf{F}_i &= -\sum_j^{\rm HI,HeI,HeII} n_j c \bar{\sigma}_{ij} N_i \; + \dot{N}_i^\star + \dot{N}_i^{rec}, && (3.5) \\ \frac{\partial \mathbf{F}_i}{\partial t} + c^2 \nabla \cdot \mathbb{P}_i &= -\sum_j^{\rm HI,HeI,HeII} n_j c \bar{\sigma}_{ij} \mathbf{F}_i, && (3.6) \\ \mathbb{P}_i &= \mathbb{D}_i N_i, && (3.7) \end{aligned}}$$

where *average cross sections* between each package i and species j are defined by[1]

$$\bar{\sigma}_{ij} = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} N_\nu \, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} N_\nu \, d\nu}. \tag{3.8}$$

The RT equations (3.5)-(3.7) now have a form which is ready for numerical discretization. For clarity and completeness we repeat the expressions for the M1 Eddington closure, here in frequency-discretized form and photon package notation: The Eddington tensor is given by

$$\boxed{\mathbb{D}_i = \frac{1-\chi_i}{2}\mathbf{I} + \frac{3\chi_i - 1}{2}\mathbf{n}_i \otimes \mathbf{n}_i,} \tag{3.9}$$

where

$$\boxed{\mathbf{n}_i = \frac{\mathbf{F}_i}{|\mathbf{F}_i|}, \qquad \chi_i = \frac{3 + 4f_i^2}{5 + 2\sqrt{4 - 3f_i^2}}, \qquad f_i = \frac{|\mathbf{F}_i|}{cN_i},} \tag{3.10}$$

are the unit direction, the Eddington factor and the reduced flux, respectively.

As we only keep track of the frequency-integrated photon densities and fluxes, the frequency distribution within each package is in principle lost as the photons propagate through the volume, interact with the gas, and mix. The average cross sections (3.8) are thus not (and cannot be) calculated on a per-cell basis, but are rather global variables that are derived from some *global frequency distribution* $J(\nu)$, which can for example be a blackbody,

[1] here we also assume the spectral shape of $\mathbf{F}_\nu$ to be identical, within each package, to that of N_ν.

AGN or stellar energy distribution (SED) model, and represents the frequency distribution, within each package frequency interval, of *all* emission sources in the simulation. The cross sections are thus in practice derived through

$$\boxed{\bar{\sigma}_{ij} = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J(\nu)\, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} J(\nu)\, d\nu}.} \tag{3.11}$$

Likewise, we must define *average photoionization energies* of each package-species couple in our simulation by

$$\boxed{\bar{\epsilon}_{ij} = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J(\nu) h\nu\, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J(\nu)\, d\nu},} \tag{3.12}$$

which are relevant to photoheating (see Eq. 3.59). In §3.6 we show how values for $\bar{\sigma}_{ij}$ and $\bar{\epsilon}_{ij}$ can be extracted on-the-fly in `RamsesRT` from SED models and in-simulation stellar populations.

These global properties imply an important approximation made in our implementation: For (3.5) and (3.6) to be valid derivations of (3.1) and (3.2), we must make the assumption that the unresolved *shape* of N_ν within each photon package frequency interval is identical in all cells (and likewise that of $\mathbf{F}_\nu$), i.e. that

$$N_\nu(\mathbf{x}) \,|_{\nu_{i0}}^{\nu_{i1}} = g_i(\mathbf{x}) J(\nu) \,|_{\nu_{i0}}^{\nu_{i1}}, \tag{3.13}$$

where $g_i(\mathbf{x})$ is a magnitude factor that represents the spatial variability in N_ν. The constant frequency distribution approximation (3.13) holds immediately after photons are emitted from the source, but as the photons transport through the volume and experience emission and absorption, the approximation starts to break down. In the limit that $M \to \infty$ and the frequency bins become narrower, Eq. 3.13 ceases to be an approximation and Eqs. (3.5)-(3.7) become better descriptions of multifrequency. Note, though, that Doppler effects cannot be appropriately modelled in this framework, no matter how many the photon packages, because of the average directionality we infer by taking moments of the RT equation.

The memory requirement and computational load both scale linearly with the number of photon packages in `RamsesRT`. Each photon package requires four additional cell variables in `RamsesRT` – $(N, F_x, F_y, F_z)_i$ – and Eqs. (3.5)-(3.7) are solved separately for each package. So while there can in principle be any number of photon packages, it is practically feasible to include only a handful. In `RamsesRT` simulations, we typically use three photon packages – HI, HeI and HeII ionizing – defined by the frequency intervals

$$\begin{aligned}
[\nu_{10}, \nu_{11}] &= [\nu_{\rm HI}, \nu_{\rm HeI}] &&= [3.29,\ 5.95]\ 10^{15}\ \ \text{Hz},\\
[\nu_{20}, \nu_{21}] &= [\nu_{\rm HeI}, \nu_{\rm HeII}] &&= [5.95,\ 13.2]\ 10^{15}\ \ \text{Hz},\\
[\nu_{30}, \nu_{31}] &= [\nu_{\rm HeII}, \infty[&&= [13.2,\ \infty[\quad 10^{15}\ \ \text{Hz}.
\end{aligned}$$

In addition to the $4 \times M$ photon package variables, `RamsesRT` contains three new variables that keep track of the ion abundances of the H and He species, namely

$$x_{\rm HII} = \frac{n_{\rm HII}}{n_{\rm H}}, \qquad x_{\rm HeII} = \frac{n_{\rm HeII}}{n_{\rm He}}, \qquad x_{\rm HeIII} = \frac{n_{\rm HeIII}}{n_{\rm He}}. \tag{3.14}$$

These ionization fractions are stored in `RamsesRT` as *passive scalars*; multiples of density that are advected with the gas.

3.2 Solving the RT equations on a grid

We will now describe how pure radiative transfer is solved on a grid structure – without yet taking into consideration the hydrodynamical coupling. The details here are not very specific to RamsesRT and are much like those of AT08. We will then in the next sections describe how it all fits inside Ramses and how we couple the radiative transfer to function on-the-fly inside Ramses simulations.

Like AT08, we solve (3.5)-(3.7) with an operator splitting strategy, which involves decomposing the equations into three steps that are executed in sequence over the same time-step Δt, which has some pre-determined length. The steps are:

(a) *Photon transport step*, where photons are propagated in space. This corresponds to solving (3.5)-(3.6) with the RHS $= 0$.

(b) *Photon injection step*, where radiation from stellar and other radiative sources (other than gas recombinations) is injected into the grid. This corresponds to the $\dot{N}_i^\star$ term in (3.5).

(c) *Thermochemistry step*, where the rest of the RHS of (3.5)-(3.6) is solved. This is where the photons and the gas couple, so here we evolve not only the photon densities and fluxes, but also the ionization state and temperature of the gas.

On a homogeneous grid, these steps could be executed in any order. However, in the AMR structure of Ramses the transport must be done ahead of the injection and thermochemistry (see §3.4.2).

The operator splitting approach means that each time the simulation time is advanced from t_i to t_f $(= t_i + \Delta t)$, the same time-jump over Δt is made three times: The photon densities and fluxes are first evolved to t_f via (a). Then we evolve again via (b) from t_i to t_f, which just means adding photons from the radiative sources on to the solution from the transport step. Finally, we go yet again back to time t_i and advance to t_f via the thermochemical step (c), using as an initial condition the t_f result from (a) and (b).

The operator splitting strategy is widely used in numerical simulation codes to break complicated equations into separate components. The reasons for using it are twofold: *First* it is to simplify the numerical solver. Solving the coupled set of equations like (3.5)-(3.7) in one go is not impossible but complicated and computationally expensive. Also, these equations have conceptually different components, each of which requires a numerical scheme adapted to its form. For example, many numerical schemes exist for efficient, stable and accurate transport, while different types of schemes exist that are adapted to solving the stiff equations of thermochemistry. This touches upon the *second* and perhaps more important main reason for using operator splitting, which is that the different components typically have different timescales attached to them. In our case, the transport is global, i.e. solving it requires synchronization between cells, which means the transport step must be restrained to a global time-step corresponding to the shortest transport timescale in any cell in our volume. The thermochemistry, on the other hand, requires no synchronization between cells, so the timescale is a local one. If the equations are attacked in one go instead of using operator splitting, the thermochemistry time-step also becomes global, because of the coupling to the transport, and thus the global time-step must be set to the minimum of

all transport and thermochemistry timescales, which would have a crippling effect on the simulation speed, since the termochemistry timescale can vary a lot.

Operator splitting is not a very physical thing to do – after all these processes happen simultaneously in reality, whereas the operator splitting strategy separates them and solves sequentially. Usually this doesn't present a problem in simulations, but it is good to be aware of what's happening in the code. In fact we have two issues with operator splitting approaches in this work which we discuss in detail in §3.2.5 and §5.2.4.

We'll now describe the three steps in detail, and having done that, explain how the time-step lengths are determined.

3.2.1 (a) Photon transport step

The equations describing free-flowing photons are

$$\frac{\partial N}{\partial t} + \nabla \cdot \mathbf{F} = 0, \tag{3.15}$$

$$\frac{\partial \mathbf{F}}{\partial t} + c^2 \nabla \cdot \mathbb{P} = 0, \tag{3.16}$$

i.e. (3.5)-(3.6) with the RHS $= 0$. Note that we have removed the photon package subscript, since this set of equations is solved independently for each package over the time-step.

We can write the above equations in vector form

$$\frac{\partial \mathcal{U}}{\partial t} + \nabla \mathcal{F}(\mathcal{U}) = 0, \tag{3.17}$$

where $\mathcal{U} = [N, \mathbf{F}]$ and $\mathcal{F}(\mathcal{U}) = [\mathbf{F}, c^2\mathbb{P}]$. To solve (3.17) over time-step Δt, we utilize an explicit conservative formulation, expressed here in 1D for simplicity,

$$\frac{\mathcal{U}_l^{n+1} - \mathcal{U}_l^n}{\Delta t} + \frac{\mathcal{F}_{l+1/2}^n - \mathcal{F}_{l-1/2}^n}{\Delta x} = 0, \tag{3.18}$$

where n corresponds to time index ($n = t$ and $n + 1 = t + \Delta t$) and l corresponds to cell index along the x-axis. $\mathcal{F}_{l+1/2}$ and $\mathcal{F}_{l-1/2} = \mathcal{F}_{(l-1)+1/2}$ are intercell fluxes evaluated at the cell interfaces. Simple algebra gives us the updated cell state,

$$\mathcal{U}_l^{n+1} = \mathcal{U}_l^n + \frac{\Delta t}{\Delta x}\left(\mathcal{F}_{l-1/2}^n - \mathcal{F}_{l+1/2}^n\right), \tag{3.19}$$

and all we have to do is determine expressions for the intercell fluxes.

Many intercell flux functions are available for differential equations of the form (3.17) which give a stable solution in the form of (3.19) (see e.g. Toro, 1999), as long as the Courant time-step condition is respected (see §3.2.4). It should be noted that the most intuitive approach, which is to linearly interpolate the intercell flux from both sides of the interface, results in an unstable and unusable method (Toro, 1999). We have followed AT08 and González et al. (2007) and implemented two flux functions which can be used in `RamsesRT`.

One is the Harten-Lax-van Leer (HLL) flux function (Harten et al., 1983),

$$(\mathcal{F}_{\mathrm{HLL}})_{l+1/2}^n = \frac{\lambda^+ \mathcal{F}_l^n - \lambda^- \mathcal{F}_{l+1}^n + \lambda^+\lambda^-\left(\mathcal{U}_{l+1}^n - \mathcal{U}_l^n\right)}{\lambda^+ - \lambda^-}, \tag{3.20}$$

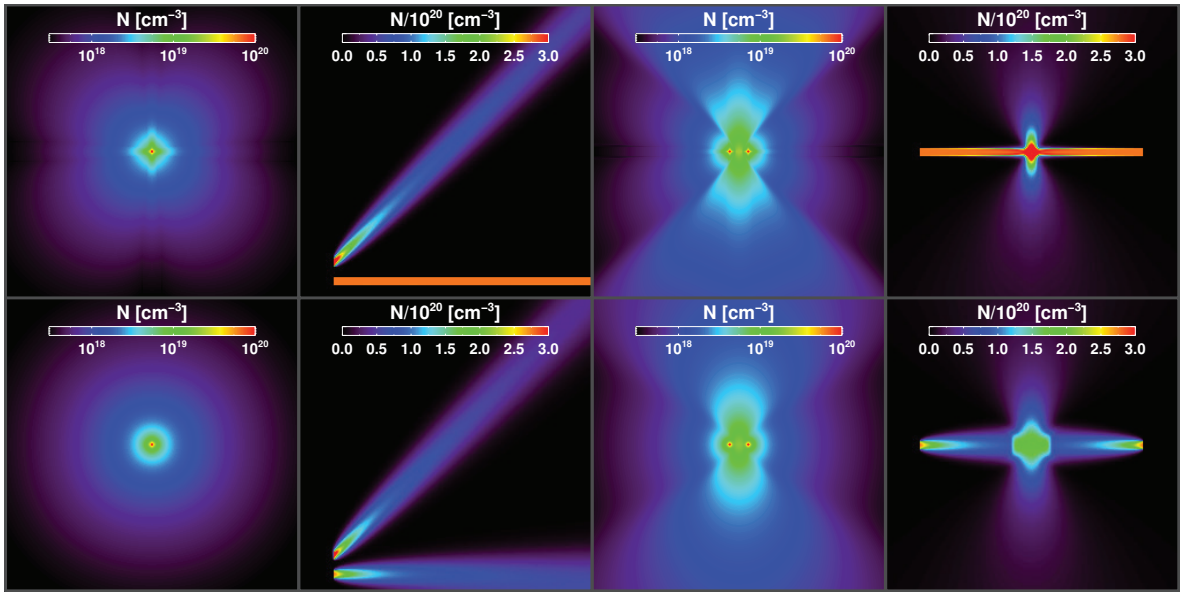


Figure 3.1: *Comparison of radiative transport in 2D* RamsesRT *runs (no gas interaction) with the HLL (**top**) and GLF (**bottom**) flux functions. The box width is* 1 *kpc and* 10^{50} *photons* s^{-1} *are injected continuously for each source. The snapshots are taken after a few light crossing times. **Far left** frames show single isotropic point sources. **Middle left** frames show attempts at creating horizontal and diagonal beams. Photons are injected into a vertical buffer of 7 cells with a unity reduced flux (i.e. no isotropic component). **Middle right** frames show two isotropic point sources and how the photons act between them. **Far right** frames show two beams of opposing directions and how a perpendicular radiation source forms where they meet.*

where, $\lambda^+ = \max(0, \lambda_l^{\max}, \lambda_{l+1}^{\max})$ and $\lambda^- = \min(0, \lambda_l^{\min}, \lambda_{l+1}^{\min})$ are maximum and minimum eigenvalues of the Jacobian $\partial\mathcal{F}/\partial\mathcal{U}$. These eigenvalues mathematically correspond to wave speeds, which in the case of 3D radiative transfer depend only on the magnitude of the reduced flux f (3.10) and the angle of incidence of the flux vector to the cell interface. This dependence has been calculated and tabulated by González et al. (2007), and we utilize their table to extract the eigenvalues.

The other flux function we have implemented is the simpler Global Lax Friedrich (GLF) function,

$$(\mathcal{F}_{\mathrm{GLF}})_{l+1/2}^n = \frac{\mathcal{F}_l^n + \mathcal{F}_{l+1}^n}{2} - \frac{c}{2}\left(\mathcal{U}_{l+1}^n - \mathcal{U}_l^n\right), \tag{3.21}$$

which corresponds to setting the HLL eigenvalues to the speed of light, i.e. $\lambda^- = -c$ and $\lambda^+ = c$, and has the effect of making the radiative transport more diffusive. Beams and shadows are therefore better modelled with the HLL flux function than with the GLF one, whereas the inherent directionality in the HLL function results in radiation around isotropic sources (e.g. stars) which is noticeably asymmetric, due to the preference of the axis directions.

Fig. 3.1 illustrates the difference between the two flux functions in some idealized 2D RamsesRT tests, where we shoot off beams and turn on isotropic sources. It can be seen that the HLL flux function fails to give isotropic radiation (far left) and that the GLF function

gives more diffusive beams (second from left). Note also how the diffusivity of beams with the HLL flux function is direction-dependent. A horizontal or vertical beam is perfectly retained while a diagonal one 'leaks' to the sides almost as much as with the GLF function, which has the advantage of being consistent on whether the beam is along-axis or diagonal. The right frames of the figure give an idea of how the radiative transport behaves in the case of multiple sources, i.e. with opposing beams and near-lying isotropic sources. The two flux functions are of course mutually exclusive: One can only pick one or the other for a simulation. We generally prefer to use GLF, since we mostly deal with isotropic sources in our cosmological/galactic simulations, but the choice of function really depends on the problem. There is no noticeable difference in the computational load, so if shadows are important, one should definitely go for HLL, as we have done in the shadowing RT tests described in §4.3.4 and §4.3.7. AT08 have compared the two flux functions in some of the benchmark RT tests of Iliev et al. (2006a) and found that they give very similar results. We do likewise for the test we describe in 4.3.7, and come to the same conclusion.

3.2.2 (b) Photon injection step

The equations to solve in this step are very simple,

$$\frac{\partial N_i}{\partial t} = \dot{N}_i^\star, \tag{3.22}$$

where $\dot{N}_i^\star$ is a rate of photon injection into photon package i, normally coming from stellar sources, but this could also include other point sources such as AGN, and also pre-defined point sources or continuous 'volume' sources.

Given the time-step length Δt and the cell injection rate $\dot{N}_i^\star$, the discrete update done for each photon package (in each cell) is

$$N_i^{n+1} = N_i^n + \frac{\Delta t}{V}\dot{N}_i^\star \tag{3.23}$$

where V is the cell volume.

Section 3.6 shows how $\dot{N}_i^\star$ and the photon cross sections and energies can be determined on-the-fly from stellar models in `RamsesRT`.

3.2.3 (c) Thermochemistry step

In the final step, we solve for the interaction between photons and gas. This is done by solving (3.5) and (3.6) with zero divergence and stellar injection terms, which leaves us with

$$\frac{\partial N_i}{\partial t} = -\sum_j^{\text{HI,HeI,HeII}} n_j c \bar{\sigma}_{ij} N_i \quad + \dot{N}_i^{rec}, \tag{3.24}$$

$$\frac{\partial \mathbf{F}_i}{\partial t} = -\sum_j^{\text{HI,HeI,HeII}} n_j c \bar{\sigma}_{ij} \mathbf{F}_i. \tag{3.25}$$

These equations however involve more than just photon densities and fluxes. Photon absorption and emission have the effect of heating and cooling the gas, so in order to self-consistently implement these interactions, we evolve along with them the thermal energy

ε of the gas and the abundances of the species that interact with our photons, here HI, HeI and HeII via photoionizations and HII, HeII (again) and HeIII via recombinations. We follow these abundances in the form of the three ionization fractions $x_{\rm HII}$, $x_{\rm HeII}$ and $x_{\rm HeIII}$, that we presented in Eqs. (3.14). Unlike RamsesRT, Ramses does not explicitly keep track of these abundances, but rather assumes that they exist in photoionization equilibrium, such that they can be calculated directly as a function of temperature and redshift. This should normally be an adequate approximation, but it won't do in the context of self-consistently following the radiative field.

The set of non-equilibrium thermochemistry equations solved in RamsesRT consists of:

$$\frac{\partial N_i}{\partial t} = -\sum_j^{\rm HI,HeI,HeII} n_j c\bar{\sigma}_{ij} N_i \quad + \sum_j^{\rm HII,HeII,HeIII} b_{ji}^{rec}\left[\alpha_j^{\rm A}(T) - \alpha_j^{\rm B}(T)\right] n_j\, n_e, \tag{3.26}$$

$$\frac{\partial \mathbf{F}_i}{\partial t} = -\sum_j^{\rm HI,HeI,HeII} n_j c\bar{\sigma}_{ij} \mathbf{F}_i, \tag{3.27}$$

$$\frac{\partial \varepsilon}{\partial t} = \mathcal{H} + \mathcal{L} + \frac{Z}{Z_\odot}\mathcal{L}_Z, \tag{3.28}$$

$$n_{\rm H}\frac{\partial x_{\rm HII}}{\partial t} = n_{\rm HI}\,\beta_{\rm HI}(T) n_{\rm e} + n_{\rm HI}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HI} c N_i \left(+n_{\rm HI}\Gamma_{\rm HI}^{\rm hom}\right) - n_{\rm HII}\,\alpha_{\rm HII}^{\rm A}(T) n_{\rm e}, \tag{3.29}$$

$$\begin{aligned} n_{\rm He}\frac{\partial x_{\rm HeII}}{\partial t} &= n_{\rm HeI}\,\beta_{\rm HeI}(T) n_{\rm e} + n_{\rm HeIII}\,\alpha_{\rm HeIII}^{\rm A}(T) n_{\rm e} + n_{\rm HeI}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeI} c N_i \left(+n_{\rm HeI}\Gamma_{\rm HeI}^{\rm hom}\right) \\ &\quad - n_{\rm HeII}\,\beta_{\rm HeII}(T) n_{\rm e} - n_{\rm HeII}\,\alpha_{\rm HeII}^{\rm A}(T) n_{\rm e} - n_{\rm HeII}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeII} c N_i \left(-n_{\rm HeII}\Gamma_{\rm HeII}^{\rm hom}\right) \end{aligned} \tag{3.30}$$

$$n_{\rm He}\frac{\partial x_{\rm HeIII}}{\partial t} = n_{\rm HeII}\,\beta_{\rm HeII}(T) n_{\rm e} + n_{\rm HeII}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeII} c N_i \left(+n_{\rm HeII}\Gamma_{\rm HeII}^{\rm hom}\right) - n_{\rm HeIII}\,\alpha_{\rm HeIII}^{\rm A}(T) n_{\rm e} \tag{3.31}$$

The discretization of these equations is scrutinized in §3.5, so we suffice here in explaining the different terms:

- The photon density and flux equations, (3.26) and (3.27), are just repeats of (3.24) and (3.25), where we've replaced the photon emission rate $\dot{N}_i^{rec}$ with the full expression for recombinative emissions from gas. Here, $\alpha_j^{\rm A}$ and $\alpha_j^{\rm B}$ represent case A and B recombination rates for electrons combining with species j (= HII, HeII, HeIII). Case A recombinations are to all levels and case B are to all except ground level. Only ground level recombinations release ionizing photons, hence the emission rate corresponds to the difference between the two rates. The b_{ji}^{rec} factor is a boolean (1/0) that states which photon package j-species recombinations emit into, and $n_{\rm e}$ is electron number density (a direct function of the H and He ionization states, neglecting the metals).

- The temperature-evolution, (3.28) is quite long and simplified here – the full version can be found in §3.5. Basically it consists of three terms: The photoheating rate $\mathcal{H}(N_i, x_{\rm HII}, x_{\rm HeII}, x_{\rm HeIII}, n_{\rm H})$, the radiative cooling rate $\mathcal{L}(T, N_i, x_{\rm HII}, x_{\rm HeII}, x_{\rm HeIII}, n_{\rm H})$, and

metal-line contributions to radiative cooling, $\mathcal{L}_Z(T, n_{\rm H})Z/Z_\odot$, where $\mathcal{L}_Z$ is a precomputed per solar-metallicity contribution, and $Z/Z_\odot$ is the metal mass fraction in the gas/sun.

- The $x_{\rm HII}$ evolution (3.29) consists of, from left to right on the RHS, HI collisional ionizations, HI photoionizations (with an optional homogeneous UV background contribution in parentheses), and HII recombinations. Here, $\beta(T)$ is a rate of collisional ionizations and $\Gamma^{\rm hom}$ is a UV background photoionization rate, pre-calculated from some assumed UV background model.
- The $x_{\rm HeII}$ evolution (3.30) consists of, from left to right, HeI collisional ionizations, HeIII recombinations, HeI photoionizations (UV background contribution in parentheses), and HeII collisional ionizations, recombinations, and photoionizations.
- Likewise, the $x_{\rm HeIII}$ evolution (3.31) consists of HeII collisional ionizations and photoionizations, and HeIII recombinations.

The on-the-spot approximation

It is optional whether or not the photon-emitting recombinative term, the second RHS sum in (3.26), is included. Excluding it is usually referred to as the on-the-spot approximation (OTSA), meaning that any recombination-emitted photons are absorbed 'on the spot' by a near-lying atom (in the same cell), and hence these photon emissions cancel out by local photon absorptions. If the OTSA is assumed, the gas is thus not photoemitting, and the case A recombination rates are replaced with case B recombination rates in (3.26)-(3.31), i.e. photon-emitting recombinations straight down to the ground level are not counted. The OTSA is in general a valid approximation in the optically thick regime but not so when the photon mean free path becomes longer than the cell width.

3.2.4 The RT time-step and the reduced speed of light

In each iteration before the 3 RT steps of photon transport, injection and thermochemistry are executed, the length of the time-step, Δt, must be determined.

As we stated in §3.2.1, we use an explicit solver for the radiative transport (3.18). An explicit advection solver is constrained by the Courant condition, which states that no information should move further than the length of one cell in a single time-step. Since the information speed is the speed of light in the case of RT, this translates to the direct constraint

$$\Delta t \leq \frac{\Delta x}{c}, \tag{3.32}$$

where Δx is the cell width. In practice, the time-step length is set to

$$\Delta t = f_{Courant}\frac{\Delta x}{c}, \tag{3.33}$$

where $f_{Courant} \leq 1$ is a global Courant factor. Compared to non-relativistic hydrodynamical simulations, this time-step constraint is severe, because of the very different timescales in hydrodynamics and RT. Hydrodynamical simulations have a similar time-step constraint,

but here the speed in question is the maximum of the sound and bulk speeds, and these are typically slower than the light speed by a factor of 100 to 1000. This means that for every hydrodynamical time-step, one would need to perform $100 - 1000$ RT time-steps, which basically makes RT impossible if one wants to couple it to the hydrodynamics[2]. Even when running postprocessing RT simulations where the hydrodynamics are turned off, this usually proves to be a very limiting constraint, because the propagation speed of ionization fronts (I-fronts), which roughly determines the simulation timescale, is usually much slower than the light speed.

In the case of radiative transfer with the moment equations, there are two well-known solutions to this problem.

One is to use an *implicit* method instead of an explicit one to solve the transport equation, which means using forward-in-time intercell fluxes in (3.18), i.e replacing $\mathcal{F}^n \equiv \mathcal{F}^t$ with $\mathcal{F}^{n+1} \equiv \mathcal{F}^{t+\Delta t}$. This seemingly simple change ensures that the solution is *always stable*, no matter how big the time-step, and we are rid of the Courant condition. However: *(i)* It doesn't mean that the solution is accurate, and in fact we still need some time-stepping condition to retain the accuracy, e.g. to restrain any quantity to be changed by more than say 10% in a single time-step. Furthermore, such a condition usually must be checked by trial-and-error, i.e. one guesses a time-step and performs a global transport step (over the grid) and then checks whether the accuracy constraint was broken anywhere. Such trial-and-error time-stepping can be very expensive since it is a global process. *(ii)* Replacing $\mathcal{F}^t$ with $\mathcal{F}^{t+\Delta t}$ is actually not simple at all. Eqs. 3.18 become a system of coupled algebraic equations that must be solved via matrix manipulation in an iterative process, which is complicated, computationally expensive, and of limited scope (i.e. can't be easily applied to any problem). Due to these two reasons we have opted out of the implicit approach. It is absolutely a valid approach however, and used by many (e.g. Petkova and Springel, 2009).

The approach we have chosen instead is to keep our solver explicit, and relax the Courant condition by changing the speed of light to a *reduced light speed* $c_r << c$, the payoff being that the time-step (3.33) becomes longer. This is generally referred to as the reduced speed of light approximation (RSLA), and was introduced by Gnedin and Abel (2001). The idea is that the timescale involved in RT simulations is not really subject to the speed of photons but rather to the speed of I-fronts. As long as the photon speed is faster than the fastest I-front, then changing the speed of light has a negligible effect on the results. We have confirmed this to be the case in most of the RT tests described in §4.3. The trick is to know the propagation speed of I-fronts. One can derive their approximate speed in an idealized setup (Strömgren, 1939), and in general the speed becomes faster as the medium becomes more diffuse, but it is difficult to translate that to non-idealized cosmological or galactic simulations. So far, we have chosen reduced light speed fractions of $f_c \sim 10^{-2} - 10^{-3}$ (with $f_c \equiv c_r/c$), that make RT experiments feasible in terms of computational time, and confirmed, by varying the light speed by a factor of a few, that the chosen light speed has a negligible effect on the results. This of course depends on the context of the simulations: In simulations of galactic feedback it seems that $f_c \sim 10^{-2} - 10^{-3}$ is a valid approximation, but in the case of simulating cosmological reionization, a correct physical speed of light appears

[2]This is actually exactly what `Aton` does – but `Aton` runs on GPUs, which are about a hundred times faster than CPUs (see §2.4), whereas `RamsesRT` runs on CPUs and thus can't afford such huge amount of RT subcycling.

to be important for correct modelling (see §4.3.5). In Appendix B we consider I-front speeds for the idealized Strömgren sphere and how the RSLA fares against them. The analysis presented there can serve as a crude yardstick to determine the minimum 'allowed' light speed in `RamsesRT` simulations.

In practice, the implementation of the RSLA consists of simply exchanging the true speed of light c with the reduced one c_r in *all* the discretized RT equations and expressions introduced in this chapter.

The 10% thermochemistry rule

The Courant condition is not the only condition that limits the time-step in the 3 RT steps. For accuracy, each thermochemistry step is also restricted by a cooling time which prohibits any of the thermochemical quantities to change by a substantial fraction in one time-step. Fortunately, this timescale is a local one: Some cells may have cooling times much longer than the global Courant condition and others much shorter, and we want to avoid reducing the global timescale to the smallest local cooling time, which would have severe consequences on the computational load. We therefore subcycle the thermochemistry step to fill in the RT time-step, using what we call *the* 10% *rule*: In each cell, the thermochemistry step is initially executed with the full RT time-step length, and then the fractional update is considered. If any of the evolved quantities (N_i, F_i, ε, ionization fractions) have changed by more than 10%, we backtrack and do the same calculation with half the time-step length. At the end of each calculation we perform the same 10% check to determine what we do next.

3.2.5 Smooth RT

It turns out in `RamsesRT` runs and tests that there are often some cells, usually it seems along I-fronts, that execute a huge number of thermochemistry subcycles in a single RT time-step. This is in part fault of the operator-splitting approach used, where the RT equations have been partly decoupled. Specifically, the photon density updates happen in three steps in this approach (see Fig. 3.2, left). The photon injection step always increases the number of photons, usually by a relatively large amount, and the transport step does the same when it feeds photons into cells along these I-fronts. The thermochemistry step in the I-front cells has the exact opposite effect: There the photon density decreases again via absorptions. If the photon-depletion time is shorter than the Courant time, we have a curious situation where the cell goes through a cycle during the thermochemistry subcycles: It starts neutral with a large abundance of photons (that have come in via the transport and/or photon injection steps). It first requires a number of subcycles to evolve to a (partly) ionized state, during which the photon density is gradually decreased. It can then reach a turnaround when the photons are depleted. If the RT time-step is not yet finished, the cell then goes into a reverse process, where it becomes neutral again. This whole cycle may take a large number of thermochemical steps, yet the cell gas ends up being in much the same state as it started. In a way, the cell is making a big fuss about nothing, and spending computational resources in doing so.

In reality, the ionization state and photon density would not cycle like this but would rather settle into a semi-equilibrium where the rate of ionizations equals that of recombinations.

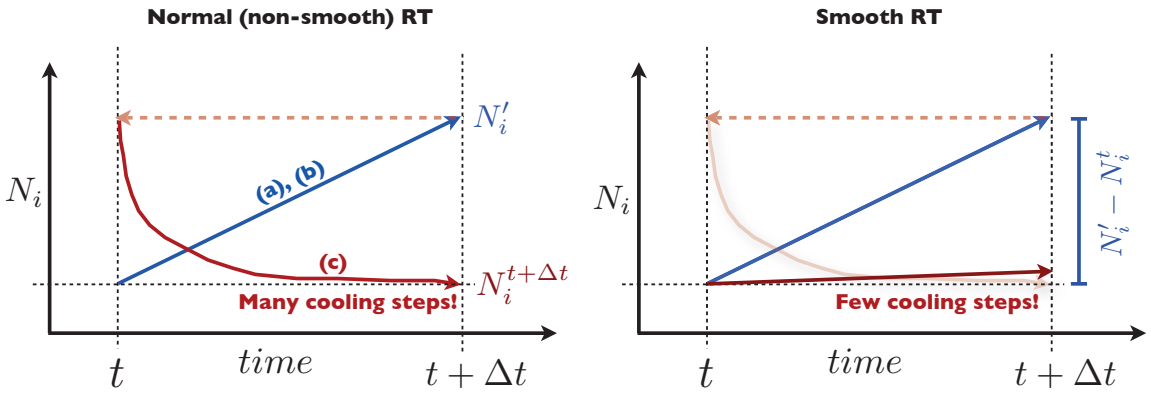


Figure 3.2: *Sketch plots showing a photon density evolution over a global RT time-step with normal RT (**left**) and smooth RT (**right**). In normal RT the photon density is updated to N' during photon transport (a) and injection (b). This is then used as an initial state for thermochemistry (c). It is often the case that the photons are depleted over the global time-step Δt, in a process which takes many thermochemistry (cooling) subcycles. In smooth RT, the photon density state is not updated by the transport and injection steps, but rather the difference is used to infer a photon injection rate for the cell, which is gradually added during each thermochemistry cycle. This can dramatically reduce the needed number of chemistry subcycles.*

For the purpose of saving up on computing time and reducing the number of thermochemistry subcycles, we have implemented an optional strategy we call *smooth RT* that roughly corrects this non-equilibrium effect of operator splitting (see Fig. 3.2, right). In it, the result of $(N_i', \mathbf{F}_i')$ from the transport and injection steps in each cell is used to infer a *rate* for the thermochemistry step, rather than being set as an initial condition. We use the pre-transport, pre-injection values of N_i^t and $\mathbf{F}_i^t$ as initial conditions for the thermochemistry, but instead update the thermochemistry equations (3.26) and (3.27) to

$$\frac{\partial N_i}{\partial t} = - \sum_j^{\text{HI,HeI,HeII}} n_j c \bar{\sigma}_{ij} N_i \; + \; \dot{N}_i^{rec} + \dot{N}_i, \tag{3.34}$$

$$\frac{\partial \mathbf{F}_i}{\partial t} = - \sum_j^{\text{HI,HeI,HeII}} n_j c \bar{\sigma}_{ij} \mathbf{F}_i + \dot{\mathbf{F}}_i, \tag{3.35}$$

where the new terms at the far right represent the *rates* at which the photon densities and fluxes changed in the transport and injection steps, i.e.

$$\dot{N}_i = \frac{N_i' - N_i^t}{\Delta t}, \tag{3.36}$$

$$\dot{\mathbf{F}}_i = \frac{\mathbf{F}_i' - \mathbf{F}_i^t}{\Delta t}. \tag{3.37}$$

The idea is that when the photons are introduced like this into the thermochemistry step, they will be introduced *gradually* in line with the subcycling, and the photon density vs. ionization fraction cycle will disappear as a result and be replaced with a semi-equilibrium, which should reduce the number of subcycles and the computational load. The total photon injection (or depletion) will still equal $N_i' - N_i^t$, so in the limit that there are no

photoionizations or photon-emitting recombinations, the end result is exactly the same photon density (and flux) as would be left at the end of the transport and injection steps without smooth RT.

The gain in computational speed is dependent on the problem at hand, and also on the reduced light speed, which determines the size of the RT time-step. We've made a comparison on the computational speed between using the smooth and non-smooth RT in a cosmological zoom simulation from the NUT simulations suite (e.g. Powell et al., 2011) that includes the transfer of UV photons from stellar sources. Here, smooth RT reduces the average number of thermochemistry subcycles by a factor of 1/6 and the computing time by a factor 2/7. So a lot may indeed be gained by using smooth RT.

The smooth method has the drawback that photons are not strictly conserved – in a way photons can exist and contribute to the thermochemistry step in two adjacent cells at the same time(-step), if the number of thermochemistry subcycles is different between the cells. Result-wise however, we have extensively verified that the difference between smooth and non-smooth RT is minimal.

One could even argue that the ionization states in I-fronts are better modelled with smooth RT, since this cycle of photon density and ionization fraction is a purely numerical effect of operator splitting. We have intentionally drawn a slightly higher end value of N_i in the smooth RT than non-smooth in Fig. 3.2: Whereas non-smooth RT can completely deplete the photons in a cell, smooth RT usually leaves a small reservoir after the thermochemistry, that more accurately represents the 'equilibrium value'. We show this difference concretely in a comparison between smooth and non-smooth RT in §5.2.7. In Chapter 5 we also extensively describe a similar effect, where operator splitting in `Ramses` hydrodynamics results in slightly underestimated temperatures in our simulations of extended Lyα emission (though there we resolve the problem by other means than 'smooth hydrodynamics').

Of course an alternative to smooth RT, and a more correct solution, is to attack the root of the problem and reduce the global time-step length, i.e. also limit the transport and injection steps to the 10% rule. Reducing the global time-step length is highly impractical though; the main reason for using operator splitting in the first place is that it enables us to separate the timescales for the different steps.

The same method of smoothing out discreteness that comes with operator splitting (in the case of pure hydrodynamics) has previously been described by Sun (1996), where it is referred to as 'pseudo-non-time-splitting'.

3.3 Cosmological settings

`Ramses` uses super-comoving variables to allow for the impact of the cosmological expansion on the Poisson equation, the equations of hydrodynamics (2.1-2.3) and particle propagation (Martel and Shapiro, 1998; Teyssier, 2002): A change is made from the physical variables to

super-comoving ones with

$$d\tilde{t} = \frac{H_0}{a^2}\, dt, \qquad \tilde{x} = \frac{H_0}{aL}\, x, \qquad \tilde{\rho} = \frac{a^3}{\Omega_m \rho_c H_0^2}\, \rho,$$
$$\tilde{\mathbf{u}} = \frac{a}{L}\, \mathbf{u}, \qquad \tilde{\varepsilon} = \frac{a^5}{\Omega_m \rho_c H_0^2 L^2}\, \varepsilon,$$

where H_0 is the Hubble constant, Ω_m the matter density parameter, L the comoving width of the simulation box (physical width at $a = 1$), and ρ_c the critical density of the Universe. When these variables are used instead of the physical ones, the cosmological expansion is accounted for, while all relevant equations remain unchanged, Euler equations included.

For consistency, and to partly account for the effect of cosmological expansion on the radiative transfer, the additional change is made in RamsesRT to super-comoving RT variables for the photon transport:

$$\tilde{N} = a^3\, N, \qquad \tilde{\mathbf{F}} = \frac{a^4}{L}\, \mathbf{F}, \qquad \tilde{c} = \frac{a}{L}\, c.$$

The dilution ($\propto a^{-3}$) of photon number density is thus accounted for, while it can easily be verified that Eqs. (3.5)-(3.7) remain unchanged with the new variables – including the reduced flux (3.10) used in the M1 tensor (3.9). In the photon injection step (§3.2.2) the injected photon densities are transformed straight to super-comoving densites, and for the thermochemistry step (§3.2.3) a transformation is made to physical scales and back to super-comoving for the relevant variables.

Note that when reduced light speed is used, the photons will be over-diluted in cosmological simulations, since the time taken for them to get from source to destination will be overestimated. Note also that wavelength stretching with redshift, which in reality adds a fourth power of a to the dilution of N_γ, is not accounted for here. This is actually non-trivial to do: One could add one power of a to the definitions of $\tilde{N}$ and $\tilde{\mathbf{F}}$, but it would be a very crude approximation of the wavelength dilution, as the wavelength shift that should feed photons from one package to the other is neglected. In any case, this effect is likely to be important only in the context of reionization, where the photons have a chance of travelling cosmological distances before they are absorbed. To the best of our knowledge, no published RT or RHD simulations (except Rosdahl and Blaizot, 2012) account at all for the effect of cosmological expansion on the photon transport.

3.4 Putting the RT into RamsesRT

We have chosen the Ramses cosmological code to host our radiative transfer implementation. Using this strategy of including the RT in an existing host code is very practical in the sense that we don't have to do everything from scratch. The hydrodynamical nature of the moment method allows us to take advantage of an existing structure of the host for the hydrodynamical solver, the grid structure and parallellization. For post-processing, the RT implementation can be used directly on the host output, and for cosmological RHD simulations we don't have to implement all the other stuff which is relevant in these simulations (DM, stars and their formation, gravity solver, hydrodynamics).

Figure 3.3: *An oct – the basic grid element in* Ramses.

We have chosen Ramses as our host because it is a widely used and thoroughly tested cosmological code, with a large and growing number of scientific publications to show for. We will first describe the basics of Ramses, where we focus on the hydrodynamics and the grid structure, which are most relevant to us, and then describe how we fit radiative transfer inside it.

3.4.1 The AMR structure of Ramses

Ramses (Teyssier, 2002) is a cosmological adaptive mesh refinement (AMR) code that can simulate the evolution and interaction of dark matter, stellar populations and baryonic gas via gravity, hydrodynamics and radiative cooling (see §2.1.1). It can run on parallel computers using the message passing interface (MPI) standard, and is highly optimized to run very large numerical experiments. It is used for cosmological simulations in the framework of the expanding Universe, and also smaller scale simulations of more isolated phenomena, such as the formation and evolution of galaxies, clusters, and stars. Dark matter and stars are modelled as collisionless particles that move around the simulation box and interact via gravity. We will focus here on the hydrodynamics of Ramses though, which is where the UV radiative transfer couples to the simulations.

Ramses solves the Euler equations (2.1)-(2.3) in their conservative form, employing a second-order Godunov method (Toro, 1999) on the discretized form we have shown in (3.18), except the intercell fluxes are time-centered as well as space-centered. The equations are solved across an AMR grid structure. Fortunately for us, operator splitting is employed for the thermochemistry source terms, i.e. they are split from the rest of the Euler equations in the numerical implementation – which makes it easy for us to modify the thermochemistry solver, i.e. change it from equilibrium to non-equilibrium.

The basic grid element in Ramses is an oct (Fig. 3.3), which is a grid composed of eight cells (in three dimensions; four cells in 2D and two in 1D). Each cell stores the gas properties of density, momentum, internal energy and metallicity, and is also used for calculating and storing the gravitational force. Each cell in the oct can be recursively refined to contain sub-octs, up to a maximum level ℓ of refinement. The whole Ramses simulation box is one oct at $\ell = 1$, which is homogeneously and recursively refined to a minimum refinement level $\ell_{\rm min}$, such that the coarse (minimum) box resolution is $2^{\ell_{\rm min}}$ cells on each side. Octs at or above level $\ell_{\rm min}$ are then adaptively refined during the simulation run, to follow the formation and evolution of structures, up to a maximum refinement level $\ell_{\rm max}$, giving the box a maximum *effective* resolution of $2^{\ell_{\rm max}}$ cells widths per box width. The octs are stored in a tree-like structure where each oct points to it's parent cell, 6 neighbouring parent cells (on each face of the oct) and up to 8 children octs. An important rule concerning the refinement is that the resolution should never change by more than one level across cell boundaries.

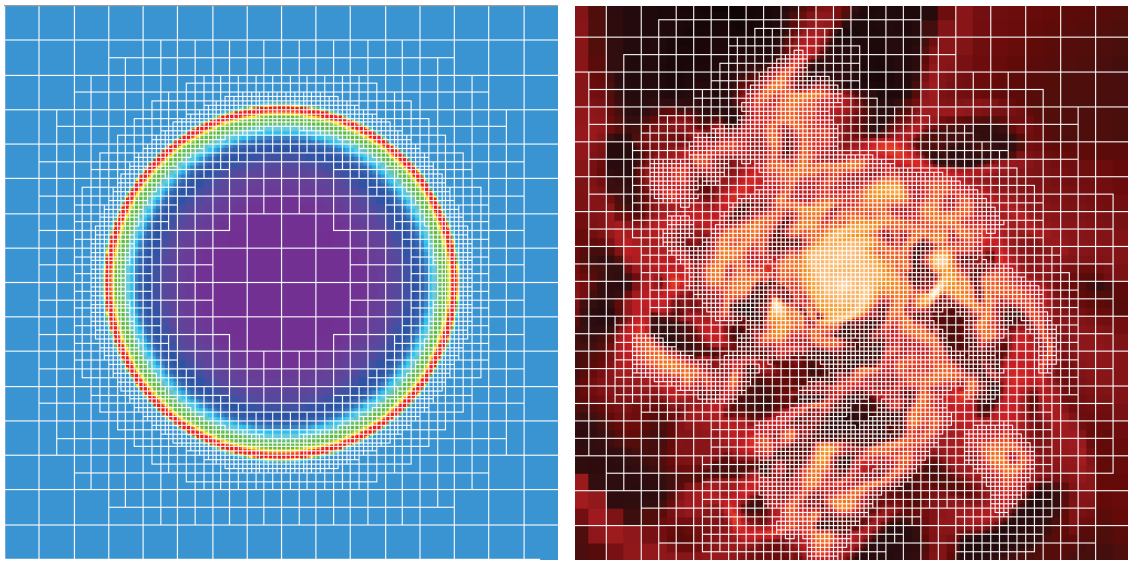


Figure 3.4: *Snapshots from two* Ramses *simulations, illustrating the AMR grid in action: A Sedov explosion on the left and a galaxy simulation to the right. The color scales show gas density and overplotted in white is the grid structure. In each simulation the minimum AMR level is 5, so at the coarsest level we see 16 grids (32 cells) spanning the box width. The maps also illustrate how the refinement must be gradual in space: It is forbidden to change by more than one refinement level across grid/cell boundaries. Credit: Julien Devriendt (left) and Yohan Dubois (right).*

This rule and the oct setup in general is illustrated in Fig. 3.4, which shows octs overplotted on Ramses simulation outputs.

Ramses multi-stepping approach

Simulation time is advanced in turns over the AMR levels in Ramses. A coarse time-step, over the whole AMR grid, is initiated at the coarse level, $\ell_{\rm min}$, as we show schematically in Fig. 3.5. First, the coarse time-step length $\Delta t_{\ell_{\rm min}}$ is estimated via (the minimum of) Courant conditions in all $\ell_{\rm min}$ cells. Before the coarse step is executed, the next finer level, $\ell_{\rm min}+1$, is told to execute the same time-step, in two substeps since the finer level Courant condition should approximately halve the time-step length. This process is recursive: The next finer level makes its own time-step estimate (Courant condition, but also $\Delta t_\ell \leq \Delta t_{\ell-1}$) and asks its next finer level to execute two substeps. This recursive call up the level hierarchy continues to the highest available level $\ell_{\rm max}$, which contains only *leaf* cells and no sub-octs. Here the first two substeps are finally executed, with step lengths $\Delta t_{\ell_{\rm max}} \leq \Delta t_{\ell_{\rm min}}/2^{\ell_{\rm max}-\ell_{\rm min}}$. When the two $\ell_{\rm max}$ substeps are done, the $\ell_{\rm max}-1$ time-step is re-evaluated to be no longer than the sum of the two substeps just executed at $\ell_{\rm max}$, and then one $\ell_{\rm max}-1$ step is executed. Then back to level $\ell_{\rm max}$ to execute two steps, and so on. The substepping continues in this fashion across the level hierarchy, ending with one time-step for the coarsest level cells (with a modified time-step length $\Delta t_{\ell_{\rm min}}$).

At the heart of Ramses lies a recursive routine called `amr_step`(ℓ) which describes a single time-step at level ℓ, and is initially called from the coarsest level ($\ell_{\rm min}$). To facilitate our descriptions on how the RT implementation is placed into Ramses, we will now go through

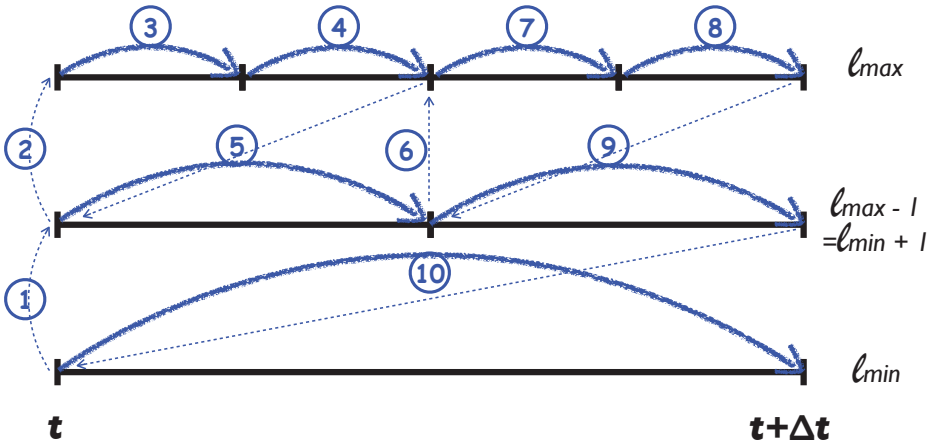


Figure 3.5: *Recursive time-stepping over one coarse time-step in the AMR levels of* `Ramses`*, here shown for a three-level AMR structure. Each solid blue arrowed line represents a time-step which is executed for all cells belonging to the corresponding AMR level. The numbers indicate the order of the time-stepping, including the calls to finer levels (1, 2, and 6). The time-stepping is initiated from the coarsest level ($\ell_{\rm min}$), but the first sub-step is executed at the finest level ($\ell_{\rm max}$).*

the operations of the `amr_step` routine, in order to later show precisely where the RT calls are placed in the scheme of things. It is shown in pseudocode format in Listing 3.1, where we have excluded details and bits not directly relevant to the hydrodynamics and RT (e.g. MPI syncing and load-balancing, adaptive refinement and de-refinement, particle propagation, gravity solver, star formation, and stellar feedback).

Let us go through the `amr_step`(ℓ) line by line. In **line 2** the maximum time-step is determined via Courant-type conditions, to be less than any kind of crossing speed in any cell at level ℓ, but not larger than half the coarser level time-step, so the coarser level will be able to catch up in one time-step. In **line 3** a buffer state $\widetilde{\mathcal{U}}$ is created for all level ℓ cells, which is a copy of their state $\mathcal{U}$, containing the hydrodynamical variables of density, velocity and internal energy, and has the purpose of maintaining consistency during the hydro transport: Whenever a cell state is updated via the intercell fluxes, these fluxes are calculated from the original state, $\mathcal{U}$, but the update is performed on $\widetilde{\mathcal{U}}$, such that the update is accumulated into this buffer via all 6 cell interfaces. If there was no $\widetilde{\mathcal{U}}$ and the update was performed instead on $\mathcal{U}$, the final result would depend on the order of which the intercell updates are performed, which would make it inconsistent and wrong. **Lines 4-6** contain the AMR recursion: If the finest available level has not been reached, all finer levels are made to perform their AMR step. Once the finer levels are done, control comes back, and in **line 7** the hydro transport is executed, where $\widetilde{\mathcal{U}}_\ell$ is updated in all level ℓ leaf cells to reflect the intercell fluxes calculated from $\mathcal{U}_\ell$. Once the hydro-transport is done, the update is made final in **line 8** with $\widetilde{\mathcal{U}}_\ell \rightarrow \mathcal{U}_\ell$. **Line 9** calls the thermochemistry step. Here the gas internal energy, $\mathcal{E}$, is updated (in $\mathcal{U}_\ell$) in all level ℓ leaf cells. Finally, **line 10** re-sets the coarser level time-step, to sync it with the two time-steps that have just been executed at the current level.

Listing 3.1: *The AMR step in* Ramses.

```
1 recursive subroutine amr_step(ℓ)
2    Δt_ℓ = min(Courant_dt(ℓ), Δt_{ℓ-1}/2)
3    set Ũ_ℓ = U_ℓ for all level ℓ cells
4    if ℓ < ℓ_max and any cells exist in ℓ+1
5       call amr_step(ℓ+1)
6       call amr_step(ℓ+1)
7    call hydro_solver(ℓ): all U_ℓ → Ũ_ℓ and neighbouring U_{ℓ-1} → Ũ_{ℓ-1}
8    set U_ℓ = Ũ_ℓ for all level ℓ cells
9    call eq_thermochemistry on all U_ℓ → U_ℓ
10   if second time here: Δt_{ℓ-1} = Δt_ℓ^previous + Δt_ℓ
11 end
```

We note an important aspect of the hydrodynamical solver in the AMR framework: As noted in line 7 in Listing 3.1, a level ℓ step also includes intermediate updates on all neighbouring coarser level $(\ell-1)$ cells. The coarser level update is only *partial* though, because it only reflects the intercell fluxes across inter-level boundaries, and fluxes across other boundaries (same level or next coarser level) will only be accounted for when the coarser level time-step is advanced. Until then, these coarser level neighbour cells have gas states that are not well defined, since they only reflect some of their intercell fluxes. This is shown schematically in Fig. 3.6 – it is very important to keep in mind when considering the coupling of RT with the hydrodynamics of Ramses.

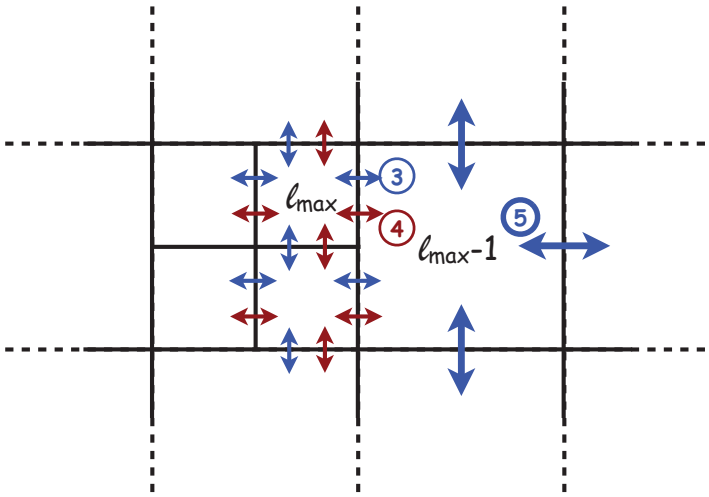


Figure 3.6: *Level ℓ gas state updates via intercell fluxes also perform partial gas updates in neighbouring cells at level $\ell-1$. The example shown corresponds to the hierarchy from Fig. 3.5. Steps 3 and 4 at the finest level also include partial updates of neighbouring $\ell_{\rm max}-1$ cells, but these neighbour cell states are not fully updated until all the intercell fluxes are taken into account, which is in step 5 from Fig. 3.5.*

Let us briefly acknowledge the precise need for the particular ordering of the lines in `amr_step`(ℓ): The time-step determination, line 2, must to be done before the recursion, such that finer cells know the maximum time-step length they are allowed. The creation of the $\widetilde{\mathcal{U}}_\ell$ buffer state must obviously be done before the hydro-solver, but it must also be done before the recursion, so that inter-level cells can be partially updated via their inter-level fluxes (Fig. 3.6). The recursion, lines 4 to 6, must be done before the hydro-solver, because the current level time-step length is not finally determined until the finer level steps are finished (line 10). The hydro-solver needs access to two (before- and after-) copies of the cell states, so it needs to come before $\mathcal{U}_\ell$ is discarded (line 8). Lines 8 - 10 are then pretty much interchangeable, except if lines 8 and 9 are interchanged, then the thermochemistry step must use $\widetilde{\mathcal{U}}_\ell$ rather than $\mathcal{U}_\ell$. The thermochemistry step must come before or after any hydro solver updates on $\widetilde{\mathcal{U}}_\ell$ are made, otherwise the cell state is not well defined; because of the $\widetilde{\mathcal{U}}_{\ell-1}$ updates, this also makes it illegal for the thermochemistry to be done before the hydro solver.

Having put down the basics of AMR hydrodynamics, we are now in a position to add radiative transfer. We will begin with `RamsesRT` postprocessing of the radiative field.

3.4.2 Radiative transfer post-processing

In `RamsesRT`, new variables are added to $\mathcal{U}$: The three ionization fractions, x_{HII}, x_{HeII} and x_{HeIII}, which are stored as multiples of the gas density, and $4M$ photon field variables, i.e. 4 variables of photon density and flux vector for each photon package. Although there is no difference in how the variables are stored, we conceptually split $\mathcal{U}$ into two parts:

- $\mathcal{U}^{\mathrm{h}}$ consists of the usual hydrodynamic variables (ρ, $\mathbf{u}$, $\mathcal{E}$, Z), plus the three ionization fractions (x_{HII}, x_{HeII}, x_{HeIII}).
- $\mathcal{U}^{\mathrm{rt}}$ consists of the $4M$ photon variables (N_i, $\mathbf{F}_i$).

The cell state vector is then $\mathcal{U} = \left(\mathcal{U}^{\mathrm{h}}, \mathcal{U}^{\mathrm{rt}}\right)$.

In RT post-processing, hydrodynamics and particle propagation are turned off, so RT is in effect performed inside an otherwise frozen simulation, and can be done by re-starting from an existing `Ramses` output. Here `amr_step` is not called at all, but instead a similar recursive routine called `rt_step` is called from the coarse level.

Listing 3.2 shows the pseudocode for this routine: The lines of execution and their ordering is basically the same as in `amr_step`. In **line 2** the time-step is calculated via one simple Courant condition. There is no worry here about the time-step changing due to finer levels, since the time-step length is purely a function of the Courant factor, cell width and light speed, all of which are constants. In **line 3** a buffer state $\widetilde{\mathcal{U}}^{\mathrm{rt}}_\ell$ is created for the RT variables. There is no need to buffer the hydro variables, as these are not transported between cells in the RT step. The recursion to finer level photon transport and thermochemistry is done in **lines 4-6**. Then the current level photon transport (§3.2.1) in **line 7**, which also partially updates inter-level boundary cells at level $\ell - 1$. Then the photon injection step (§3.2.2) in **line 8**, where photons are added to the grid ($\widetilde{\mathcal{U}}^{\mathrm{rt}}_\ell$) via e.g. stellar emission. Note that this also involves some $\widetilde{\mathcal{U}}^{\mathrm{rt}}_{\ell-1}$ updates due to stars that move across level boundaries in the second RT sub-step at level ℓ. The photon transport and injection is made final in

Listing 3.2: *The RT postprocessing step in* RamsesRT.

```
1 recursive subroutine rt_step(ℓ)
2    Δt_ℓ = f_c Δx_ℓ / c_r
3    set Ũ^rt_ℓ = U^rt_ℓ for all level ℓ cells
4    if ℓ < ℓ_max and any cells in ℓ+1
5       call rt_step(ℓ+1)
6       call rt_step(ℓ+1)
7    call rt_transport(ℓ): all U^rt_ℓ → Ũ^rt_ℓ and neighbouring U^rt_{ℓ-1} → Ũ^rt_{ℓ-1}
8    call photon_injection_step(ℓ): → Ũ^rt_ℓ and → Ũ^rt_{ℓ-1}
9    if not smooth RT: set U^rt_ℓ = Ũ^rt_ℓ for all level ℓ cells
10   call neq_thermochemistry on all U_ℓ → U_ℓ
11 end
```

line 9 by writing the state variables. Finally in **line 10**, non-equilibrium thermochemistry is performed on all level ℓ cells. It is important to remember that here not only the RT variables are being updated, but also the hydro variables of temperature (or internal energy) and ionization states. This is the vital point where RT couples with the hydrodynamics and this update is actually what makes it very tricky to implement radiation hydrodynamics in an AMR framework, as will soon come clear.

When smooth RT is used (see §3.2.5), line 9 in Listing 3.2 is skipped, i.e. the new state is not made final before the thermochemistry step. This allows the thermochemistry to use the before-transport-and-injection photon densities and fluxes and also allows it to determine their change rates, $(\widetilde{\mathcal{U}}^{\rm rt}_\ell - \mathcal{U}^{\rm rt}_\ell)/\Delta t_\ell$, which are the 'smooth' terms that are added to the thermochemistry calculation. The effect of photon transport and injection thus only comes through after the thermochemistry step.

Let us briefly again consider the ordering of commands: Here the transport step can actually be called before the recursion (the two calls to `rt_step` in lines 5 and 6), since all the time-step lengths are known in advance. This means the transport could in principle be done from the coarsest level up instead of from the finest level down as with the hydrodynamics. However, this would only further complicate the buffer states, as the coarser levels would still have to wait for the finer levels to do those inter-level updates. For simplicity and consistency with the `amr_step` we keep the finest level down stepping. Once that is set, we still need to create the buffer state before the recursion and do the transport after. As with the hydrodynamics the thermochemistry must be done after the transport. The photon injection step could in principle be anywhere after the recursion, as the $\mathcal{U}^{\rm rt}$ update here doesn't depend on $\widetilde{\mathcal{U}}^{\rm rt}$.

Although we do not show it in the pseudocode, adaptive refinement can be performed on the grid via gradient criteria on photon densities and ionization fractions. We demonstrate this in some of the tests described in Chapter 4. The details, e.g. of refinement, the AMR grid bookkeeping and MPI communication is made very easy by the fact that we have simply added new variables to the pre-existing structure of `Ramses`, so while we have to change some indexing and be careful to make calls in the right places to keep the bookkeeping in order (which are basically the same places as in `amr_step`), the effort is minimal considering

the fact that we can perform radiative transfer in an AMR structure on parallel computers. RT postprocessing works well in RamsesRT, we have thoroughly tested it with the benchmark tests of Iliev et al. (2006a) (see Chapter 4) and used it for some preliminary applications, like the mapping of gas ionization state and Lyα emissivity in simulations of isolated galaxies. We note that RamsesRT can perform in two as well as three dimensions – having the possibility of two dimensional simulations is very convenient for making simple tests and debugging.

3.4.3 Radiation hydrodynamics

Then for on-the-fly radiative transfer in RamsesRT. Here we need to do both RT and hydrodynamics (and DM propagation, gravity etc.) at the same time. This is made tricky by the combination of a timescale difference and the nested time-stepping in the AMR structure. We describe first the obvious and preferred way of doing things, which alas proves impossible so far to implement in a self-consistent manner. Then we present two alternative coupling setups that do work, but are less optimal speed-wise.

First a mention about the advection of the ionization fractions: These variables are stored as multiples of the gas densities, in the form of passive scalars that are advected with the gas. The gas metallicity is already such a passive scalar that is advected in Ramses, and one can simply add further variables (always multiples of density) that work in the same way. This is also the reason that these ionization fractions belong in $\mathcal{U}^{\rm h}$ rather than $\mathcal{U}^{\rm rt}$.

RT subcycling: The preferred but illegal coupling method

In view of the large timescale difference between RT and hydrodynamics, which is due to the large speed of light compared to the maximum wave-speeds in the gas, it appears natural and ideal to subcycle the radiative transfer within each hydrodynamical time-step at the finest hydro-stepping scale. For each $\ell_{\rm max}$ hydrodynamical step, an RT step should be performed *over the whole grid, over all levels*. The problem is that during the coarse hydro time-step, in between the finer level hydro substeps, there are cells with undefined hydrodynamical states, which makes it non-trivial for RT-thermochemistry to update the temperature and the ionization states.

We can take a concrete look by attempting to insert the RT step inside the AMR step so that it is subcycled between the fine-scale hydrodynamics. There is really only one place where calling the RT step fits, which is after the recursion but before the hydro-solver. It has to be after the recursion because until then the final time-step length for the RT subcycling is not determined. It has to be before the hydro-solver for consistency: The RT step includes thermochemistry, so if it is called after the hydro-solver, thermochemistry happens after hydro-transport at $\ell_{\rm max}$ but before it at coarser levels (where hydro-transport has yet to happen). Thus we arrive at the pseudocode setup shown in Listing 3.3.

It seems simple enough, but here is the problem: At the point that the RT step is called, every cell in the grid exists with two hydrodynamic states; the 'source' state $\mathcal{U}^{\rm h}$ and the 'destination' state $\widetilde{\mathcal{U}}^{\rm h}$. Which one of those should be updated via RT thermochemistry? Choosing one but not the other leads to an obvious and severe inconsistency between the source and destination states. If thermochemistry does the update $\widetilde{\mathcal{U}}^{\rm h} \rightarrow \widetilde{\mathcal{U}}^{\rm h}$, then a gas element which is transported from one cell to it's neighbour during the following hydro

Listing 3.3: *The AMR step with RT step subcycling, which is ideal but illegal*

```
recursive subroutine amr_step(ℓ)
   Δt_ℓ = min(Courant_dt(ℓ), Δt_{ℓ−1}/2)
   set Ũ^h_ℓ = U^h_ℓ for all level ℓ cells
   if ℓ < ℓ_max and any cells in ℓ+1
      call amr_step(ℓ+1)
      call amr_step(ℓ+1)
   else if RT: call rt_step(ℓ_min)
   call hydro_solver(ℓ): all U^h_ℓ → Ũ^h_ℓ and neighbouring U^h_{ℓ−1} → Ũ^h_{ℓ−1}
   if not RT: call neq_thermochemistry on all Ũ_ℓ → Ũ_ℓ
   set U^h_ℓ = Ũ^h_ℓ for all level ℓ cells
   Δt_{ℓ−1} = Δt^pre_ℓ + Δt_ℓ
end
```

transport is not thermochemically evolved over the time-step, because it originates from $\mathcal{U}^{\mathrm{h}}$. If instead the update is done on $\mathcal{U}^{\mathrm{h}} \rightarrow \mathcal{U}^{\mathrm{h}}$, a gas element which stays still in a cell over the following hydro transport step is not thermochemically evolved over the time-step. One might then just update both states via thermochemistry, i.e. apply it on each cell twice. This does not really make sense for these inter-level intercell boundary cells that have $\mathcal{U}^{\mathrm{h}} \neq \widetilde{\mathcal{U}}^{\mathrm{h}}$, as $\widetilde{\mathcal{U}}^{\mathrm{h}}$ doesn't represent a true state but is rather an intermediate and temporary quantity that exists between well-defined times. Also, it would be really non-trivial to implement: Applying thermochemistry on each of the states also implies transporting the photons through two different states in each cell, which creates alternate time-lines for the radiative transfer! We have thus abandoned this pre-supposed ideal way of RT-hydro coupling and gone for two alternatives which sidestep the problem of undefined gas states. In one method we prohibit hydro-subcycling over the AMR levels, and in the other we prevent the RT subcycling within the fine-scale hydro step by limiting the hydrodynamical time-step down to the RT one.

Hydro-single-stepping with RT subcycling

The only way of retaining the subcycling of the RT steps within the hydro step, while avoiding the problem of ill-defined hydro-states, is to do the hydrodynamical update in a single time-step over the whole AMR grid hierarchy (which is an option in `Ramses`). An `amr_step` is performed over the whole grid, on a hydro time-step length which is set by the finest level, and then RT is subcycled over the grid to match the hydro time-step. Here the hydro states are final and well-defined everywhere when the RT is called. We show the pseudocode for `amr_step` and `rt_step` in Listings 3.4 and 3.5, and a schematic example of a coarse time-step is shown in Fig. 3.7. The basics of the coarse time-step consist of:

```
call amr_step(ℓ_min) → Δt_{ℓ_max}
while Δt_{ℓ_max} > 0:
   call rt_step(ℓ_min, Δt_{ℓ_max})
   Δt_{ℓ_max} = Δt_{ℓ_max} − Δt^rt_{ℓ_min}
end
```

Listing 3.4: *Hydro single-stepping: The AMR step.*

```
recursive subroutine amr_step(ℓ)
    Δt_ℓ = min(Courant_dt(ℓ), Δt_{ℓ-1})
    set Ũ^h_ℓ = U^h_ℓ for all level ℓ cells
    if ℓ < ℓ_max and any cells in ℓ+1
        call amr_step(ℓ+1)
    call hydro_solver(ℓ): all U^h_ℓ → Ũ^h_ℓ and neighbouring U^h_{ℓ-1} → Ũ^h_{ℓ-1}
    if not RT: call neq_thermochemistry on all Ũ_ℓ → Ũ_ℓ
    set U^h_ℓ = Ũ^h_ℓ for all level ℓ cells
    Δt_{ℓ-1} = Δt_ℓ
end
```

Listing 3.5: *Hydro single-stepping: The RT step.*

```
recursive subroutine rt_step(ℓ, Δt)
    Δt_ℓ = min(f_c Δx_ℓ/c_r, Δt)
    set Ũ^rt_ℓ = U^rt_ℓ for all level ℓ cells
    if ℓ < ℓ_max and any cells in ℓ+1
        call rt_step(ℓ+1, Δt_ℓ/2)
        call rt_step(ℓ+1, Δt_ℓ/2)
    call rt_transport(ℓ): all U^rt_ℓ → Ũ^rt_ℓ and neighbouring U^rt_{ℓ-1} → Ũ^rt_{ℓ-1}
    call photon_injection_step(): → Ũ^rt_ℓ and → Ũ^rt_{ℓ-1}
    if not smooth RT: set U^rt_ℓ = Ũ^rt_ℓ for all level ℓ cells
    call neq_thermochemistry on all U_ℓ → U_ℓ
end
```

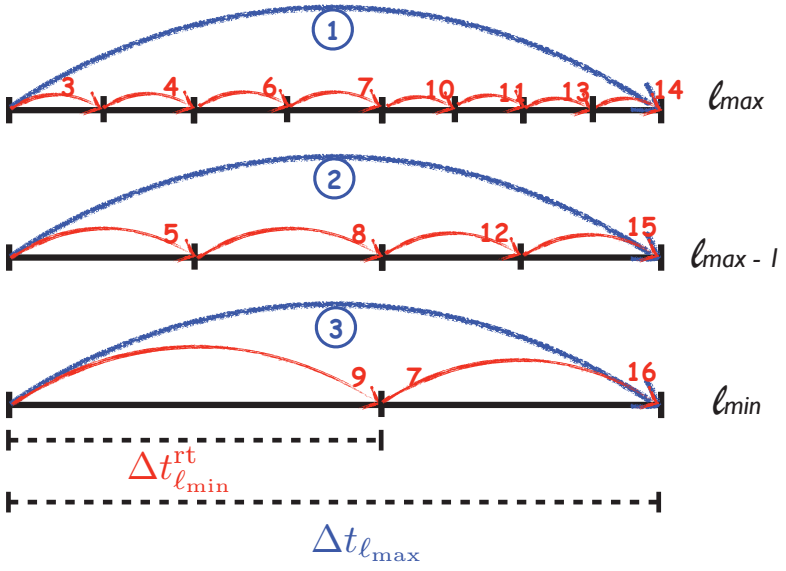


Figure 3.7: *Hydro-single-stepping with RT subcycling. We show an example hydro step where the finest-level time-step length is almost two times larger than the coarse level RT time-step length. The hydro step is first done on all levels, without sub-stepping. Once finished, RT fills the same time-step over the whole grid, with subcycling on each level.*

Not allowing subcycling of the hydrodynamics is a very severe limitation though, and in effect removes one of the two main advantages of AMR: It is not just about adaptive resolution in space but also in time, and here we are throwing away the time-refinement. In practice this has a devastating effect on the run-time of simulations when the AMR hierarchy spans many levels. If there is only a handful of levels however, this is a good option. In fact, we have run most of the RT benchmark tests with this method, as we only do one level refinement in these tests. For large simulations with lots of refinement levels, the method we describe next is a better option.

RHD with an RT time-step

The other alternative for RamsesRT radiation-hydrodynamics, is to abandon the RT-within-hydro subcycling. This can be achieved by either expanding the RT time-step length to match the hydrodynamical time-step, or vice-versa. We're not allowed to relax the Courant time-step condition on the RT unless we switch to an implicit RT transport – and we haven't done that. So the option we take here is to reduce the hydro time-step down to the RT one. Of course if we stick to the physical speed of light, this means reducing the hydro time-step by ~ 3 orders of magnitude, which is unrealistic in simulations, and likely creates a lot of hydrodynamical diffusion to add. However, if we match this with a speed of light which is reduced by, say, a factor of 100, this only means reducing the hydro time-step by a factor of ten or so, which is starting to become borderline manageable. This is in fact the most feasible way we have found to run large scale RHD simulations – the time increase when using a light speed fraction of $1/100$ and three photon packages typically is about a factor of ten, and this can basically be reduced to $\lesssim$ a factor of two if the light speed fraction is low enough that the hydro time-step doesn't have to be modified at all. A schematic of this setup is shown in Fig. 3.8 and the pseudocode for the amr_step, which now has been 'merged' with the rt_step to include the RT transport, injection and thermochemistry steps, is shown in Listing 3.6.

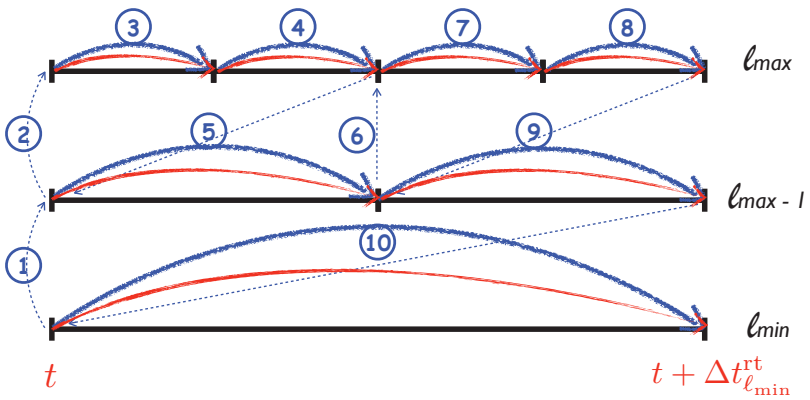


Figure 3.8: *RHD with RT time-step. This is much like the normal* amr_step *in* Ramses*, except that the time-step length has the extra constraint of the light speed Courant condition, and each level ℓ step also performs RT transport, photon injection and thermochemistry over the same time-step and level.*

Listing 3.6: *RHD with RT time-step in* `RamsesRT`.

```
recursive subroutine amr_step(ℓ)
   Δt_ℓ = min(Courant_dt(ℓ), Δt_{ℓ-1}/2, f_c Δx_ℓ/c_r)
   set Ũ_ℓ = U_ℓ for all level ℓ cells
   if ℓ < ℓ_max and any cells in ℓ+1
      call amr_step(ℓ+1)
      call amr_step(ℓ+1)
   call hydro_solver(ℓ): all U^h_ℓ → Ũ^h_ℓ and neighbouring U^h_{ℓ-1} → Ũ^h_{ℓ-1}
   call rt_transport(ℓ): all U^rt_ℓ → Ũ^rt_ℓ and neighbouring U^rt_{ℓ-1} → Ũ^rt_{ℓ-1}
   call photon_injection_step(): → Ũ^rt_ℓ and → Ũ^rt_{ℓ-1}
   set U^h_ℓ = Ũ^h_ℓ for all level ℓ cells
   if not smooth RT: set U^rt_ℓ = Ũ^rt_ℓ for all level ℓ cells
   call neq_thermochemistry on all U_ℓ → U_ℓ
   Δt_{ℓ-1} = Δt_ℓ^previous + Δt_ℓ
end
```

3.5 Thermochemistry

Here we describe the thermochemistry in `RamsesRT`. To establish the framework, we'll begin with a basic description of the *equilibrium* thermochemistry that is already in place in `Ramses`. We'll then argue why it is insufficient for the inclusion of RT and finally describe in detail the numerical implementation of the *non-equilibrium* thermochemistry we've made from scratch for `RamsesRT`.

The game is to solve for the radiative cooling and chemistry terms of the conservation of energy equation (2.3), i.e. to evolve the internal energy density – or temperature – of a single cell via the radiative processes of heating and cooling. This is purely a single cell process with no outside effect from neighbouring cells, which is a big advantage as the cooling in each cell can trivially be computed in parallel (which is not at all the case for the advection).

3.5.1 `Ramses` equilibrium thermochemistry - what's already there

We take a cell of gas with a given internal energy density ε, mass density ρ and metallicity Z, and we want to update the cell energy over a time-step Δt. Z and ρ are updated elsewhere (during the advection step) and can be taken as constants during the thermochemistry step. First, we convert the internal energy to a temperature. The temperature of the cell is given by

$$T = \varepsilon \, \frac{(\gamma - 1) m_H}{\rho k_{\mathrm{B}}} \, \mu, \tag{3.38}$$

where γ is the ratio of specific heats (usually given the value of $5/3$ in `Ramses`, corresponding to monatomic gas), m_H the proton mass, k_{B} the Boltzmann constant and μ is the average mass per particle in the gas, in units of m_H. Since the ion species are not stored in the normal version of `Ramses`, μ is not readily available. Thus, the quantity that is really evolved in

Ramses thermochemistry is

$$T_\mu \equiv \frac{T}{\mu}, \tag{3.39}$$

which can be directly extracted from the variables stored in the cell. We start at time t with temperature T_μ^t and wish to evolve it to $T_\mu^{t+\Delta t}$, where the superscript is for denoting *when* T_μ is evaluated.

Cooling and heating rates of the gas are functions of temperature, density, redshift (via redshift-dependent UV background and CMB radiation), metallicity, and the abundances of each primordial ion species, $n_{\rm HI}$, $n_{\rm HII}$, $n_{\rm HeI}$, $n_{\rm HeII}$, $n_{\rm HeII}$, and $n_{\rm e}$. In (non-RT) Ramses however photoionization equilibrium (PIE) is assumed, which means the primordial ion abundances are surjective functions of temperature, density, and redshift, that can be calculated with a simple iterative process that involves equating the rates of photo-ionization, collisional ionization and recombination. The cooling and heating rates in the equilibrium thermochemistry are then reduced to being functions only of temperature, density, redshift and metallicity.

These rates are pre-computed and stored in tables every coarse time-step for a range of $(T_\mu, n_{\rm H})$-bins, where $n_{\rm H} = X\rho/m_H$ is hydrogen number density and X is the hydrogen mass fraction in the gas (a global constant, typically set to the value of 0.76). These rates can then be quickly picked up when evolving the cell temperature. A redshift-dependent UV background of ionizing radiation is assumed, but since it is homogeneous, i.e. exactly the same in every cell, it suffices to recompute the cooling tables every few time-steps, as the redshift changes.

Using T_μ and $n_{\rm H}$ as look-up indexes, the following rates, all given in [erg cm^3 s^{-1}], are fetched on-the-fly from the precomputed tables:

- Heating rate $\mathcal{H}(T_\mu, n_{\rm H})$: The heating contribution of the UV background at the current redshift.
- Primordial cooling rate $\mathcal{L}(T_\mu, n_{\rm H})$, i.e. cooling rate of a mixture of H and He (and electrons) of primordial composition.
- Metal cooling rate contribution $\mathcal{L}_Z(T_\mu, n_{\rm H})$, containing the per-solar-metallicity cooling contribution of metals in the gas.

With these three rates in hand the temperature is updated by solving:

$$\frac{\partial T_\mu}{\partial t} = \frac{(\gamma-1)m_{\rm H}}{\rho k_{\rm B}}\,\Lambda, \tag{3.40}$$

where $\Lambda \equiv \dot{\varepsilon} = (\mathcal{H} + \mathcal{L} + Z/Z_\odot\,\mathcal{L}_Z)\,n_{\rm H}^2$. The update is done with a semi-implicit Euler formulation (SIE, see Press et al., 1992):

$$T_\mu^{t+\Delta t} = T_\mu^t + \frac{\Lambda K \Delta t}{1 - \Lambda' K \Delta t}, \tag{3.41}$$

where $K = \frac{(\gamma-1)m_{\rm H}}{\rho k_{\rm B}}$ is the conversion factor between ε and T_μ and $\Lambda' \equiv \frac{\partial \Lambda}{\partial T_\mu}$ can be estimated by finite-differencing the rate tables.

So the integration of T_μ over the time-step is quick and painless, and basically just involves interpolations of the rate-tables to retrieve $\mathcal{L}$, $\mathcal{H}$ and $\mathcal{L}_Z$ for the given temperature and density. The expensive calculations happen between the time-steps, where the rate tables are updated, the advantage being they are done a lot less than if they were done on-the-fly for every cell.

Filling in the rate tables

First photoionization rates Γ_i $[\mathrm{s}^{-1}]$ and photoheating rates $\mathcal{H}_i$ $[\mathrm{erg\,s}^{-1}]$ are calculated for HI, HeI, HeII and e (Compton heating in the case of electrons). These rates depend *only* on the UV background, so they are homogeneous functions of redshift, i.e. apply to every cell.

The abundances table: Then, since PIE is assumed, calculation and tabulation is performed per (T_μ, n_H)-bin of the abundances n_i, of each of the 6 primordial species (e, HI, HII, HeI, HeII, HeIII). Here rates are used for all the possible interactions involving these species - these are functions of Γ_i, T and n_i, and amount to a closed set of equations that are converged iteratively to an equilibrium solution, such that the creation rate equals the destruction rate for each species. The species abundances also give the value of μ per (T_μ, n_H)-bin, which can be retrieved by

$$\mu = [\, X(1 + x_\mathrm{HII}) + Y/4(1 + x_\mathrm{HeII} + 2x_\mathrm{HeIII})\,]^{-1}\,, \tag{3.42}$$

where $Y = 1 - X$ is the helium mass fraction, $x_\mathrm{HII} \equiv n_\mathrm{HII}/n_\mathrm{H}$, $x_\mathrm{HeII} \equiv n_\mathrm{HeII}/n_\mathrm{He}$ and $x_\mathrm{HeIII} \equiv n_\mathrm{HeIII}/n_\mathrm{He}$. The tabulation of the abundances also provides a direct mapping between T_μ and T which is useful in generating the rest of the tables.

The cooling rates table: Given the abundances, it is then a straightforward matter to calculate and tabulate $\mathcal{L}(T_\mu, n_\mathrm{H})$ – the cooling rate is a sum of bremsstrahlung-, collisional excitation-, collisional ionization-, recombination-, dielectric recombination- and Compton- cooling rates, all fitted analytic functions of temperature and abundances, and fetched from various sources (e.g. Cen, 1992).

The heating rates table: It is also straightforward to tabulate $\mathcal{H}(T_\mu, n_\mathrm{H})$. Each bin contains

$$\mathcal{H} = \sum_i \mathcal{H}_i n_i, \tag{3.43}$$

where the sum is over the primordial ion species, and $\mathcal{H}_i$ are photoheating rates for the individual species (n_HI, n_HeI, n_HeII, and e).

The metal-cooling-contributions table: `Ramses` keeps a hard-coded table of a precomputed metal-cooling rate contribution, $\mathcal{L}_Z^{CIE}(T)$, which is the difference between zero metallicity and solar metallicity cooling rates calculated assuming collisional ionization equilibrium (CIE) [3] with the `Cloudy` software package (Ferland et al., 1998). That is,

$$\mathcal{L}_Z^{CIE}(T) \;=\; \mathcal{L}_Z^{Cloudy}(T,\, Z = Z_\odot,\, UV = 0) \;-\; \mathcal{L}_Z^{Cloudy}(T,\, Z = 0,\, UV = 0). \tag{3.44}$$

[3] i.e. chemical equilibrium in zero ionizing radiation.

These rates are computed for a photoionization-free environment so they don't depend on gas density. Using this, the photoionization equilibrium (PIE) metal-cooling rates are approximated and tabulated as

$$\mathcal{L}_Z(T_\mu, n_\mathrm{H}) \;=\; \mathcal{L}_Z^{CIE}(T) \;\times\; f(T, n_\mathrm{H}, z), \tag{3.45}$$

where f is a dimensionless analytic function built by Stephanie Courty, that corrects for density n_H and UV background photoionization at redshift z.

3.5.2 RamsesRT non-equilibrium thermochemistry

In the RamsesRT implementation, it becomes questionable to stick to the approximation of photoionization equilibrium. Assuming PIE would still allow one to approximately follow the I-fronts, but photon conservation would not be met as the evolution from neutral to ionized state would not be tracked in time. In order to self-consistently evolve the photons and their interactions with the gas we therefore implement non-equilibrium chemistry that keeps track of the abundances of all ion species of hydrogen and helium. These abundances are tracked in the form of the three ionization fractions, x_HII x_HeII, and x_HeIII. The non-equilibrium thermochemistry evolves these ionization fractions along with the temperature and photon densities and fluxes (via absorption/emission) on a per-cell basis.

Here the species abundances are not just functions of T_μ and n_H, but also of their own history, and of inhomogeneous photoionization rates. As a result it becomes unfeasible to keep redshift-dependent tables of cooling rates, unless we want also to add ionization states and photon fluxes as new dimensions to the tables (and we don't want to do that!). So here we do all calculations for hydrogen and helium cooling and chemistry on the fly.

For the metal contribution to cooling, we still assume the same redshift-evolving homogeneous PIE at this point, in the same way as described in the previous section. Equilibrium may actually be a much better approximation for the metals than for hydrogen and helium: As discussed in Anninos et al. (1997), the abundances of rare species (the metals) tend to quickly adapt to changes in the more abundant (primordial) species, and are thus better described by PIE. Also, the metals contribute negligibly to the electron reservoir. Another matter is the fact that the metal contribution to cooling still assumes a homogeneous redshift-dependent UV background (from the model of Haardt and Madau, 1996): Coupling metal-cooling to the inhomogeneous UV photons is a future task for RamsesRT development.

Now the thermochemistry is not just about evolving the temperature. Here we must

numerically advance the state of a gas cell

$$\mathcal{U} = \begin{pmatrix} T_\mu \\ x_{\mathrm{HII}} \\ x_{\mathrm{HeII}} \\ x_{\mathrm{HeIII}} \\ N_1 \\ .. \\ N_M \\ F_1 \\ .. \\ F_M \end{pmatrix}. \tag{3.46}$$

where in addition to T_μ from equilibrium-`Ramses` we have the three ionization fractions, and $2 \times M$ variables of cell photon densities N_i and flux magnitudes F_i for each of the M photon packages (which are typically three). As before the gas density and metallicity is included in the calculation, but not evolved. Also the change-rate terms of photon density and flux are included as source terms in the case of smooth-RT (§3.2.4).

In each thermochemistry time-step Δt we need to solve

$$\frac{\partial \mathcal{U}}{\partial t} = \mathcal{S}, \tag{3.47}$$

where the rates in $\mathcal{S} \equiv \dot{\mathcal{U}}$ are given by inter-dependent expressions, meaning the $4 + 2 \times M$ equations in (3.47) are coupled. The thermochemistry equations (3.47) are also *stiff*, meaning that explicit numerical solutions like the explicit (or forward) Euler one,

$$\mathcal{U}^{t+\Delta t} = \mathcal{U}^t + \mathcal{S}^t \Delta t,$$

tend to be unstable and barge away in wrong directions. The way to go to achieve stability is to use the forward Euler formulation, where the forward-in-time (FW) state appears on the right hand side:

$$\mathcal{U}^{t+\Delta t} = \mathcal{U}^t + \mathcal{S}^{t+\Delta t} \Delta t. \tag{3.48}$$

This is quite tricky - the FW state is exactly what we're looking for in the first place! The fully implicit solution involves algebraically solving (3.48), which is a daunting (even hopeless) task for this set of equations. A way around that, where we *estimate* the final state by differentiation, is the SIE formulation (also used in equilibrium-`Ramses`, though note $\mathcal{U}$ only includes one variable there),

$$\mathcal{U}^{t+\Delta t} = \mathcal{U}^t + \Delta t (I - \mathcal{J} \Delta t)^{-1} \mathcal{S}^t, \tag{3.49}$$

where $\mathcal{J} = \frac{\partial \mathcal{S}}{\partial \mathcal{U}}$ is the Jacobian matrix, and I the identity matrix. This method is stable for the thermochemistry equations (by experience, though it is not guaranteed to be stable, according to Press et al., 1992), but generally not considered very accurate, being first order. However there are a lot of factors in $\mathcal{S}$ that are far more error-inducing than the first-order integration method so it is perfectly adequate for our purposes[4]. The computational bottleneck of (3.49) is taking the inverse of a matrix that has the same number of dimensions as

[4]Stability of the method is most important, then speed and accuracy comes in last, within sensible limits of course.

there are variables in $\mathcal{U}$, i.e. $4 + 2 \times M$ in our case. It makes the computation expensive and also quite tricky and error-prone, so we have not ventured the SIE path. Instead we have gone for the following compromise.

The RamsesRT thermochemistry sub-steps

Our thermochemistry implementation is inspired by Anninos et al. (1997). Instead of solving the whole set of equations (3.47) in one coupled go like in the SIE formula of (3.49), the idea is to solve one equation at a time in a specific order, and on the RHS use FW values wherever available, but otherwise backwards-in-time (BW) values. So for the first variable we choose to advance in time, there are no FW variables available. For the next one, we can use the FW state of the first variable, and so on. In that sense the method can be thought of as being partially implicit.

We call cell-thermochemistry once every RT time-step of length Δt_{RT}, but in each cell we split that into local sub-steps of length Δt that adhere to the condition

$$max\left(\left|\frac{\Delta \mathcal{U}}{\mathcal{U}}\right|\right) \leq 0.1, \tag{3.50}$$

where $\Delta\mathcal{U}$ is the change in $\mathcal{U}$ during the sub-step. The RT step thus contains a loop for each cell, which calls the `thermo_step`$(\mathcal{U}, \texttt{Δt})$ routine once or more often: First with $\Delta t = \Delta t_{RT}$, then possibly again a number of times to fill in Δt_{RT} if the first guess at Δt proves too long to meet the condition set by (3.50).

The `thermo_step`$(\mathcal{U}, \texttt{Δt})$ routine performs the following tasks:

(a) N and F update

(b) T_μ update

(c) Hydrogen ionized fraction update

(d) Helium ionized fractions update

(e) Check if we are safe to use a bigger time-step

Tasks **(b)** to **(d)** are in the same order as in Anninos et al. (1997), but they don't include radiative transfer in their code, so there is no photon update. The argument we have for putting it first rather than anywhere else is that the photon densities appear to be the most dynamic variables and so are also most likely to break the time-step condition (3.50). This we want to catch early on in the thermochemistry step so we avoid doing calculations of tasks **(b)** to **(e)** that turn out to be useless because of the too-long time-step.

Now for a more elaborate description of the individual tasks:

(a) Photon density and flux update

We update the photon number densities and fluxes, N_i and F_i, one photon package i at a time. We can do it in any order we like since the packages don't interact. For the photon

density the equations to solve are

$$\frac{\partial N_i}{\partial t} = \dot{N}_i + C_i - N_i\, D_i, \tag{3.51}$$

where $\dot{N}_i$ represents the time derivative of N_i given by the RT transport solver (which is nonzero only if the *smooth RT* option is used), C_i represents photon-creating re-combinations, and D_i represents photon-destroying absorptions. The creation term is non-existent if the OTSA is used (emitted photons are assumed to be immediately reabsorbed), but is otherwise given by

$$C_i = \sum_{j}^{\text{HII,HeII,HeIII}} b_{ji}^{rec}\ (\alpha_j^{\text{A}} - \alpha_j^{\text{B}})\ n_j\ n_e, \tag{3.52}$$

where the b_{ji}^{rec} factor is a boolean (1/0) that states which photon package j-species recombinations emit into and α_j^{A} and α_j^{B} are the temperature dependent case A and B recombination rates for the recombining species (see Appendix C). The photon destruction factor is given by

$$D_i = \sum_{j}^{\text{HI,HeI,HeII}} c_r\ \bar{\sigma}_{ij}\ n_j, \tag{3.53}$$

where c_r is the (reduced) light speed and $\bar{\sigma}_{ij}$ is the cross-section between species j and photons in package i.

Photon emission from recombination is assumed to be spherically symmetric, i.e. to go in all directions. It is therefore purely a diffusive term, and the photon flux equation only includes the photoabsorbtions:

$$\frac{\partial F_i}{\partial t} = \dot{F}_i - F_i\, D_i, \tag{3.54}$$

where $\dot{F}_i$ is the time derivative used only in smooth RT and the destruction factor remains as in (3.53).

Equations (3.51) and (3.54) are solved numerically using a partly SIE formulation, in the sense that they are semi-implicit in the photon density and flux but otherwise explicit (in temperature and the ion abundances). A tiny bit of algebra gives:

$$\boxed{\begin{aligned} N_i^{t+\Delta t} &= \frac{N_i^t + \Delta t(\dot{N}_i + C_i)}{1 + \Delta t D_i}, \\ F_i^{t+\Delta t} &= \frac{F_i^t + \Delta t \dot{F}_i}{1 + \Delta t D_i}, \end{aligned}} \tag{3.55}$$

where all the variables at the RHS are evaluated at the beginning of the time-step, i.e. at t.

Note that it is imperative that the photon fluxes are updated completely in sync with the photon densities. Neglecting to do this, e.g. updating the flux only after all thermochemistry substeps have been completed to fill in the Δt_{RT} step, causes the RT solver to break down, with random appearances of 'light explosions'. In the case of the smooth RT solver, we need to explicitly check and enforce the constraint

$$F_i \leq c_r N_i \tag{3.56}$$

after every update via (3.55) – failing to do so can also result in a simulation crash. This is essentially because the nonzero $\dot{N}$ and $\dot{F}$ terms can lead the density and flux to be slightly out of sync. There is thus a potential danger for these light explosions to happen with smooth RT, but we have never found this to happen, as long as (3.56) is enforced – this may seem a bit reckless, but there is a clear computational gain in using smooth RT in most cases.

At the end of each photon package update, the time-stepping condition is checked: If

$$\frac{\left|N_i^{t+\Delta t} - N_i^t\right|}{N_i^t} > 0.1,$$

the `cool_step` routine returns with an un-updated state but instead a recommendation for a new time-step length $\Delta t_{new} = 0.5\ \Delta t$, so the routine can be called again with a better chance of completing. If the condition holds for all M photon packages, we are ok to continue, and the forward-in-time photon densities and fluxes can be used in the following calculations.

(b) T_μ update

Next we update the temperature by solving

$$\frac{\partial T_\mu}{\partial t} = \frac{(\gamma - 1) m_{\rm H}}{\rho k_{\rm B}}\ \Lambda, \tag{3.57}$$

where $\Lambda \equiv \dot{\varepsilon} = (\mathcal{H} + \mathcal{L} + Z/Z_\odot \mathcal{L}_Z)$, $\mathcal{L}$ is the cooling rate, $\mathcal{H}$ the photoheating rate and $\mathcal{L}_Z$ the metal-contributed cooling rate from (3.45) (here with the added factor of $n_{\rm H}^2$). These rates are calculated as follows:

The photoheating rate $\mathcal{H}$ is set by collecting the heating contributions from all photoionization events:

$$\mathcal{H} = \sum_j^{\rm HI,HeI,HeII} n_j \int_0^\infty \sigma_j(\nu) F(\nu) \left[h\nu - \epsilon_j\right] d\nu, \tag{3.58}$$

where ν is photon frequency, $F(\nu)$ local photon flux and ϵ_j photoionization energies. With the discretization into M photon packages, (3.58) becomes

$$\mathcal{H} = \sum_j^{\rm HI,HeI,HeII} n_j \sum_{i=1}^{M} \bar{\sigma}_{ij} c_r N_i \left(\bar{\epsilon}_{ij} - \epsilon_j\right) \qquad \left(+ \sum_j^{\rm HI,HeI,HeII} n_j \mathcal{H}_j^{\rm hom}\right), \tag{3.59}$$

where $\bar{\sigma}_{ij}$ and $\bar{\epsilon}_{ij}$ are the global average cross sections and photon energies, respectively, for ionization events between package i and species j (see §3.1 for their definitions and §3.6 on how they optionally relate to stellar population models). If so desired, a homogeneous UV background contribution to heating can also be included, as indicated by the term in parentheses, where $\mathcal{H}_j^{\rm hom}$ is a precomputed redshift dependent UV background heating rate corresponding to species j.

The primordial cooling rate $\mathcal{L}$ is given by

$$\begin{aligned}\mathcal{L} = & \left[\zeta_{\mathrm{HI}}(T) + \psi_{\mathrm{HI}}(T)\right]\ n_{\mathrm{e}}\ n_{\mathrm{HI}} \\ & + \zeta_{\mathrm{HeI}}(T)\ n_{\mathrm{e}}\ n_{\mathrm{HeI}} \\ & + \left[\zeta_{\mathrm{HeII}}(T) + \psi_{\mathrm{HeII}}(T) + \eta^{\mathrm{A}}_{\mathrm{HeII}}(T) + \omega_{\mathrm{HeII}}(T)\right]\ n_{\mathrm{e}}\ n_{\mathrm{HeII}} \\ & + \eta^{\mathrm{A}}_{\mathrm{HII}}(T)\ n_{\mathrm{e}}\ n_{\mathrm{HII}} \\ & + \eta^{\mathrm{A}}_{\mathrm{HeIII}}(T)\ n_{\mathrm{e}}\ n_{\mathrm{HeIII}} \\ & + \theta(T)\ n_{\mathrm{e}}\left(n_{\mathrm{HII}} + n_{\mathrm{HeII}} + 4n_{\mathrm{HeIII}}\right) \\ & + \varpi(T)\ n_{\mathrm{e}},\end{aligned} \tag{3.60}$$

where the various cooling processes are collisional ionizations ζ, collisional excitations ψ, recombinations η, dielectronic recombinations ω, bremsstrahlung θ and Compton cooling ϖ, all analytic (fitted) functions of temperature taken from various sources. The complete expressions are listed (with references) in Appendix C. If the OTSA us used, the η^{A} coefficients are replaced with η^{B}.

The metal contribution $\mathcal{L}_Z$ is unchanged at this point from the homogeneous UV PIE modelling in `Ramses`. The metal contribution $\mathcal{L}_Z$ is thus interpolated from a (T_μ, n_{H})-table which is pre-calculated every coarse time-step with Eq. 3.45.

The temperature update (3.57) is solved numerically using semi-implicit formulation in T_μ, using FW values of photon densities and BW values of H and He species abundances, since these have not been updated yet. The temperature is updated to

$$\boxed{T_\mu^{t+\Delta t} = T_\mu^t + \frac{\Lambda K \Delta t}{1 - \Lambda' K \Delta t'}} \tag{3.61}$$

where $K \equiv \frac{(\gamma-1)m_{\mathrm{H}}}{\rho k_{\mathrm{B}}}$. The temperature derivative,

$$\Lambda' \equiv \frac{\partial \mathcal{L}}{\partial T_\mu} + \frac{Z}{Z_\odot} n_{\mathrm{H}}^2 \frac{\partial \mathcal{L}_Z}{\partial T_\mu},$$

is found by algebraically differentiating each of the primordial cooling rate expressions in the case of $\mathcal{L}$ (and using $\frac{\partial \mathcal{L}}{\partial T_\mu} = \mu \frac{\partial \mathcal{L}}{\partial T}$), and finite-differencing the metal-cooling table in the case of $\mathcal{L}_Z$. The temperature derivative of the heating rate is zero.

With $T_\mu^{t+\Delta t}$ in hand, the time-stepping condition is checked. If

$$\frac{\left|T_\mu^{t+\Delta t} - T_\mu^t\right|}{T_\mu^t} > 0.1,$$

the `cool_step` routine returns with an un-updated $\mathcal{U}$ and a recommendation for halving the time-step length. In tests we've found that the usual time-step constraint given here is not enough to ensure stability, as the temperature in some cases oscillates, occasionally even in a divergent way. Λ and Λ' are both evaluated backwards in time, i.e. at t, and the large difference that can exist in these values from t to $t + \Delta t$ appears to cause these instabilities.

To fix that we include also a first-order time-step constraint on the temperature, meaning we return with an un-updated state if

$$\frac{|K\Lambda\Delta t|}{T^t_\mu} > 0.1. \tag{3.62}$$

With this fix, we have not seen further temperature oscillations, but there is no guarantee that numerical instabilities are eliminated.

(c) Hydrogen ionized fraction update

The HII abundance is controlled by collisional ionizations, photoionizations, and recombinations, i.e.

$$\frac{\partial n_{\rm HII}}{\partial t} \;=\; n_{\rm HI}\,\beta_{\rm HI}(T)n_{\rm e} \;+\; n_{\rm HI}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HI}c_r N_i \quad \left(+n_{\rm HI}\Gamma^{\rm hom}_{\rm HI}\right) \;-\; n_{\rm HII}\,\alpha^{\rm A}_{\rm HII}(T)n_{\rm e}, \tag{3.63}$$

where $\beta_{\rm HI}$ is the rate of collisional ionizations by electrons (see Appendix C) and $\alpha^{\rm A}_{\rm HII}$ the case A hydrogen recombination rate, which is replaced here by $\alpha^{\rm B}_{\rm HII}$ if the OTSA is used. The term in parentheses represents an optional homogeneous UV background, with $\Gamma^{\rm hom}_{\rm HI}$ a precomputed redshift dependent HI photoionization rate. In terms of ionization fraction, (3.63) becomes

$$\begin{aligned}\frac{\partial x_{\rm HII}}{\partial t} &= (1-x_{\rm HII})\left[\beta_{\rm HI}(T)n_{\rm e} \;+\; \sum_{i=1}^{M}\bar{\sigma}_{i\rm HI}c_r N_i \quad \left(+\Gamma^{\rm hom}_{\rm HI}\right)\right] \;-\; x_{\rm HII}\,\alpha^{\rm A}_{\rm HII}(T)n_{\rm e} \\ &= (1-x_{\rm HII})\,C - x_{\rm HII}\,D \\ &= C - x_{\rm HII}\,(C+D),\end{aligned} \tag{3.64}$$

where we have in the second line separated the rates into HII creation C and destruction D, and in the third line collected multiples of $x_{\rm HII}$.

Anninos et al. (1997) recommend solving the HII update partly-implicitly, with something like

$$x^{t+\Delta t}_{\rm HII} = \frac{x^t_{\rm HII} + C\Delta t}{1+(C+D)\Delta t}, \tag{3.65}$$

where the updated, $t+\Delta t$ values, of temperature and photon densities are used. We've tried this approach initially but found that it can in some cases be unstable, leading to diverging oscillations in $\mathcal{U}$. The problem here is that the temperature varies substantially with the variation in $x_{\rm HII}$ through μ, and hence also the temperature dependent rates of recombinations and collisional ionizations, causing a particular stiffness in this equation. We have therefore gone for an approach which is semi-implicit in $x_{\rm HII}$:

$$\boxed{x^{t+\Delta t}_{\rm HII} = x^t_{\rm HII} + \Delta t\frac{C - x^t_{\rm HII}(C+D)}{1-J\Delta t},} \tag{3.66}$$

where

$$J \equiv \frac{\partial \dot{x}_{\rm HII}}{\partial x_{\rm HII}} = \frac{\partial C}{\partial x_{\rm HII}} - (C+D) - x_{\rm HII}\left(\frac{\partial C}{\partial x_{\rm HII}} + \frac{\partial D}{\partial x_{\rm HII}}\right), \tag{3.67}$$

and the creation and destruction derivatives are given by

$$\frac{\partial C}{\partial x_{\rm HII}} = n_{\rm H}\beta_{\rm HI} - n_{\rm e}T_\mu\mu^2 X\frac{\partial \beta_{\rm HI}}{\partial T} \tag{3.68}$$

$$\frac{\partial D}{\partial x_{\rm HII}} = n_{\rm H}\alpha^{\rm A}_{\rm HII} - n_{\rm e}T_\mu\mu^2 X\frac{\partial \alpha^{\rm A}_{\rm HII}}{\partial T}. \tag{3.69}$$

With this approach, $x_{\rm HII}$ instabilities no longer appear in our tests.

We end with the usual check if the 10% rule is broken. If

$$\frac{\left|x^{t+\Delta t}_{\rm HII} - x^t_{\rm HII}\right|}{x^t_{\rm HII}} > 0.1,$$

`cool_step` returns with an un-updated $\mathcal{U}$ and a recommendation for halving the time-step length. Like with the temperature a first-order check is included as well, i.e. if

$$\frac{\left|C - x^t_{\rm HII}(C+D)\right|}{x^t_{\rm HII}}\Delta t > 0.1.$$

(d) Helium ionized fractions update

The HeI fraction is not a variable that is stored in `RamsesRT`, but we evolve it nontheless, starting with $x^t_{\rm HeI} = 1 - x^t_{\rm HeII} - x^t_{\rm HeIII}$, in order to make a consistency check at the end of the helium updates. Before each of these updates, we recalculate $n_{\rm e}$ and μ to reflect the new FW abundances. The HeI fraction is controlled by

$$\begin{aligned}\frac{\partial x_{\rm HeI}}{\partial t} &= x_{\rm HeII}\,\alpha^{\rm A}_{\rm HeII}(T)n_{\rm e} - x_{\rm HeI}\,\beta_{\rm HeI}(T)n_{\rm e} - x_{\rm HeI}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeI}c_r N_i \quad \left(-x_{\rm HeI}\,\Gamma^{\rm hom}_{\rm HeI}\right)\\ &= C - x_{\rm HeI}\,D,\end{aligned} \tag{3.70}$$

i.e. HeII recombinations and collisional- and photo-ionizations of HeI. As usual $\alpha^{\rm A}$ is replaced by $\alpha^{\rm B}$ in the case of the OTSA, and the term in parentheses represents an optional homogeneous UV background. In the second line of (3.70) we've separated the RHS into HeI creation C and destruction D.

Here we follow Anninos et al. (1997) and do the HeI update with

$$\boxed{x^{t+\Delta t}_{\rm HeI} = \frac{x^t_{\rm HeI} + C\Delta t}{1 + D\Delta t},} \tag{3.71}$$

which is derived by using $x^{t+\Delta t}_{\rm HeI}$ on the RHS of the discretized version of (3.70), and solving algebraically for $x^{t+\Delta t}_{\rm HeI}$. The update is partly implicit in the other variables since it uses

updated values of $N_i^{t+\Delta t}$, $T_\mu^{t+\Delta t}$, and $x_{\rm HII}^{t+\Delta t}$ ($\rightarrow \mu$ and $n_{\rm e}$), but un-updated values of $x_{\rm HeII}^t$ and $x_{\rm HeIII}^t$. No tests have revealed any instabilities here, so it appears stiffness is less of a problem than with temperature and $x_{\rm HII}$.

We then evolve the HeII fraction. The differential equation to solve is

$$\begin{aligned}\frac{\partial x_{\rm HeII}}{\partial t} &= x_{\rm HeI}\,\beta_{\rm HeI}(T)n_{\rm e} + x_{\rm HeIII}\,\alpha^{\rm A}_{\rm HeIII}(T)n_{\rm e} + x_{\rm HeI}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeI}c_r N_i \quad \left(+x_{\rm HeI}\Gamma^{\rm hom}_{\rm HeI}\right)\\ &- x_{\rm HeII}\,\beta_{\rm HeII}(T)n_{\rm e} - x_{\rm HeII}\,\alpha^{\rm A}_{\rm HeII}(T)n_{\rm e} - x_{\rm HeII}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeII}c_r N_i \quad \left(-x_{\rm HeII}\Gamma^{\rm hom}_{\rm HeII}\right)\\ &= C - x_{\rm HeII}\,D. \end{aligned} \tag{3.72}$$

The RHS terms are, in order of appearance, HeI collisional ionizations, HeIII recombinations, HeI photoionizations (with an optional homogeneous background in parentheses), HeII collisional ionizations, HeII recombinations and HeII photoionizations (UV background optional). In the third line we have grouped the terms into a creation term C (first RHS line) and destruction terms D (second RHS line).

The discrete update is done with the same formulation as (3.71), i.e.

$$\boxed{x_{\rm HeII}^{t+\Delta t} = \frac{x_{\rm HeII}^t + C\Delta t}{1 + D\Delta t},} \tag{3.73}$$

using updated values of $N_i^{t+\Delta t}$, $T_\mu^{t+\Delta t}$, $x_{\rm HII}^{t+\Delta t}$, and $x_{\rm HeI}^{t+\Delta t}$ ($\rightarrow \mu$ and $n_{\rm e}$), and the un-updated value only of $x_{\rm HeIII}^t$.

The only variable left to update is the HeIII fraction. The differential equation is

$$\begin{aligned}\frac{\partial x_{\rm HeIII}}{\partial t} &= x_{\rm HeII}\,\beta_{\rm HeII}(T)n_{\rm e} + x_{\rm HeII}\sum_{i=1}^{M}\bar{\sigma}_{i\rm HeII}c_r N_i \quad \left(+x_{\rm HeII}\Gamma^{\rm hom}_{\rm HeII}\right)\\ &- x_{\rm HeIII}\,\alpha^{\rm A}_{\rm HeIII}(T)n_{\rm e}\\ &= C - x_{\rm HeIII}\,D. \end{aligned} \tag{3.74}$$

The RHS terms are, in order of appearance, HeII collisional ionizations, HeII photoionizations (with an optional homogeneous background in parentheses), and HeIII recombinations. In the third line we have as usual grouped the terms into creation and destruction.

Again the update follows the same formulation,

$$\boxed{x_{\rm HeIII}^{t+\Delta t} = \frac{x_{\rm HeIII}^t + C\Delta t}{1 + D\Delta t},} \tag{3.75}$$

which is implicit in the sense that $x_{\rm HeII}^t$ is used on the RHS and $t + \Delta t$ values of all the other variables.

Conservation of helium density is then enforced, i.e. that

$$x_{\rm HeI} + x_{\rm HeII} + x_{\rm HeIII} = 1, \tag{3.76}$$

by lowering the largest of these fractions accordingly (in the case of x_{HeI} being the largest there is no update, since it only exists in terms of the other two fractions).

The 10% rule is not applied to the helium fractions. Instead, the final 10% check is done on the electron density[5], which is retrieved from all the ionization fractions with

$$n_{\mathrm{e}} = x_{\mathrm{HII}} n_{\mathrm{H}} + (x_{\mathrm{HeII}} + 2x_{\mathrm{HeIII}}) n_{\mathrm{He}}. \tag{3.77}$$

If

$$\frac{\left| n_{\mathrm{e}}^{t+\Delta t} - n_{\mathrm{e}}^{t} \right|}{n_{\mathrm{e}}^{t}} > 0.1,$$

`cool_step` returns with an un-updated $\mathcal{U}$ and a recommendation for halving the time-step length.

(e) Time-step check

All the variables have been updated, from $\mathcal{U}^t$ to $\mathcal{U}^{t+\Delta t}$, and the 10% rule is not violated over the time-step. The time-step length just used may have been unneccessarily short, and if so, there is a large probability that it is also unneccessarily long for the call to `cool_step`, to fill the total Δt_{RT}.

Therefore, `cool_step`$(\mathcal{U}, \mathtt{\Delta t})$ makes a final time-step check before finishing up, of how close it was to breaking the 10% rule. If the maximally changed variable in $\mathcal{U}$ has changed by less than 5%, i.e. if

$$\max\left(\left| \frac{\mathcal{U}^{\mathrm{t}+\Delta \mathrm{t}} - \mathcal{U}^{\mathrm{t}}}{\mathcal{U}^{\mathrm{t}}} \right| \right) < 0.05, \tag{3.78}$$

then the routine returns with a recommendation for a time-step length Δt that is twice the one just used. Note that the recommendation is for the current cell only, i.e. it is made on a cell-by-cell basis, and the recommended time-step for each cell is only stored during the subcycling and lost at the end of each Δt_{RT} cycle. At the beginning of each cell-cycle the first guess at a time-step is always Δt_{RT}. If this is too large for the 10% rule to be obeyed, successive calls to `cool_step` will quickly fix that by halving the time-step until the rule is no longer broken, and only then will `cool_step` start to return updated values of $\mathcal{U}$.

3.5.3 Thermochemistry performance and prospectives

The non-equilibrium thermochemistry of `RamsesRT` has been thoroughly checked with tests described in the next chapter. These include one-cell tests that check for instabilities and accuracy of the numerical methods over a range of densities, temperatures and ionization states, and also the benchmark radiative transfer tests which allow direct comparison of temperatures and ionization states with results from other codes, though unfortunately these tests include hydrogen only.

The non-equilibrium thermochemistry of `RamsesRT` is slower than that of equilibrium `Ramses` – understandably so since there is only one variable to evolve in `Ramses`, i.e. the

[5] This is just following Anninos et al. (1997), but since the electron density predominantly follows the hydrogen ionized fraction, which has already been checked, it would be more sensible replace this in the future with a direct 10% rule on the helium fractions.

temperature, and many, typically 10, in RamsesRT, which makes for an increase in the number of operations to perform in each step. We have not checked quantitatively the difference in speed between the two, but using RamsesRT, even with no photon packages, will slow down the thermochemistry by a factor of $\sim 2-3$ compared to equilibrium Ramses.

One way of speeding up the RamsesRT thermochemistry is to tabulate the cooling and reaction rate coefficients, which are given in Appendix C. These are all given by exponential expressions, which are expensive in terms of computing time. Tabulating these into (T, ρ)-bins at RamsesRT startup and interpolating values from the tables when needed should give some speedup, with a negligible sacrifice in accuracy.

Metal cooling should also at some point be updated to reflect the inhomogeneous UV radiation – which in reality has a large impact on metal cooling by determining which metal ions, and thus which metal cooling lines, are dominant.

3.6 Stellar UV emission and photon package properties

In the photon injection step in RamsesRT (§3.2.2), the task is to inject photons into each grid cell corresponding to the stellar particles that reside in it. Here we describe how we derive these injection rates from stellar energy distribution (SED) models, along with the cross sections and photon energies for each package.

The rate of photon injection from stellar particles into a cell over a time-step Δt is

$$\dot{N}_i^\star = f_{\rm esc} \sum_\star^{\rm cell\ stars} \frac{m_\star}{\Delta t} \left[\Pi_i({\tau_\star}^{t+\Delta t}, Z_\star) - \Pi_i({\tau_\star}^{t}, Z_\star) \right], \tag{3.79}$$

where $f_{\rm esc}$ is an escape fraction, $m_\star$, $\tau_\star$ and $Z_\star$ are mass, age and metallicity of the stellar particles, respectively, and Π_i is some model for the accumulated number of package i photons emitted per solar mass over the lifetime (so far) of a stellar particle. This goes into (3.22) in the photon injection step.

Stellar particles in Ramses represent stellar populations, so it makes sense to use SED models to infer their luminosities. We have implemented functions in RamsesRT that read SED tables at startup and derive from them stellar UV luminosities for photon injection, as well as photon package properties that can optionally be updated to reflect the average emission from the stellar particles populating the simulation.

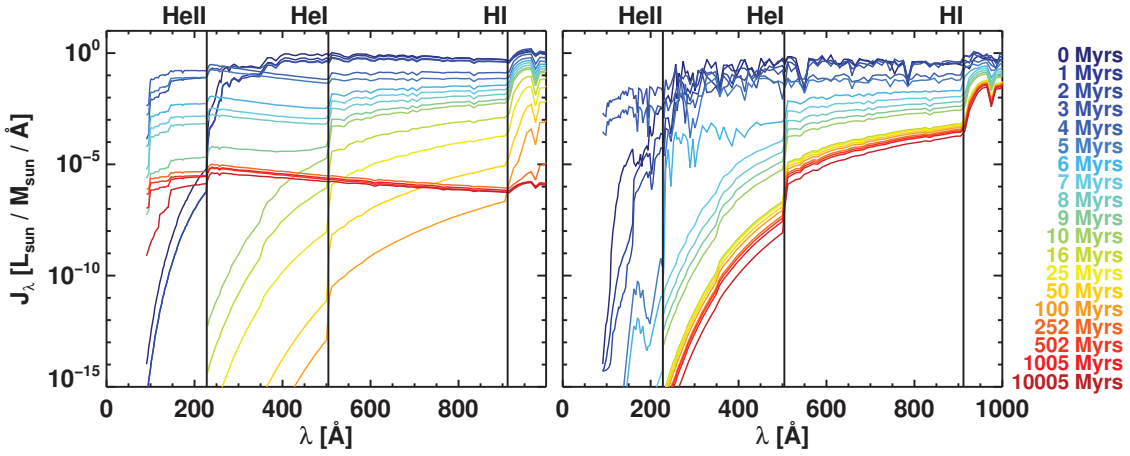
We have hitherto used the SED model of Bruzual and Charlot (2003) (BC03), but it can be replaced with any other model, e.g. Starburst99 (Leitherer et al., 1999), as long as the file format is the same, so RamsesRT can read the tables. The model should come in the form of spectra, $J_\lambda(\tau, Z)$, binned by stellar population age and metallicity. In Fig. 3.9 we show J_λ from BC03 and Starburst99 at solar metallicity for various population ages.

The per-package ionizing luminosities of stellar particles are derived from the SED model by

$$J_i(\tau, Z) = \int_{\nu_{i0}}^{\nu_{i1}} J_\nu(\tau, Z)\, d\nu, \tag{3.80}$$

where $J_\nu = c/\nu^2 J_{\lambda(\nu)}$, and the cumulative luminosities are calculated with

$$\Pi_i(\tau, Z) = \int_0^\tau J_i(t, Z)\, dt. \tag{3.81}$$



(a) BC03 *(b) Starburst99*

Figure 3.9: *SED plots from (**a**) BC03 and (**b**) Starburst99 for solar metallicity at different stellar population ages. The spectral luminosity is given in solar luminosities ($3.8\,10^{33}$ erg s^{-1}) per solar mass ($2\,10^{30}$ g) per wavelength. Vertical lines mark the ionization wavelengths for* HI, HeI *and* HeII*, which correspond to the wavelengths marking the three photon packages we typically use in our simulations. The Starburst99 spectra are generated with the instantaneous formation of 10^6 solar masses and a Salpeter initial mass function.*

Since both the photon injection and the calculation of package properties are done on the fly, $J_i(\tau, Z)$ and $\Pi_i(\tau, Z)$ must be evaluated as quickly as possible for given stellar ages and metallicities. Values of J_i and Π_i are therefore only calculated from the SED spectra via (3.80) and (3.81) at simulation startup, and tabulated with equally-spaced logarithmic bins of age and metallicity, so that they can be retrieved with minimum computational effort via linear interpolation.

Package properties

There are two sets of global properties for each photon package, which reflect *average cross sections* and *average photoionization event energies*, that were defined in §3.1. For an age and metallicity dependent reference spectrum $J_\nu(\tau, Z)$, these are, respectively

$$\bar{\sigma}_{ij}(\tau, Z) = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J_\nu \, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} J_\nu \, d\nu}, \tag{3.82}$$

$$\bar{\epsilon}_{ij}(\tau, Z) = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J_\nu h\nu \, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_{\nu j} J_\nu \, d\nu}. \tag{3.83}$$

Since there are three ionizeable species in the current implementation of `RamsesRT`, each photon package has three average energies and three average cross sections associated with it. These properties can be set as run parameters to reflect some typical stellar spectra, e.g. a blackbody or a SED. It can also be left to `RamsesRT` to set them on the fly to reflect the stellar populations in the simulation, using (3.83) and (3.82), with the loaded SED spectra representing J_ν and the expressions from Verner et al. (1996) for $\sigma_{\nu j}$ (see Appendix C). Due

to the averaged nature of the photon packages, we must however suffice to set the package properties to reflect the *average* stellar emission in the simulation. If this option is used, the photon package properties are updated every n coarse time-steps (where n is an adjustable parameter) by polling all the stellar particles in the simulation and setting for each package i and species j,

$$\bar{\sigma}_{ij} = \frac{\sum_{\star}^{\text{all stars}} \bar{\sigma}_{ij}(\tau_\star, Z_\star)\ m_\star\ J_i(\tau_\star, Z_\star)}{\sum_{\star}^{\text{all stars}} m_\star\ J_i(\tau_\star, Z_\star)}, \tag{3.84}$$

$$\bar{\epsilon}_{ij} = \frac{\sum_{\star}^{\text{all stars}} \bar{\epsilon}_{ij}(\tau_\star, Z_\star)\ m_\star\ J_i(\tau_\star, Z_\star)}{\sum_{\star}^{\text{all stars}} m_\star\ J_i(\tau_\star, Z_\star)}. \tag{3.85}$$

The values of each star particle's $J_i(\tau_\star, Z_\star)$, $\bar{\sigma}_{ij}(\tau_\star, Z_\star)$ and $\bar{\epsilon}_{ij}(\tau_\star, Z_\star)$ are interpolated from tables that are generated at startup via (3.80), (3.82) and (3.83). For quick interpolation these tables are binned with equally spaced logarithmic bins of age and metallicity. If no stellar particles exist yet in the simulation, which is normally the case at startup, the average energies and cross sections are set corresponding to the minimum tabulated age and metallicity, which should accurately reflect the first-born stellar particles when they appear.

Note that the on-the-fly update of photon properties according to (3.84) and (3.85) infers that existing photon's properties are changed, i.e. the properties of photons that have already been emitted change in mid-air. We must live with this: In the limit of the moment method, all photons sharing the same frequency interval are 'the same' and there is no trivial way of differentiating between them. In the limit that emitted photons do not travel over cosmological volumes before being absorbed, this should not however be such a bad approximation.

Although one can in principle use as many photon packages as one wants, it is practical to only use a handful, due to limitations in memory and computation. We typically use three photon packages in our simulations, representing HI, HeI and HeII ionizing photons, as indicated by vertical lines in the plots of Fig. 3.9. The stellar luminosities, instantaneous and accumulated, average cross sections and energies for these packages are plotted in Fig. 3.10 for BC03 (top) and Starburst99 (bottom), as calculated via (3.80)-(3.83). From the luminosity plots (top rows), it can be seen that the stellar populations emit predominantly for the first $\sim 3-6$ Myrs and the luminosity drastically goes down as the most massive stars begin to expire. During the bright early phase, roughly a third of the ionizing photons are HeI-ionizing and only a very small fraction of them are HeII-ionizing. At ~ 4 Myrs there is a sudden and short lived peak in the production of HeII-ionizing photons, which is due to Wolf-Rayet stars. In the BC03 model there is a stable but declining production of HeII-ionizing photons starting at ~ 250 Myrs, which is due to an accumulation of low-mass post-asymptotic giant branch stars (this feature does not appear in Starburst99).

As the first few Myrs dominate the photon production, the photon properties are actually always dominated by these young stellar populations, and stay pretty stable. (The HeII-ionizing photons are an exception to this, as both their properties and luminosity changes

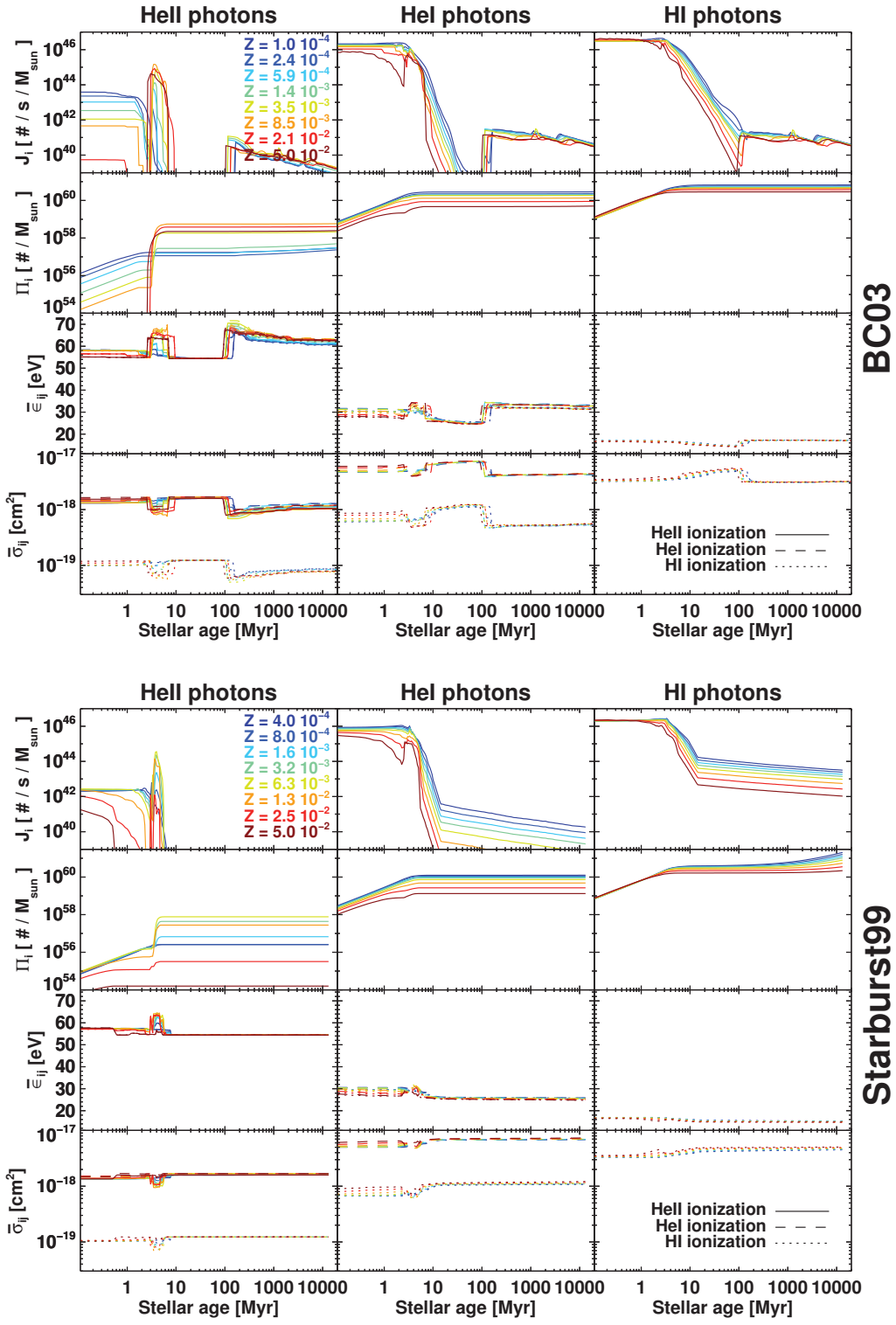


Figure 3.10: *HeII, HeI, and HI ionizing photon package properties derived from the BC03 (**top panel**) and Starburst99 SED models (**bottom panel**), as functions of age (x-axis) and metallicity (colors). The plot columns represent the three photon packages. **Top rows** show stellar luminosity, in the number of photons that goes into each package per second per solar mass. **Second rows** show accumulated number of photons emitted. **Third rows** show the average photon energies per interaction. **Bottom rows** show average cross sections per interaction.*

quite a lot both in time and metallicity, and to some extent also the HeI-ionizing photons at high metallicities.) One could thus arguably do without on-the-fly use of SED models and instead just use constant luminosities during 6 Myrs, after which they're shut off, and package properties derived from newborn stellar populations. We have tabulated those values in Table 3.1 for the two SED models. For both models, most of the photons emitted in the early phase of the stellar population lifetime are accounted for in the first 6 Myr; counting up the total emitted photons at 10 Myr only adds $\lesssim 10\%$ over the 6 Myr mark. The number of photons which is slowly accumulated over a Hubble time is usually also $\lesssim 10\%$ of what comes out during the early phase, though it can in some cases be more considerable, even up to an order of magnitude larger, as can be seen for metal-poor HeII photons in the BC03 model and metal-poor HI photons in the Starburst99 model.

Photon	J_i $[s^{-1}M_\odot^{-1}]$		$\bar{\epsilon}_{ij}$ [eV]			$\bar{\sigma}_{ij}$ [cm^2]		
package	Low Z	High Z	HeII	HeI	HI	HeII	HeI	HI
BC03 model								
HeII	$8.0\ 10^{42}$	$1.0\ 10^{44}$	58.3	58.4	58.2	$1.3\ 10^{-18}$	$1.4\ 10^{-18}$	$9.9\ 10^{-20}$
HeI	$1.5\ 10^{46}$	$2.0\ 10^{45}$	-	31.7	30.5	-	$4.7\ 10^{-18}$	$6.0\ 10^{-19}$
HI	$3.0\ 10^{46}$	$1.5\ 10^{46}$	-	-	17.0	-	-	$3.2\ 10^{-18}$
Starburst99 model								
HeII	$1.0\ 10^{42}$	$8.0\ 10^{39}$	57.3	57.5	57.4	$1.4\ 10^{-18}$	$1.5\ 10^{-18}$	$1.0\ 10^{-19}$
HeI	$6.5\ 10^{46}$	$7.0\ 10^{44}$	-	30.5	29.6	-	$5.0\ 10^{-18}$	$6.7\ 10^{-19}$
HI	$1.8\ 10^{46}$	$8.5\ 10^{45}$	-	-	16.9	-	-	$3.3\ 10^{-18}$

Table 3.1: *Typical UV luminosities (based on the first 6 Myrs) and photon package properties (based on zero-age, low-metallicity). Constant values in this ballpark may be used instead of on-the-fly evaluation from SED tables, with the stellar particles simply emitting at constant luminosities for the first 6 Myrs of their lifetime and then being shut off.*

4
Code tests

4.1 Introduction

Testing is a vital part of code development, both for ensuring that the code is free of serious bugs and for convincing potential users and competitors alike that it is safe to use and produces robust results. We have run two main sets of tests, which we describe in this chapter. Thermochemistry equations are notoriously prone to instabilities in numerical implementations, so we have run tests, presented in the first part of this chapter, that carefully check for stable convergence towards known results. The second part is then dedicated to the radiative transfer as a whole, where we run the tests from the so-called 'Cosmological radiative transfer codes comparison project', which are presented in two papers, Iliev et al. (2006a) (Il06) and Iliev et al. (2009) (Il09), and have become accepted today as the benchmark tests for ionizing RT implementations.

4.2 Thermochemistry tests

Here we validate the non-equilibrium thermochemistry that we've implemented for `RamsesRT` (§3.5.2). To this purpose, we've run *one-cell* thermochemistry tests, where we start at some initial state (temperature, ionization state, photon flux) and evolve over roughly a Hubble time. We are interested here in verifying that our implementation is correct and error free and also in comparing equilibrium vs. non-equilibrium cooling – e.g. Cen and Fang (2006) report that the methods can produce significantly different results. We thus compare against the original equilibrium thermochemistry of `Ramses` which has been modified such that it uses the exact same heating, cooling and interaction rates as `RamsesRT`.

We test to see *(i)* whether the thermochemistry of `RamsesRT` is stable, i.e. if the stiffness of the equations results in any sudden divergence or 'wiggles' in the evolution of the gas, *(ii)* whether `RamsesRT` evolves the ionization fractions towards the correct states predicted

by the equilibrium solver of Ramses, and *(iii)* whether the Ramses and RamsesRT evolve the temperature towards the same final value.

There are four tests: First we disable cooling and evolve only the ionization states of hydrogen and helium at different constant temperatures in a zero UV radiation field, and see if we reach equilibrium ionization states (predicted by Ramses). Then we turn on a constant UV radiation field and again see if we reach equilibrium states. Then we turn on cooling, and for two sets (zero, nonzero radiation field) see if the temperature evolution is comparable to Ramses equilibrium cooling from the same initial conditions.

4.2.1 Ionization convergence at constant temperature and zero ionizing photon flux

In the first test, cooling is turned off and we check for a range of densities, temperatures and initial ionization states whether we get a convergence of the ionized fractions towards their equilibrium states, as predicted by Ramses, assuming zero flux of ionizing photons.

Fig. 4.1 shows the results. Each panel of 3×6 plots in the figure represents an evolution given the constant temperature written to the right of the panel, and shows how the ionized fractions, x_{HII}, x_{HeII} and x_{HeIII}, evolve from different (color-coded) starting states $x_i = x_{\mathrm{HII}} = x_{\mathrm{HeIII}}$ (the HeII fraction always starts at zero). A black dashed line in each plot shows the equilibrium ionization fraction for the given temperature and species (which is gas density independent in the case of zero ionizing flux). Each column of plots represents a (non-evolving) hydrogen number density.

The non-equilibrium ionization fractions always evolve towards the equilibrium ones, at a rate which depends on gas density, as expected. It can even take longer than the age of the Universe to reach equilibrium for the most diffuse gas ($n_{\mathrm{H}} \lesssim 10^{-6}$ cm^{-3}), which indeed is a significant difference from the equilibrium assumption! If we zoom in around the equilibrium states we find a difference between the calculated equilibrium state and the evolved one which is typically around one in ten-thousand - this simply corresponds to the allowed error in the iterative equilibrium calculation, and can be decreased at will by reducing this error margin.

4.2.2 Ionization convergence at constant temperature and nonzero ionizing photon flux

This is the same as the previous test, except now we apply a constant flux of 10^5 ionizing photons s^{-1} cm^{-2} through the cell, assuming the spectrum of a blackbody at 10^5 K.

Fig. 4.2 shows the results. The black dashed lines in each plot show the equilibrium state which now is density dependent - the denser the gas the harder it is for the radiation field to battle against recombinations. Again the non-equilibrium ionized state always evolves towards the equilbrium one at a gas density dependent rate, though note that here it takes a maximum of ~ 10 Myr, which is much shorter than it can take in the zero photon flux case.

4.2.3 Temperature convergence with zero ionizing photon flux

Now cooling is turned on, and we compare the RamsesRT non-equilibrium temperature evolution with that of equilibrium Ramses (though keep in mind it has been adjusted to contain the exact same cooling rates as RamsesRT). Each of the 5 rows in fig. 4.3 shows

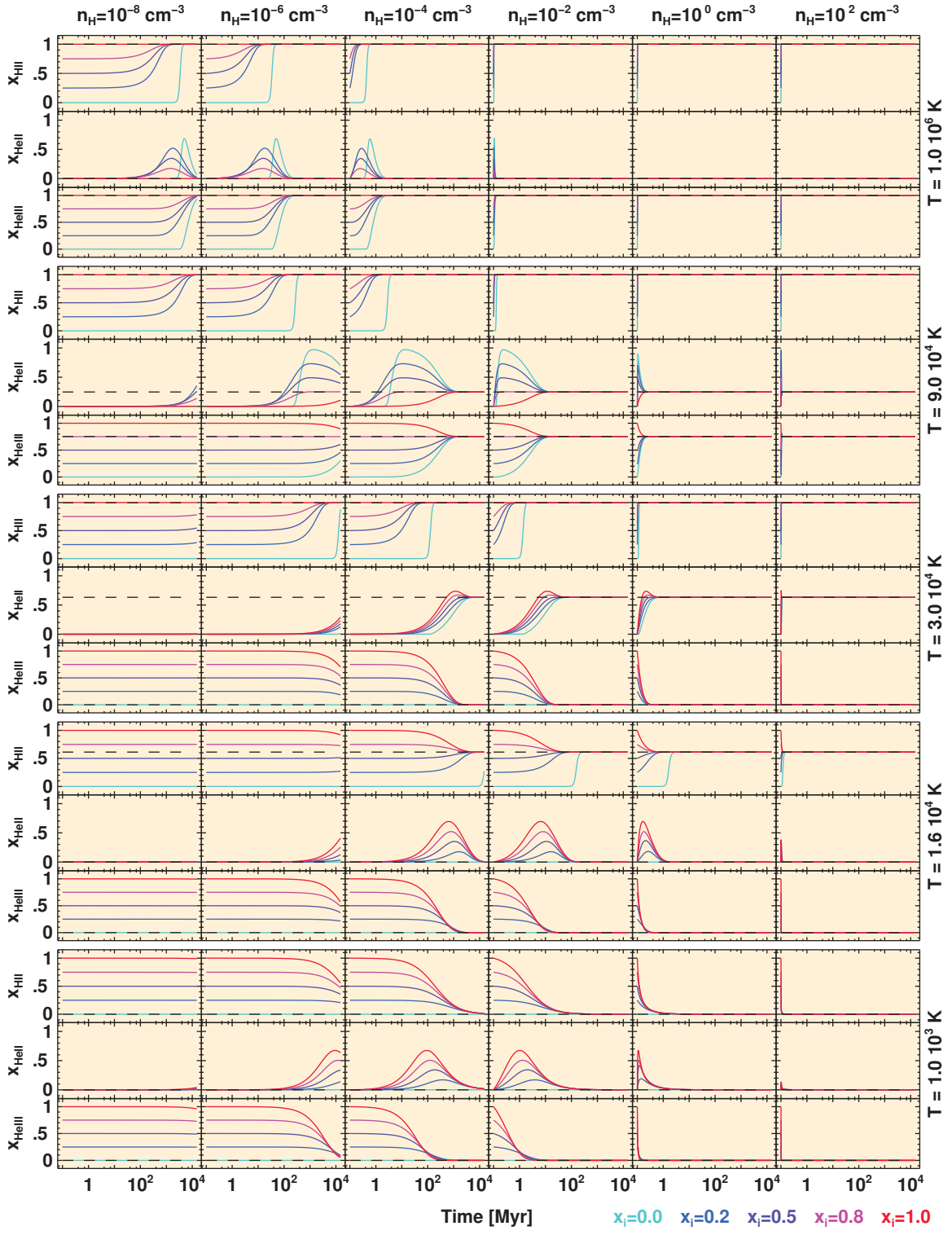


Figure 4.1: *Ionization convergence test with constant T and zero ionizing photon flux. Coloured lines show non-equilibrium evolution of the ionization fractions, given constant T (right) and n_{H} (top). Black dashed lines show the corresponding equilibrium ionization fractions as calculated in* `Ramses`.

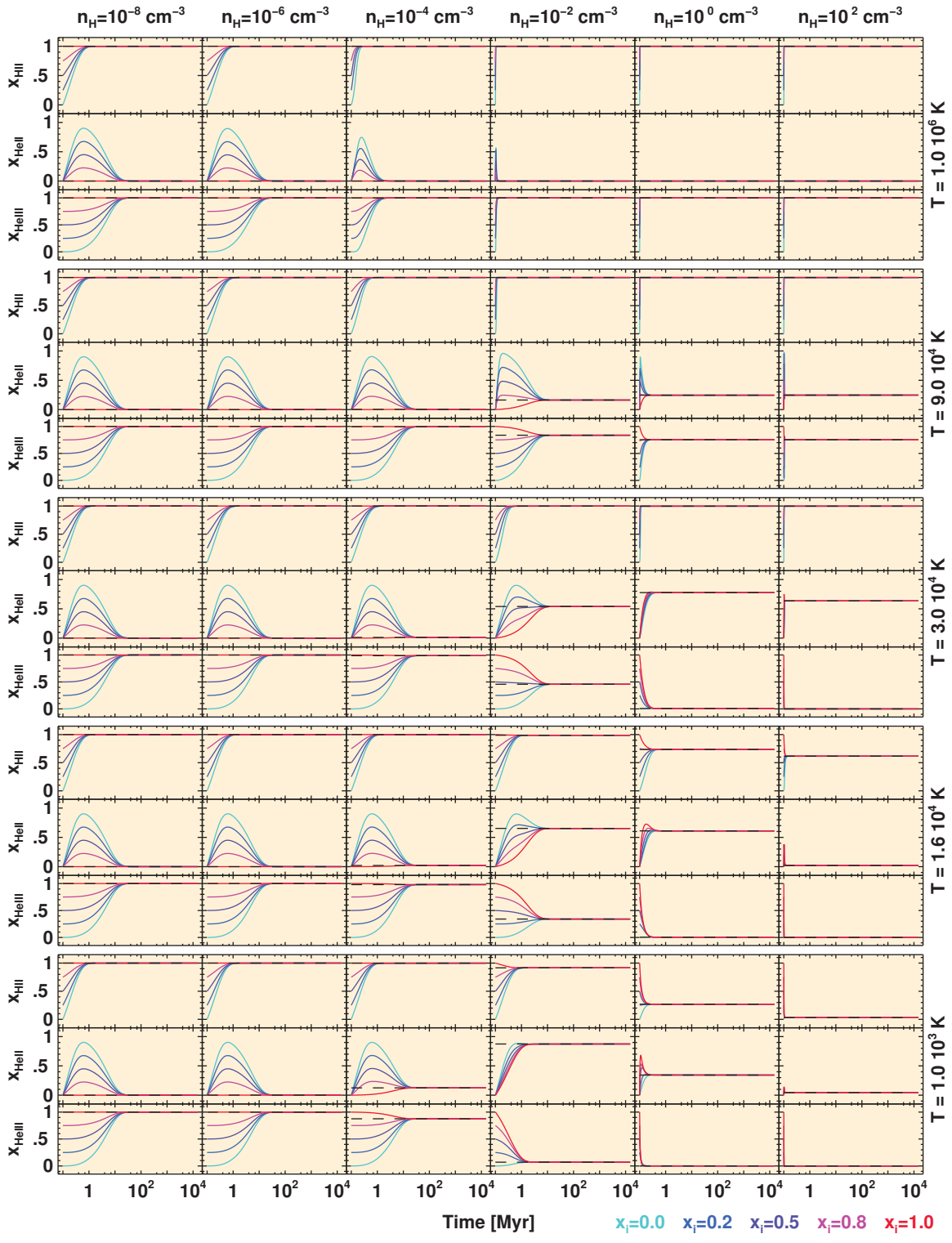


Figure 4.2: *Ionization convergence test with constant temperature and an ionizing photon flux of* $10^5\ \mathrm{s}^{-1}\,\mathrm{cm}^{-2}$.

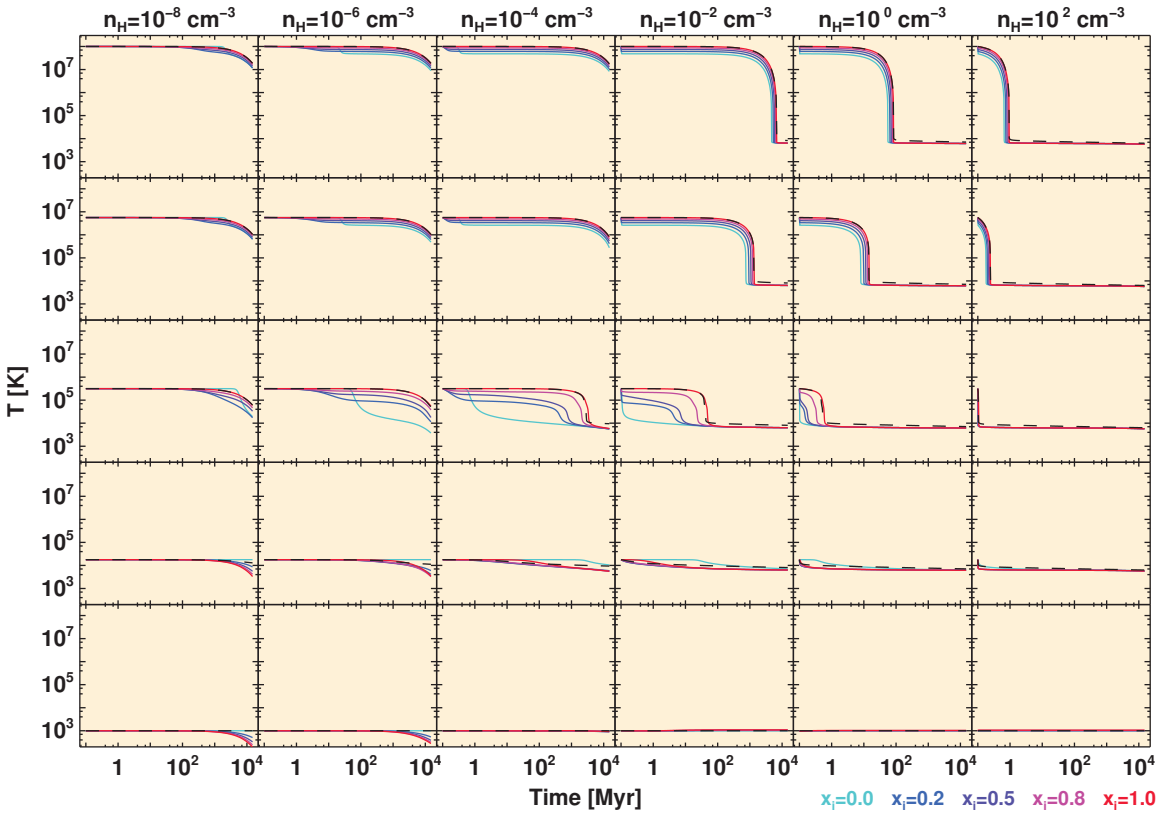


Figure 4.3: *Temperature convergence with zero ionizing flux. Color coded lines show different initial states of x_{HII} and x_{HeIII}, as indicated by the color legend at bottom right. Black dashed curves show the equilibrium evolution from* `Ramses`.

cooling for a range of decreasing initial temperatures, from top to bottom. The color-codings (initial ionization states) and columns (hydrogen number densities) are the same as before. The solid colored lines show non-equilibrium cooling in `RamsesRT` and the black dashed lines represent equilibrium cooling in `Ramses` starting from the same temperature.

Clearly the temperature evolution is quite similar between equilibrium/non-equilibrium cooling, especially if the initial ionization fraction is 'correct', i.e. if it matches the equilibrium one at the initial temperature.

The final temperature reached in the non-equilibrium case is usually a bit lower than in the equilibrium case. This is independent of gas density and initial temperature (as long as the initial temperature allows for cooling to happen). The reason for this is that the non-equilibrium ionization evolution lags behind the instantaneous equilibrium one, so there is always a somewhat larger reservoir of electrons in the non-equilibrium case. Electrons are the primary cooling agents, and complete electron depletion completely stops cooling, so it makes sense that if the electrons deplete more slowly, cooling is more effective and can bring the gas to a lower final temperature.

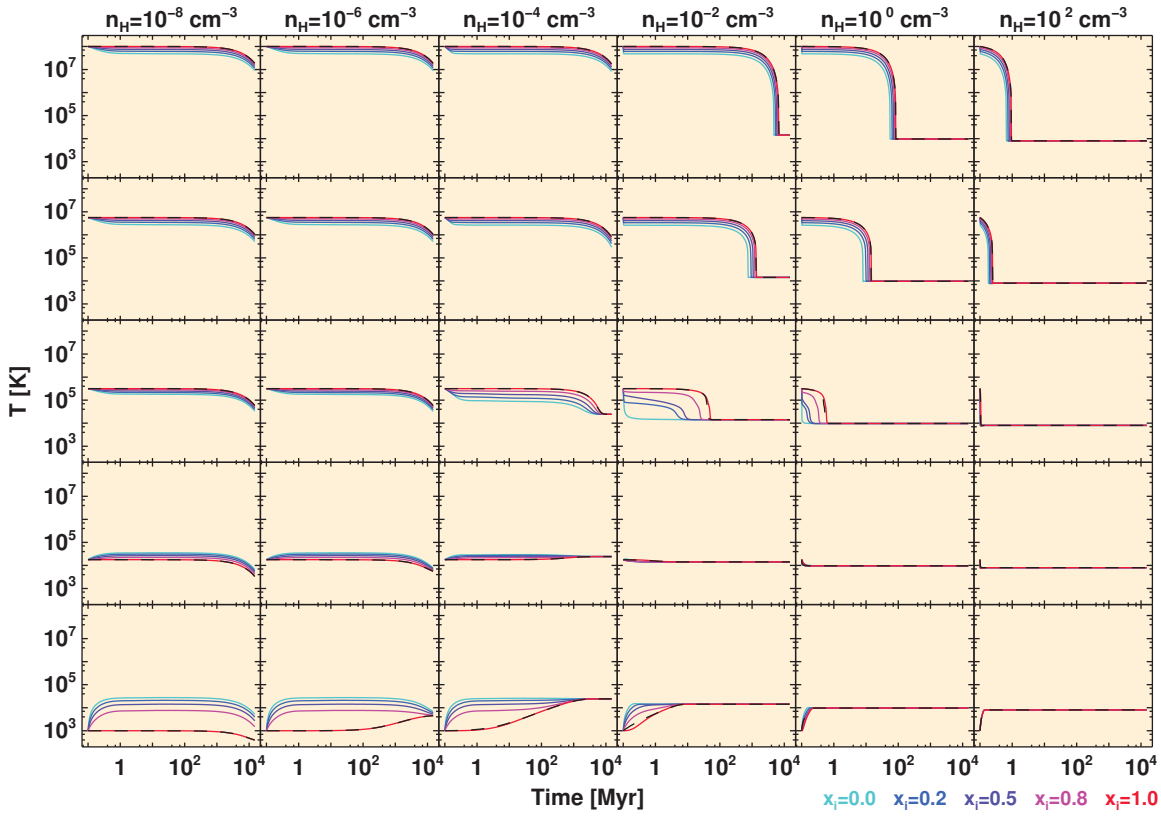


Figure 4.4: *Temperature convergence with an ionizing photon flux of* 10^5 $\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$.

4.2.4 Temperature convergence with nonzero ionizing photon flux

This is the same as the previous test, except now we apply a constant flux of 10^5 ionizing photons $\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$, assuming the spectrum of a blackbody at 10^5 K. The results are shown in Fig. 4.4. Things are much the same as before, except that the non-equilibrium temperature seems to converge to a value which is much closer to the the equilibrium one - because of the ionizing flux there is always a reservoir of electrons both in the equilibrium and non-equilibrium evolution, which makes for a much closer match in the final temperature.

Although the final temperature reached is identical between the two methods, the evolution towards that final temperature can be quite different, depending on the initial ionization states.

A zoom-in on one of the plots is shown in Fig. 4.5, and reveals that there is very little difference between the final temperatures reached. The little difference there is results from interpolation from cooling-rate tables in `Ramses` equilibrium cooling and it can be decreased further by increasing the size of these tables.

4.2.5 Thermochemistry tests conclusions

The main conclusions of the one-cell thermochemistry tests are:

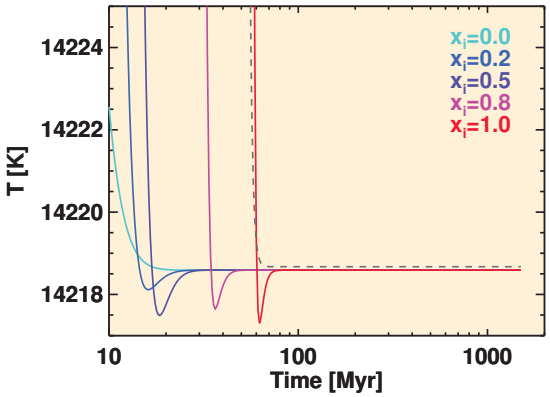


Figure 4.5: *Close-up of temperature convergence, for the UV inclusive test with initial temperature* $T \approx 10^5$ *K and* $n_{\rm H} = 10^{-2}\,{\rm cm}^{-3}$

- We always eventually reach the equilibrium ionization state with the non-equilibrium method...
- ...but this can take a very long time to happen for diffuse gas, even more than a Hubble time.
- Non-equilibrium temperature evolution of the gas is quite dependent on the initial ionization fraction of the gas at intermediate temperatures and low densities...
- ...but in the end we reach the same or at least a very similar temperature as in the equilibrium case.
- The convergence of the non-equilibrium solver towards the results of the equilibrium solver of `Ramses`, given the same cooling rate expressions, suggests that our thermochemistry solver is robust and correct.

4.3 The benchmark RT tests

The tests described in this section come from two papers that were born out of a series of workshops on radiative transfer. Tests with simple analytic solutions to compare to are hard to engineer in radiative transfer, so the solution was to instead make simple tests where the correct solution is not necessarily well known but the results of many different codes can instead be compared. Thus it is likeliest that the correct results are usually where most of the codes agree, and if a code stands out from all or most of the others in some way, this would most likely be a problem with that particular code. These tests have become sort of benchmark tests for RT codes, and most publications that present new implementations use some or all of these tests for validation.

The first paper is Iliev et al. (2006a), hereafter known as Il06 – it describes four RT post-processing tests, i.e. with the hydrodynamic advection turned off, and shows the results for

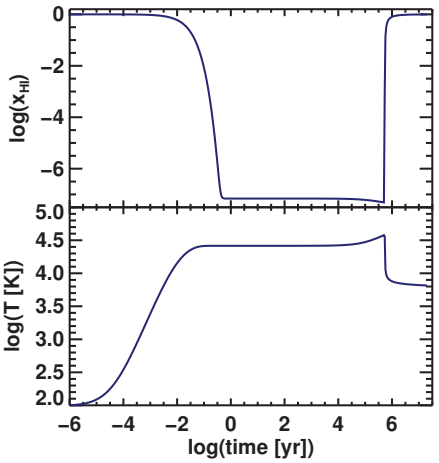


Figure 4.6: *Il06 test 0. Single-zone photoheating and ionization with subsequent cooling and recombinations.*

11 RT codes. The second paper is Iliev et al. (2009), hereafter know as Il09 – it describes three additional tests, and results for 9 codes, where the RT is coupled to the hydrodynamics.

The tests results from Il06 and Il09 are normally downloadable on the web, but at the time of this writing the links have been down for some time. However, Ilian Iliev has been nice enough to provide all test results for one of the codes, the grid based short characteristics code C^2-Ray, which is described in detail in Mellema et al. (2006). We thus present here RamsesRT results with comparisons to those of C^2-Ray. The inclusion of the C^2-Ray results in the plots shown here should be useful to guide the eye if one then wants to compare with the other codes in Il06 and Il09.

As prescribed by the test papers, all tests use hydrogen only gas. We use smooth RT in the RamsesRT runs for all tests (which is the default configuration), but remark that turning off the smoothing has no discernible effect on the results (only speed).

4.3.1 Il06 test 0: The basic thermochemistry physics

This is essentially a one-cell test of the non-equilibrium thermochemistry and not radiative transfer per se, so it doesn't really count with the rest of the comparison project tests (hence test *zero*). It is important nontheless since thermochemistry is a major new component in RamsesRT.

We start with completely neutral hydrogen gas with density $n_H = 1$ cm^{-3} and temperature $T = 100$ K at $t = 0$. A photo-ionizing flux of $F = 10^{12}$ s^{-1} cm^{-2} with a 10^5 K blackbody spectrum is applied to the gas and maintained until $t = 0.5$ Myr at which point it is switched off. The run is continued for a further 5 Myr, allowing the gas to cool down and recombine. The resulting evolution of the neutral fraction and temperature of the gas is shown in Fig. 4.6. The evolution closely follows that of the codes described in Il06, with the exception of SimpleX and FFTE which stand out somewhat, and we don't see any sign of the stiffness-induces oscillations that can be seen in the Crash code test.

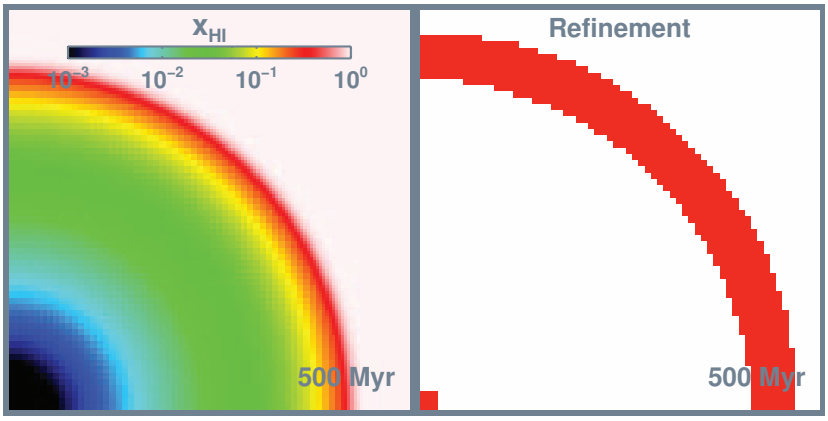


Figure 4.7: *Il06 test 1. Maps showing the neutral fraction (**left**) and grid refinement (**right**) in a box slice at $i_z = 1/128$ (first cell-slice out of 128), at 500 Myr. In the refinement map, white corresponds to the minimum effective resolution of 64^3 cells and red to the maximum effective resolution of 128^3 cells. As the map shows, the radiative source and the outer side of the I-front are refined on-the-fly during the experiment.*

4.3.2 Il06 test 1: Pure hydrogen isothermal HII region expansion

A steady monochromatic ($h\nu = 13.6$ eV) source of radiation is turned on in a homogeneous neutral gas medium, and we follow the resulting expansion of a so-called Strömgren sphere of ionized gas. Heating and cooling is turned off and the temperature is set to stay fixed at $T = 10^4$ K. The OTSA is applied in this test and we use GLF intercell fluxes.

The box is a cube of width $L_{box} = 6.6$ kpc. The gas density is $n_{\rm H} = 10^{-3}\,{\rm cm}^{-3}$ and the initial ionization fraction is $x_{\rm HI} = 1.2\,10^{-3}$, corresponding to CIE. The radiative source is in the corner of the box and the emission rate is $\dot{N}_\gamma = 5\,10^{48}$ photons s^{-1}. We use the RSLA in this test with the light speed fraction set to $f_c = 10^{-2}$. AT08 contains an analysis of the effect of different light speeds in this test and others in Il06, and finds the results start diverging non-negligibly somewhere between $f_c = 10^{-2}$ and 10^{-3}. The simulation time is $t_{sim} = 500$ Myr. As with all the tests in Il06 and Il09 (except for test 0 which doesn't involve a grid), the prescribed grid resolution is 128^3 cells, but here we do things slightly differently to demonstrate on-the-fly AMR at work. Our box resolution is 64^3 cells, but we allow for one level of further refinement, i.e. to an effective resolution of 128^3 cells that matches the prescribed resolution. The refinement criterion is applied on gradients in $x_{\rm HI}$ and $x_{\rm HII}$. According to it, two adjacent cells at positions i and $i+1$ are refined if

$$2\left|\frac{x^i - x^{i+1}}{x^i + x^{i+1}}\right| > 0.8, \tag{4.1}$$

where x is either $x_{\rm HI}$ or $x_{\rm HII}$.

The Strömgren radius, $r_{\rm S}$, is the radius of the ionization front (I-front) from the center when steady state has been reached, and in the case of fixed density and temperature it has the simple analytical solution

$$r_{\rm S} = \left[\frac{3\,\dot{N}_\gamma}{4\pi\,\alpha^{\rm B}_{\rm HII}(T)\,n_H^2}\right]^{1/3}. \tag{4.2}$$

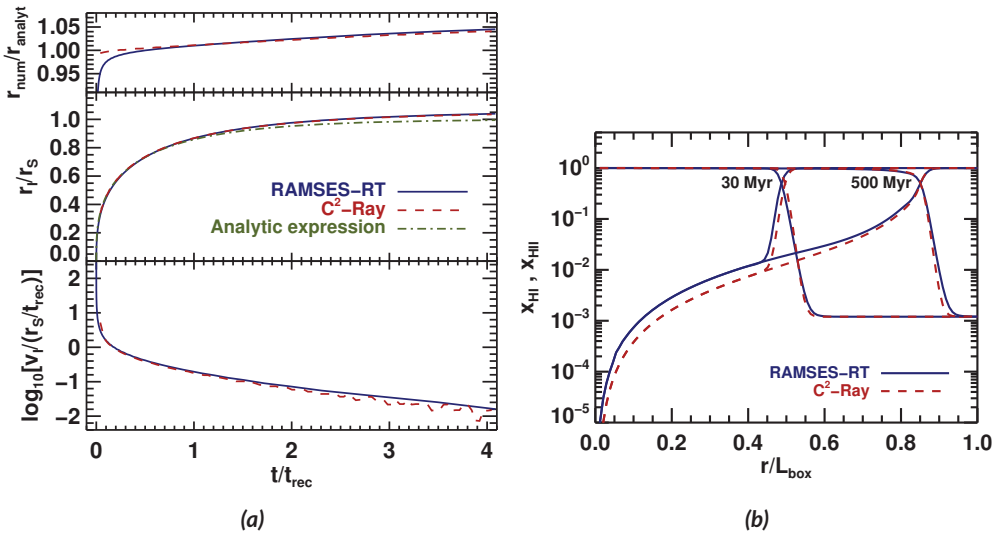


Figure 4.8: *Il06 test 1.* ***(a)*** *Evolution of the I-front position and velocity. Blue solid lines show our result, red dashed lines show the* C^2-Ray *result and green dot-dashed lines show the analytic expression.* ***(b)*** *Spherically averaged profiles for neutral fractions $x_{\rm HI}$ and ionized fractions $x_{\rm HII}$ at 30 and 500 Myr versus radius (in units of the box width $L_{\rm box}$).*

In this solution the I-front evolves in time according to

$$r_{\rm I} = r_{\rm S}\left[1 - e^{-t/t_{\rm rec}}\right]^{1/3}, \tag{4.3}$$

where

$$t_{\rm rec} = \left[\alpha^{\rm B}_{\rm HII}(T)\, n_H\right]^{-1} \tag{4.4}$$

is the recombination time. A derivation of these expressions is given in Appendix B. For the parameters of this experiment, $t_{\rm rec} = 122.4$ Myr and $r_{\rm S} = 5.4$ kpc.

Fig. 4.7 shows maps at 500 Myr of the neutral fraction (left) and grid refinement (right) in a box slice at $i_z = 0$ (out of 128, where the AMR grid has been expanded to a 128^3 homogeneous grid). The Strömgren sphere is nice and symmetric and qualitatively it can be seen to agree well with results from the RT codes described in Il06 (their Fig. 6).

Fig. 4.8a shows the evolution of the I-front position and velocity with `RamsesRT` (solid blue), compared with the analytic expression (green dot-dashed) and the result for the C^2-Ray code (red dashed), which is typical for the RT code solutions presented in Il06 and does not stand out particularly in this test. Our result can be seen to match the C^2-Ray one, though we have an initial lag due to the reduced speed of light that can best be seen in the top plot showing the fraction of the numerical result's I-front radius versus $r_{\rm S}$. The analytic $r_{\rm I}$ is typically ahead of $r_{\rm S}$ by $\lesssim 5\%$, which is simply because the analytic solution is step-like with complete ionization within $r_{\rm S}$ and none outside, whereas the real solution has a gradually evolving ionization profile with radius.

Fig. 4.8b shows spherically averaged radial profiles of the gas ionization state at 30 and 500 Myr. Again we see a good match with the C^2-Ray result. There is still a little lag in the I-front position at 30 Myr due to the RSLA and $x_{\rm HI}$ is somewhat lower inside the Strömgren

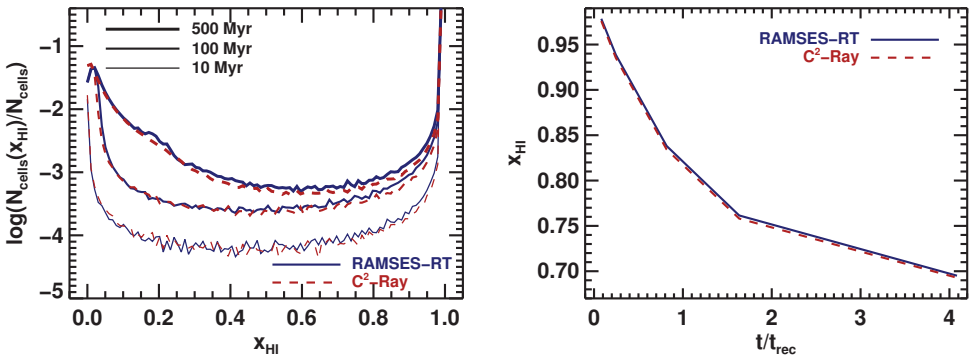


Figure 4.9: *Il06 test 1.* ***(a)*** *Histogram showing fractions of cells within bins of $x_{\rm HI}$ at three simulation times.* ***(b)*** *Evolution of the globally averaged neutral fraction.*

sphere in `RamsesRT`. However, the `C`^{2}`-Ray` result stands out a little in this test in Il06 as being most effective at ionizing the gas within the Strömgren sphere (i.e. has the lowest values of $x_{\rm HI}$), and the `RamsesRT` result is firmly typical of the Il06 codes' results in this plot.

A further comparison is made in Fig. 4.9a, here comparing ionization fraction histograms at three simulation times. Again the `RamsesRT` result closely matches the `C`^{2}`-Ray` one, whose histograms fall into a group with the codes `IFT`, `Flash-HC` and `FFTE` that stand out a little in Il06 (Fig. 9) as having less frequent intermediate neutral fractions than the other codes.

Finally for this test, Fig. 4.9b shows a comparison with `C`^{2}`-Ray` of the globally averaged neutral fraction as a function of time. It is a close match, and the `C`^{2}`-Ray` result is here firmly typical for the Il06 codes.

All in all, there is nothing out of the ordinary in the `RamsesRT` result for Il06 test 1, except for a slight initial delay of the I-front which is to be expected due to the RSLA.

4.3.3 Il06 test 2: HII region expansion and the temperature state

The setup here is the same as in Il06 test 1, except for the following points:

- We allow for cooling and photo-heating of the gas, i.e. the temperature is no longer constant, and the analytic solution, Eq. 4.2 no longer applies (because of the non-constant recombination rate).
- The initial temperature is 100 K.
- The initial ionization fraction of the gas is $x_{\rm HII} = 10^{-6}$. It should be fully neutral according to the test recipe, but this is (the default) minimum value for $x_{\rm HII}$ in `RamsesRT`, that exists in order to keep bounds on the subcycling of the thermochemistry.
- The radiation source is modelled as a blackbody with temperature $T = 10^5$ K. The emission rate is the same as before, $\dot{N}_\gamma = 5\ 10^{48}$ photons s^{-1}.

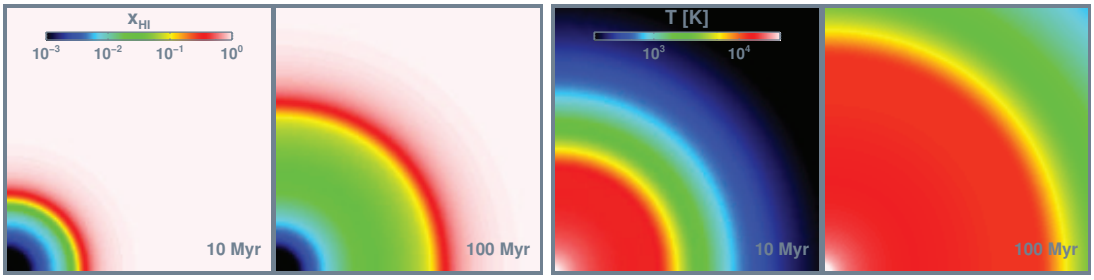


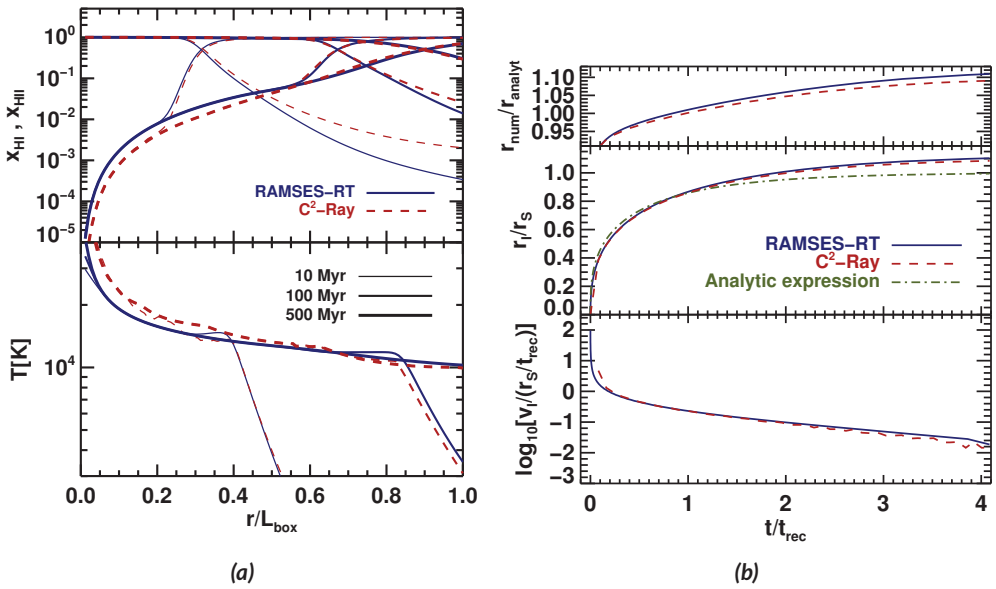
Figure 4.10: *Il06 test 2. Maps showing slices at i_z=1/128 of the neutral fraction and temperature at 10 Myr and 100 Myr.*

- Multi-frequency is approximated with three photon packages of HI, HeI and HeII ionizing photons, as in all following tests. The light speed fraction is $f_c = 10^{-2}$.
- We don't use grid refinement in this test. The grid is homogeneous and the resolution is 128^3 grid cells, as prescribed in Il06.

Slice maps at $i_z = 0$ of the neutral fraction and temperature are shown in Fig. 4.10. Both the ionization and heating fronts are smooth and symmetric, and the maps agree qualitatively with other codes in Il06 (Figs. 11-14). In comparison with the same test with `Aton` (AT08, Fig. 3), both fronts are clearly much thicker here, which is due to our multi-frequency implementation (whereas `Aton` used one photon package). More detailed comparison with the Il06 codes can be made through the ionization state and temperature plots in Fig. 4.11a, where we include the C^2-Ray result. The ionization state profile develops very similarly to that of C^2-Ray, though we have less ionization on both sides of the front, especially on the outer side where the difference in $x_{\rm HII}$ is as high as a factor of ten. Presumably this is due to the different implementations of multi-frequency photo-heating and cooling. The thermal profiles are also similar to C^2-Ray, though we have considerably lower (up to a factor of two) temperatures on the inside of the I-front, and conversely higher temperatures on the outside. As can be seen in Fig. 17 in Il06, C^2-Ray has the strongest heating of any code on the inside of the I-front in this test and most codes have stronger heating on the outside, so our thermal profiles (as the ionization state profiles) are fairly typical of the ones presented in Il06 for this test.

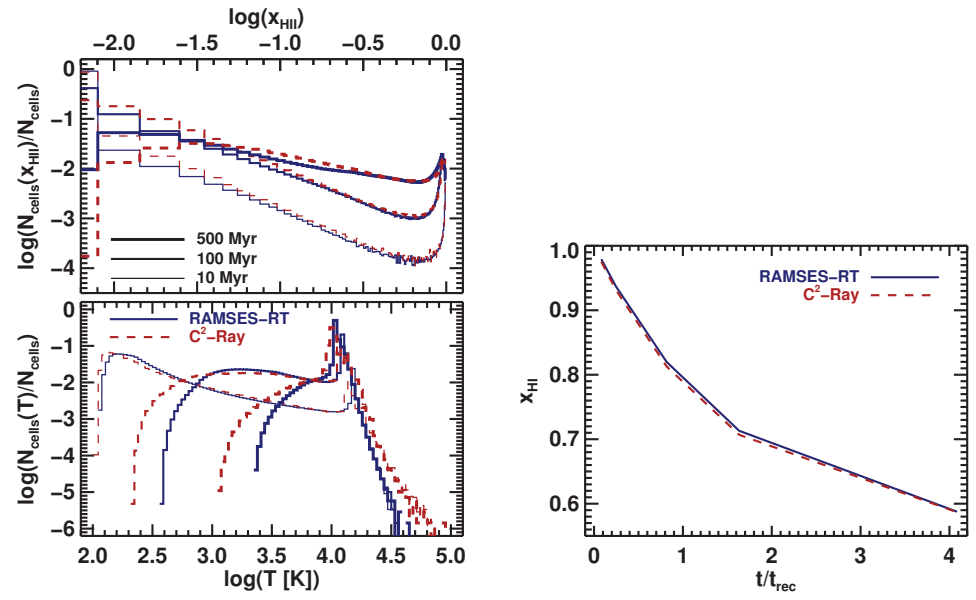
Fig. 4.11b shows the evolution with time of the ionization front, compared with C^2-Ray and the analytic solution from test 1. The front moves more slowly here than in test 1 due to the lower initial temperature, so we no longer lag behind in the initial front propagation. Our front propagates slightly further than in C^2-Ray, and ends at almost exactly the same radius as the `FFTE` code, which has the furthest travelling I-front of any code in this test in Il06. Still the difference between the codes is small, with the ratio between the numerical and analytic results ($r_{\rm num}/r_{\rm analyt}$) ranging between 1.01 and 1.11, so no cause for alarm here.

Fig. 4.12a shows histograms of the ionized fraction and temperature at different times in the test for `RamsesRT` and C^2-Ray. The ionized fraction histograms are quite similar, the biggest difference being a higher fraction of almost completely neutral gas $x_{\rm HII} \lesssim 10^{-2}$ in `RamsesRT`, which we already saw in Fig. 4.11a (top) beyond the I-front. The temperature



(a) *(b)*

Figure 4.11: *Il06 test 2.* ***(a)*** *Evolution of the temperature and ionization state profiles.* ***(b)*** *Evolution of the ionization front. The top plot shows the ratio of the radius of the I-front in the tests, $r_{\rm num}$ versus the time-evolving radius $r_{\rm analyt}$ in the analytic solution from test 1 (Eq.4.3). The middle plot shows the ratio of the test I-front radius versus the steady-state radius in the same analytic solution (Eq.4.2). The bottom plot shows the speed of the I-front, $v_{\rm I}$ in units of a 'characteristic' speed, given by $r_{\rm S}\,/\,t_{\rm rec}$.*



(a) *(b)*

Figure 4.12: *Il06 test 2.* ***(a)*** *Histograms of temperature and ionized fraction.* ***(b)*** *Time evolution of the volume average neutral fraction.*

histogram for RamsesRT differs a bit from C^2-Ray in having less extreme temperatures (C^2-Ray has both hotter gas and colder gas) but are very similar to those for the codes Art, RSPH and Crash in Il06.

Finally, Fig. 4.12b shows the time evolution of the volume averaged neutral fraction in RamsesRT and C^2-Ray, and here we see a close match. There is quite a lot of difference between the different codes in the analogue plot in Il06 (Fig. 20), with 3 groups sort of appearing, and our result closely follows those of C^2-Ray, Crash and RSPH.

As with test 1, there is nothing out of the ordinary in the RamsesRT result for Il06 test 2, except perhaps for an ever so slightly further advanced I-front than most codes in Il06 have.

4.3.4 Il06 test 3: I-front trapping in a dense clump and the formation of a shadow

This test considers self-shielding within a dense gas cloud bombarded on one side by UV radiation, and the shadow trailing on the 'dark' side - something which may find place with clouds close to sites of star-formation.

The setup is as follows: The simulation box has width $L_{\rm box} = 6.6$ kpc. We place a spherical cloud of gas in the center of the (y,z)-plane, with radius $r_{\rm cloud} = 0.8$ kpc, and it's center at $(x_c, y_c, z_c) = (5, 3.3, 3.3)$, as seen in Fig. 4.13, top left, showing an $(x,y)-$slice of gas density through the middle of the box. Outside the gas cloud we have $n_{\rm H}^{\rm out} = 2\ 10^{-4}\,{\rm cm}^{-3}$, $T^{\rm out} = 8000$ K and $x_{\rm HII}^{\rm out} = 0$, and inside we have $n_{\rm H}^{\rm cloud} = 200\, n_{\rm H}^{\rm out} = 4\ 10^{-2}\,{\rm cm}^{-3}$, $T^{\rm cloud} = 40$ K and $x_{\rm HII}^{\rm cloud} = 10^{-6}$. We apply a constant ionizing photon flux $F = 10^6\ {\rm s}^{-1}\,{\rm cm}^{-2}$ from the $x = 0$ boundary of the box (left in the Fig. 4.13 maps), and run for 15 Myr. The spectrum is a blackbody with effective temperature $T_{\rm eff} = 10^5$ K, and as before we approximate multi-frequency by splitting into three photon packages bordered by the HI, HeI, HeII photoionization frequencies. We use a light speed fraction of $f_c = 10^{-2}$. As usual the resolution prescribed by Il06 is 128^3 cells, but here we apply static AMR refinement such that the coarse resolution is 64^3 cells, but a rectangular region that encompasses the gas cloud and the shadow behind it has one level of additional refinement, making the effective resolution in the cloud and behind it 128^3 cells. The refinement region is shown in the top right panel of Fig. 4.13. As usual we use the OTSA, and in order to best capture the formation of a shadow behind the cloud, we apply the HLL flux function in this test rather than the usual GLF function. We have run identical tests though, one with the GLF flux function, and one where we use the HLL flux function but don't assume the OTSA, and we show maps of those experiments for a qualitative comparison.

Fig. 4.13 show slices at $i_z = 64/128$ of the neutral fraction and temperature at 1 and 15 Myr. From second top to bottom row are shown RamsesRT+HLL, RamsesRT+HLL without the OTSA, RamsesRT+GLF (with the OTSA) and C^2-Ray. The I-front travels fast through the diffuse medium outside the cloud, but moves much more slowly inside it, and a shadow is cast behind it. As the UV radiation eats its way into the cloud, ionizing and heating it, the shadow also very slowly diminishes in width because some photons manage to cross through the edges of the cloud. The RamsesRT+HLL maps compare very well with C^2-Ray, though the shadow is slightly thinner at 15 Myr and there is stronger heating inside the shadow; this could be due to differences in the multifrequency approach and/or photoeating. Without the OTSA, the shadow is diminished from the sides due to photons being cast from the surrounding gas. AT08 showed the same effect in their Aton tests, though

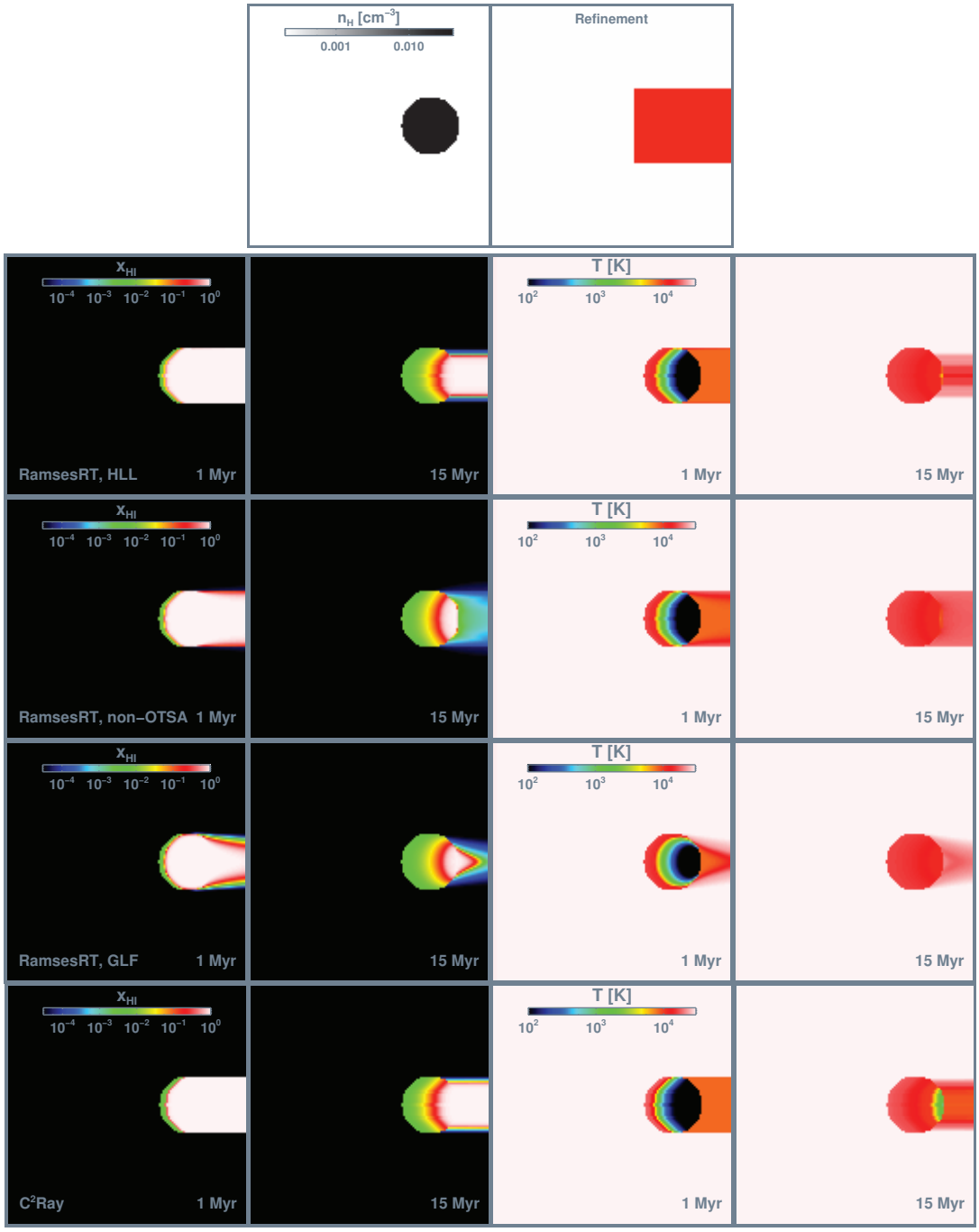


Figure 4.13: *II06 test 3. Maps showing slices at i_z=64/128. The **top row** maps show the (constant) density field (left) and the static refinement (right) applied in the* RamsesRT *run, where white/red marks effective resolution of 64^3/128^3 cells. The **second row** shows the* RamsesRT*+HLL results in terms of neutral fraction (left) and temperature (right) at 1 and 15 Myr. The **third row** Shows the* RamsesRT*+HLL results without the on-the-spot approximation. The **fourth row** shows the* RamsesRT*+GLF results. The **bottom row** shows the* C^2-Ray *results for comparison.*

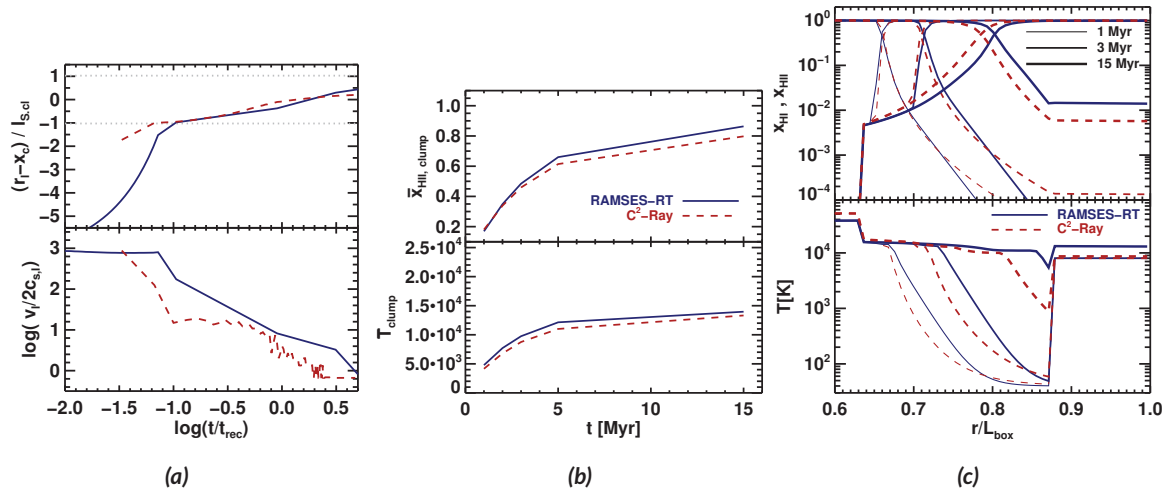


Figure 4.14: *Il06 test 3:* RamsesRT*+HLL versus* C^2*-*Ray *comparinson.* ***(a)*** *Evolution of the position and speed of the I-front along the x-axis through the center of the box. The position plot (top) shows the x-position where* $x_{\rm HII} = 0.5$*, with respect to the center of the cloud,* $x_C = 5$ *kpc, in units of the Strömgren length inside the cloud,* $\ell_{S,cl} = 0.78$ *kpc. The dotted horizontal lines mark the edges of the cloud. The speed (bottom) is plotted in units of twice the isothermal sound speed in the cloud at* $T = 10^4$ *K,* $2c_{s,I}(10^4\,\mathrm{K}) = 2.35\,10^6$ *cm/s.* ***(b)*** *Evolution of the average ionized fraction (top) and temperature (bottom) inside the dense cloud.* ***(c)*** *Profiles along the x-axis through the box center of the ionization state (top) and temperature (bottom), at 1, 3 and 15 Myr.*

there the shadow seems more diminished (this is very likely due to their monofrequency approximation, which makes their recombination photons more energetic than ours). Using the GLF flux function has much the same effect as not assuming the OTSA, though the shadow is considerably more diminished here. The result with HLL but without the OTSA is the most physical of the RamsesRT results, as one should expect recombination photons to be cast into the shadow.

Fig. 4.14a shows the evolution of the position and speed of the I-front through the center of the $(y,z)-$plane. In solid blue we plot the RamsesRT result and in dashed red is the C^2-Ray result for comparison. Horizontal dotted lines mark the edges of the cloud. There is a large initial delay in the I-front compared to C^2-Ray, which is because in the diffuse gas outside the cloud, the I-front speed is limited by the reduced speed of light. After the I-front gets into the cloud (lower dotted line) it quickly catches up and then evolves in a similar fashion in the two codes. If compared to the rest of the codes in Il06, it turns out that the evolution of the I-front in C^2-Ray slightly stands out from the rest of the codes (e.g. a small upwards 'bump' in the front position at $log(t/t_{\rm rec}) \sim 0$, and a slightly shorter distance of the I-front from the origin at the end of the simulations), and most of the others in fact evolve very similarly to that of RamsesRT. The comparison appears best with RSPH, which has the furthest extended I-front at the end-time of 15 Myr. The same can be said for the speed of the front. If we look away from the initial ~ 0.2 Myr, when our I-front has to catch up, the speed compares reasonably to C^2-Ray, and quite well to the other codes in Il06.

Fig. 4.14b shows the evolution of the mean ionized fraction and temperature inside the

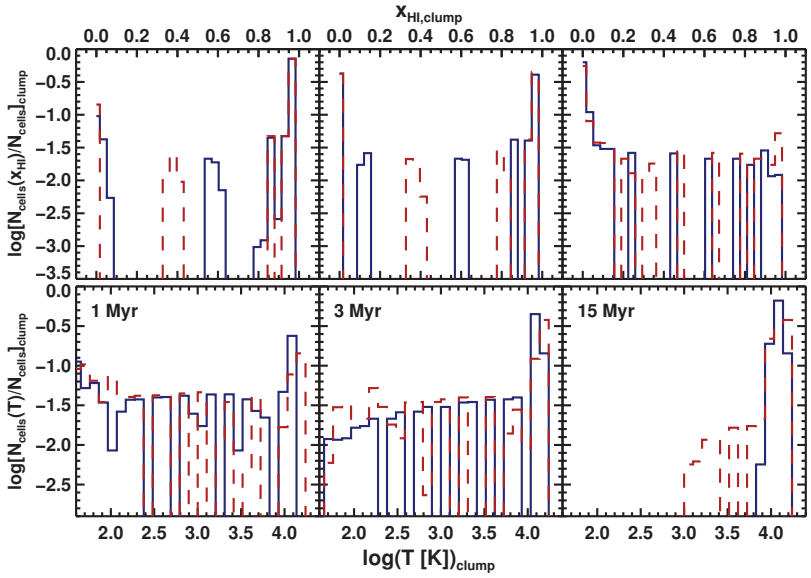


Figure 4.15: *Il06 test 3:* RamsesRT*+HLL versus* C^2-Ray *comparison. Histograms of neutral fraction (top row) and temperature (bottom) inside the dense cloud at 1, 3 and 15 Myr (from left to right).*

cloud, compared between RamsesRT and C^2-Ray. The evolution is similar between the two codes in both cases. Compared with the other codes in Il06, the evolution of the ionized fraction is most similar to RSPH, IFT and Coral, while the temperature in RamsesRT is consistently a little higher than in most codes (all except Coral and Flash which stand out quite a lot in mean temperature).

Fig. 4.14c shows profiles of the ionization state and temperature along the x-axis at the center of the $(y,z)-$plane at 1, 3 and 15 Myr. The ionization state profile in RamsesRT is similar in most respects to that of C^2-Ray, though it extends a bit further at the end of the run-time. There is initially less ionization on the far side of the front in RamsesRT, but at the end of the run this is reversed and we have slightly more ionization on the far side in RamsesRT. This 'shift' can be explained by the temperature profiles: At early times the cloud is efficiently shielding the far side from even the high-energy photons in both codes, but at the end of the RamsesRT run the shielding buffer in the cloud is thin enough that the high-energy photons can get through, hence efficiently heating the gas inside the buffer as well as in the shadow, and the gas in the shadow becomes slightly ionized as a consequence. The analogue ionization state profiles for the other codes in Il06 are mostly similar to ours. Most of them are actually closer to the RamsesRT than the C^2-Ray profile, with the exception of Crash which has a much more underdeveloped I-front and less ionization, and FFTE and IFT which have an almost step-wise x_{HII}-profile on the far side of the I-front. The temperature profiles differ pretty widely between the codes. RamsesRT doesn't particularly stand out, though, and is most similar to that of Coral at 15 Myr. The temperature profile for RamsesRT also differs notably from that of Aton, where the shielded region inside the cloud is thicker and more step-like both in the ionized fraction and temperature, due to the monochromatic radiation.

Finally, Fig. 4.15 shows histograms of the neutral fraction and temperature at 1, 3 and 15 Myr for RamsesRT and C^2-Ray. The comparison (also with the other codes in Il06) is qualitatively OK, though there is quite a difference between the individual codes in these plots.

As with the previous tests, RamsesRT performs well here and we don't really have anything out of the ordinary in our results. One should keep in note though that here we've use the non-diffusive HLL flux function, whereas in most cosmological simulations it would be more natural to use the more diffusive GLF function to have better spherical symmetry around radiative stellar sources, which comes with the price of less pronounced and shorter lived shadows than HLL. As discussed in AT08 though, shadows are not really pronounced or long-lived in a cosmological context, partly because radiative sources are typically spread around and partly because of recombination-emitted UV photons. Our analogue shadow experiment with the GLF intercell flux function gives results similar to the same experiment with HLL and non-OTSA: The shadow is short-lived. One should therefore perhaps not be overly worried that using moment-based RT with the GLF flux function doesn't cast long-lived shadows in cosmological or galactic simulations – these shadows are likely never long-lived anyway.

4.3.5 Il06 test 4: Multiple sources in a cosmological density field

This test involves the propagation of ionization fronts in a static hydrogen-only density field taken from a cosmological simulation snapshot at redshift 9. The density cube is 128^3 cells and its width is $500\,h^{-1}$ co-moving kpc (corresponding to $50\,h^{-1}$ physical kpc). The Hubble factor is $h = 0.7$. The initial temperature is fixed at 100 K everywhere. 16 radiative sources are picked out corresponding to the most massive halos in the box and these are set to radiate continuously for 0.4 Myrs, assuming a black-body spectrum with effective temperature $T = 10^5$ K. The mass-dependent radiation intensity for each halo is given in a downloadable table (from the RT comparison project website). Our computations are run with the GLF solver and the OTSA is not applied. Multi-frequency is approximated with three frequency bins of HI, HeI and HeII ionizing photons. A full light speed is used (i.e. $f_c = 1$), and an analogue run is made with a hundred-fold light speed ($f_c = 100$).

Fig. 4.16 shows box slices, at i_z=64/128, of the neutral fraction and temperature at times 0.05 and 0.4 Myr. Shown are our two runs with different light speed fractions (top and bottom row), and for comparison we show the result for the C^2-Ray code, from Il06[1]: The I-fronts and photo-heating in our $f_c = 1$ run clearly lag behind the C^2-Ray result, and there is also less heating of the ionized gas. This is in accordance with the Aton results described in AT08, where a similar delay was found. They prescribed this delay to the fact that Aton is monochromatic, but since our multi-frequency approximation (three photon packages) gives results that are still much more similar to the Aton results than those of C^2-Ray, especially in terms of the neutral fraction maps, we are inclined to blame the delay on another factor, which is the speed of light. Our results with the speed of light set to one-hundred times the physical value are shown in the bottom row of Fig. 4.16 and here the

[1]Note that Il06 have likely mislabeled the maps showing the results from this test; their text and captions indicate the maps to be at 0.2 Myr, but judging from the downloadable data they are at 0.4 Myr.

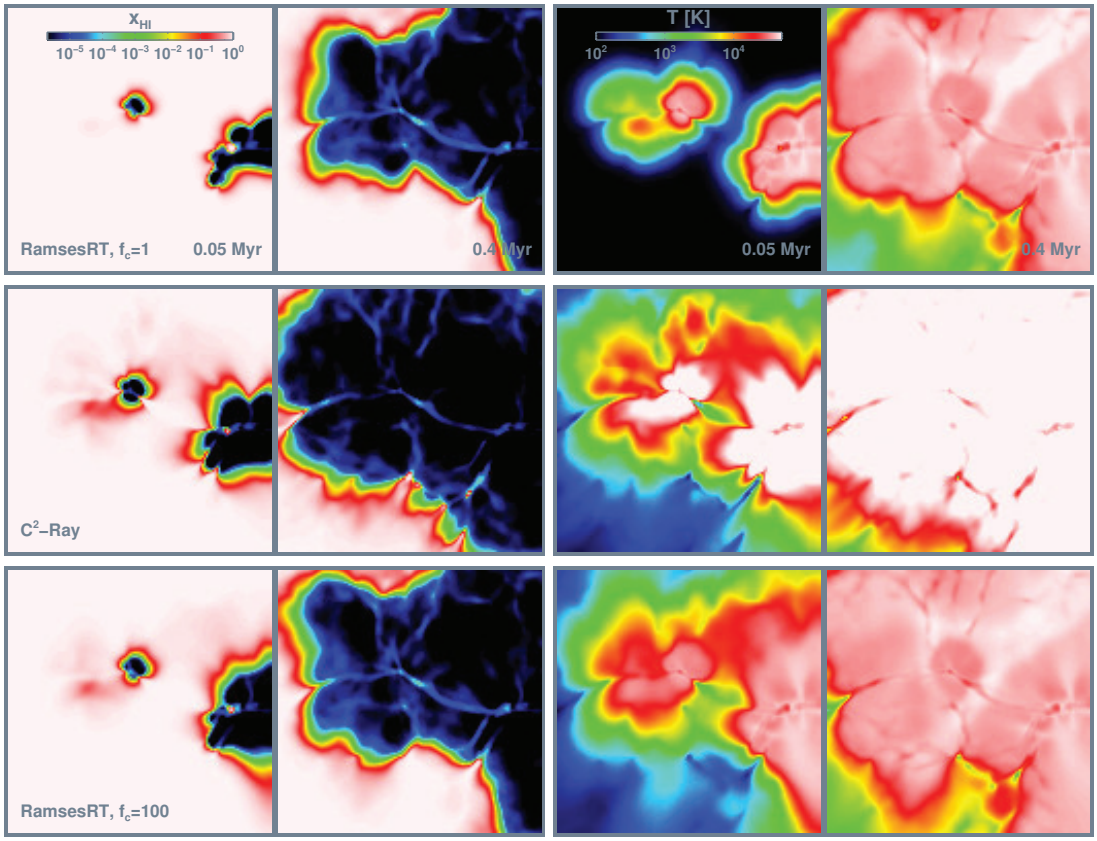


Figure 4.16: *Il06 test 4. Maps showing slices at i_z=64/128 of the neutral fraction and temperature at times 0.05 Myr and 0.4 Myr. Top row shows* `RamsesRT` *results with physical light speed. The middle row shows the* C^2`-Ray` *results (infinite light speed). The bottom row shows the* `RamsesRT` *results with one hundred times the physical light speed.*

results are considerably closer to those of C^2`-Ray` in terms of the propagation of heating- and I-fronts, although the maximum temperature in the ionized gas is still colder in comparison. All four codes considered in the Il06 4 test use an infinite effective speed of light and this may give premature fronts in the immediate vicinity of the sources and also further away in under-dense regions. Thus we are perhaps not really dealing with a delay in `RamsesRT`, but rather premature fronts in the Il06 codes. As AT08 note, we are far from reaching a static state in the fronts in this experiment in the run-time of 0.4 Myr and we should expect the different light speed runs to converge to similar solutions when static state is reached.

The smaller degree of photo-heating in the ionized gas compared to the C^2`-Ray` results is in line with the temperature profiles from the previous tests (e.g. Fig. 4.11a), and presumably results from the different ways multi-frequency is approximated.

Another notable difference in the maps in Fig. 4.16 is that our fronts are smoother and less jagged than those in C^2`-Ray`. This is an effect of the photon diffusion inherent in the GLF flux function used here. Like AT08 we find that using HLL instead gives more jagged fronts.

Fig. 4.17a shows the evolution of the mass- and volume-weighted ionized fractions, compared for the different runs. The `RamsesRT` run with the physical light speed gives ionized

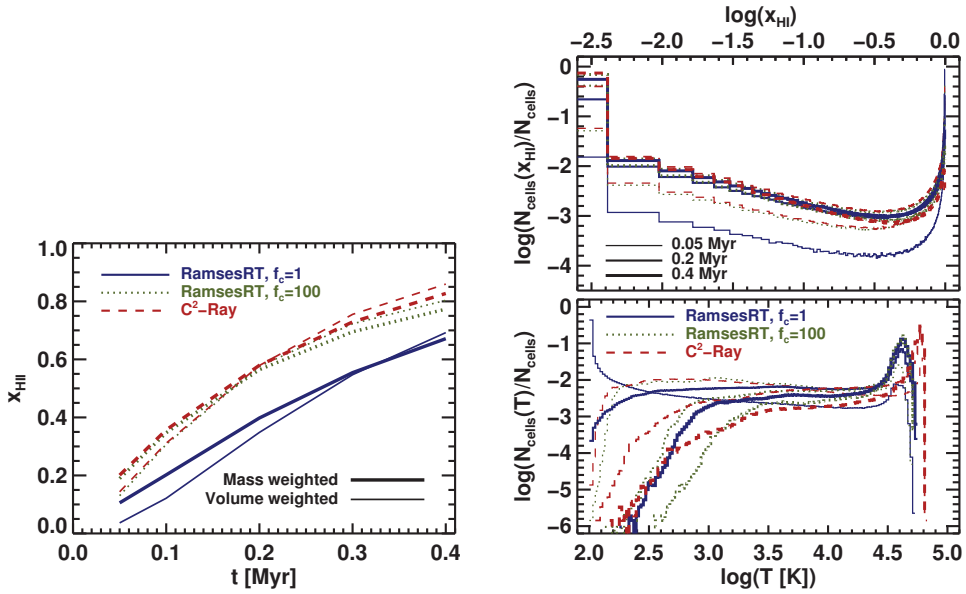


Figure 4.17: *Il06 test 4.* ***(a)*** *Time evolution of the mass weighted and volume weighted average ionized fractions.* ***(b)*** *Histograms of neutral fraction (top) and temperature (bottom).*

fractions which are close (both mass- and volume-weighted) to the `Aton` ones, whereas increasing the light speed by a factor of hundred from the physical value gives results closer to `C`^{2}`-Ray` (as well as the three other codes that ran this test in Il06). Presumably we would converge further towards `C`^{2}`-Ray` in the limit of infinite light speed, but computational time constraints do not allow to pursue that investigation. This is a further hint that the correct speed of light is important in the non-steady regime of ionization fronts.

Finally, Fig. 4.17b shows neutral fraction and temperature histograms at three times in the test. Again there is a strong discrepancy between the `RamsesRT` run with $f_c = 1$ and `C`^{2}`-Ray`, especially at early times, and the gap all but closes when $f_c = 100$ is used instead with `RamsesRT`. There remains some difference though in the minimum/maximum temperature, being smaller/larger for `C`^{2}`-Ray` than for our $f_c = 100$ run, presumably because of our rather crude multi-frequency approximation.

To summarize, there is notable discrepancy between the `RamsesRT` results and those presented in Il06, in that the `RamsesRT` ionization front lags behind, which appears to be due to a finite speed of light. In Appendix B we consider idealized I-front speeds and indeed find that they are likely to be over-predicted by infinite speed-of-light RT implementations in underdense regions such as the cosmological voids in this test. This is corroborated to some degree by others in the literature: Wise and Abel (2011) use a finite light speed and seem to get results which are slightly lagging as well, and Pawlik and Schaye (2008) specifically do a comparison between finite and infinite light speed, with the finite one resulting in a delay which is substantial, though it appears a bit less than ours.

Other sources conflict our conclusion: Petkova and Springel (2010a) use a finite light

speed, though with a softer UV spectrum than us, and get results which seem to compare very well with those of C^2-Ray.

Thus the evidence is inconclusive and the culprit in our lagging I-front may be something other than the speed of light. One possibility is that the moment method is having problems in dealing with multiple sources.

4.3.6 Il09 test 5: Classical HII region expansion

We now come to the tests described in second radiative transfer codes comparison paper by Iliev et al. (2009), which we denote as Il09. This paper provides 3 code comparison tests to add to those in Il06, but with the important difference that whereas the Il06 tests are pure radiative transfer post-processing tests with fixed density fields, the tests in Il09 are RHD tests, i.e. with the radiative transfer directly coupled to the gas-dynamics. Thus we now switch from the context of post-processing RT to hydro-coupled RHD. Here, the pressure buildup in photo-heated gas causes it to expand. Typically, the I-front is initially *R-type*, where it expands much faster than the gas response to it, which means RT-postprocessing is a fairly good approximation. The I-front then begins to slow down when it approaches the Strömgren radius, but gets moving again when the gas catches up to it, and then the front is *D-type*, i.e. moves along with the expanding gas.

As before we compare our RamsesRT tests results with those of the grid-based short characteristics ray-tracing code C^2-Ray (Mellema et al., 2006), here coupled to the Capreole code, which employs a Riemann solver for the hydrodynamics. As the Capreole+C^2-Ray combination performs badly on Il09 test 6 due to instabilities, we compare also in that particular test to C^2-Ray coupled to the Eulerian TVD solver of Trac and Pen (2004). The test numbers continue from the Il06 paper, thus we now come to Il09 test 5, which concerns the expansion of an ionization front due to a point source in an *initially* uniform-density medium. The initial setup, much like that of Il06 test 2, is as follows:

The box cube is $L_{\rm box} = 15$ kpc in width. The gas is hydrogen only as usual, initially homogeneous with density $n_{\rm H} = 10^{-3}\,{\rm cm}^{-3}$, temperature 100 K, and ionization fraction $x_{\rm HI} = 10^{-6}$ (Il09 prescribes $x_{\rm HI} = 0$). The radiative source is in the corner of the box and the emission rate is $\dot{N}_\gamma = 5\ 10^{48}$ photons s^{-1}. The spectrum is that of a blackbody at an effective temperature of 10^5 K, and as usual we bin the spectrum into three packages of HI-, HeI-, and HeII-ionizing photons. To keep the computational cost reasonable we set the light speed fraction to $f_c = 1/100$. We use the GLF flux function and don't apply the OTSA, i.e photons are emitted from gas recombinations. The box boundaries are reflective at the sides touching the radiative source and transmissive elsewhere. The simulation time is 500 Myr. The base resolution of the box is 64^3 cells and we apply on-the-fly refinement on $n_{\rm H}$ and $x_{\rm HII}$ gradients (see Eq. 4.1), so that the ionization front has the prescribed effective resolution of 128^3 cells.

We first compare volume dissections at $i_z = 1/128$ in the simulation cubes at 100 and 500 Myr, for the RamsesRT and C^2-Ray results, shown in Fig. 4.18. The maps show, from left to right, the neutral fraction, pressure, temperature, density and mach number, $M \equiv v/c_S$, where $c_s = \sqrt{1.4\ P/\rho}$ is the sound speed. (Unfortunately the M output is missing from the C^2-Ray results we've downloaded.) In these maps, the RamsesRT results look very similar to those of C^2-Ray. The $x_{\rm HI}$-maps show stronger ionization immediately around the corner

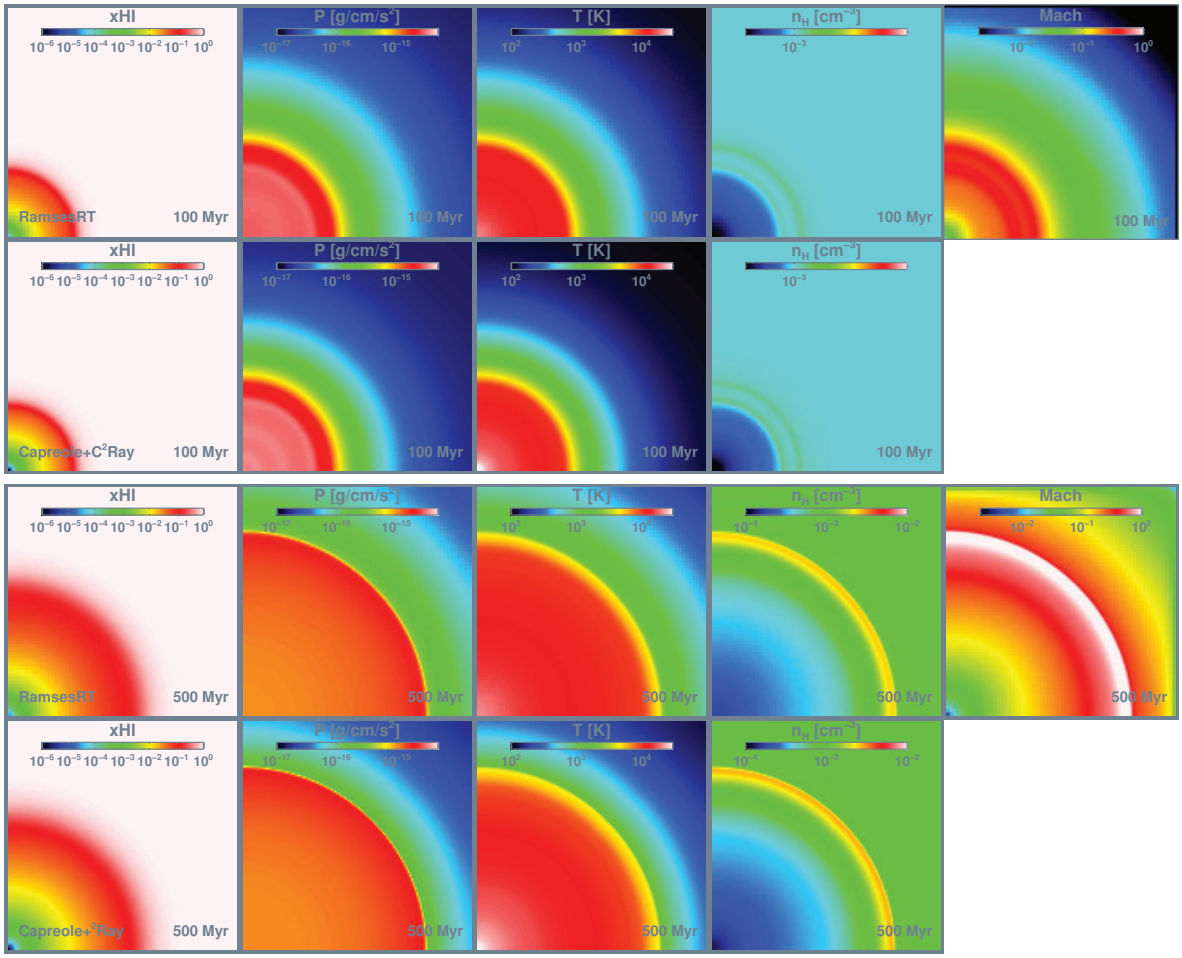


Figure 4.18: *Il09 test 5. Maps showing slices at i_z=1/128 of various quantities at 100 Myrs (top panel) and 500 Myrs (lower panel). In each panel, the top row shows the* `RamsesRT` *results and the lower row shows the* `Capreole+C`^{2}`-Ray` *results for comparison.*

source in the `C`^{2}`-Ray` result, and correspondingly the temperature and density maps show this corner gas is also hotter and more diffuse in the `C`^{2}`-Ray` result than in `RamsesRT`. Conversely, the photo-heating region is somewhat further-reaching in the `RamsesRT` result than in `C`^{2}`-Ray`, as can be seen in the pressure and temperature maps. These small differences are likely due to the different approaches in approximating multi-frequency. Notably, the `C`^{2}`-Ray` maps stand out in a very similar way when compared to most of the corresponding maps from other codes in Il09, i.e. a stronger effect close to the radiative source but shorter-reaching photo-heating.

To get a more quantitative picture, Fig. 4.19 compares radial profiles of the same quantities ($x_{\rm HI}$, P, T, $n_{\rm H}$ and M) for `RamsesRT` and `C`^{2}`-Ray` at 10, 200 and 500 Myr. The ionization state profiles (top left) indeed show `C`^{2}`-Ray` to ionize the gas more strongly close to the radiative source, but `RamsesRT` to ionize more strongly beyond the I-front. The I-front itself is however at very similar positions at all times. The pressure and temperature plots show the same thing, but apart from these minor differences at the extreme ends the shapes are very similar.

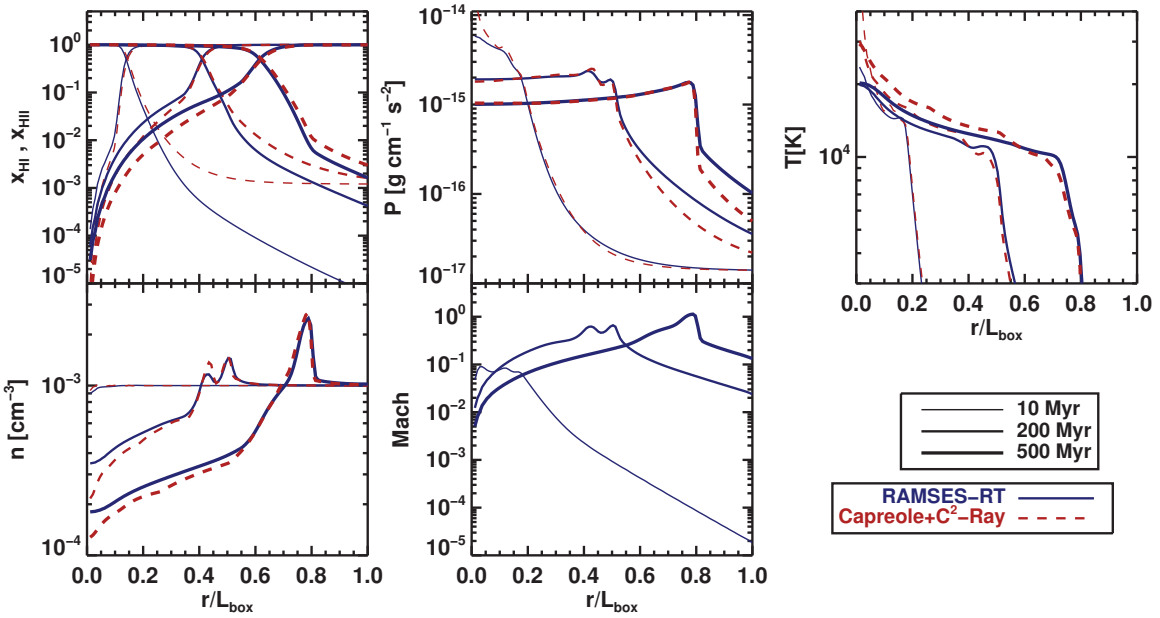


Figure 4.19: *Il09 test 5. Radial profiles at 10, 200 and 500 Myrs, compared to the* `Capreole+C`^{2}`-Ray` *results. Clockwise from top left: Ionization fractions, pressure, temperature, Mach number, atom number density.*

The density plots show that `C`^{2}`-Ray` has more has more diffuse gas close to the source as a result of the stronger photoheating, and also it appears to have a more pronounced backflow peak around 200 Myr (this double peak is a temporary effect of photo-heating by high-energy photons beyond the I-front). The smaller backflow peak in `RamsesRT` is perhaps in part a relic of on-the-fly refinement, though most of the codes in Il09 actually have backflow peaks similarly smaller than that of `C`^{2}`-Ray`. Unfortunately we can't compare the Mach profiles directly, but the `RamsesRT` profiles do look very similar in shape to those presented in Il09 (see their Fig. 15).

Finally, Fig. 4.20a shows how the position and velocity of the I-front (defined as where the radial average of $x_{\rm HII}$ is equal to 0.5), for `RamsesRT` and `C`^{2}`-Ray`. The plots for the two codes are virtually identical, the only noticeable difference being a slight initial lag in the front speed. One might attribute this to the reduced speed of light in the `RamsesRT` run, but actually most other codes described in Il09 have a very similar lag in the initial front speed compared to `C`^{2}`-Ray`.

All in all, the `RamsesRT` results for this test compare very well with most of the codes presented in Il09. The `RamsesRT` result differs slightly from that of `C`^{2}`-Ray` in some aspects, most notably in the form of weaker photo-heating and ionization close to the radiative source and wider I-fronts. However, these are precisely the aspects where `C`^{2}`-Ray` stands out from the other codes presented in Il09.

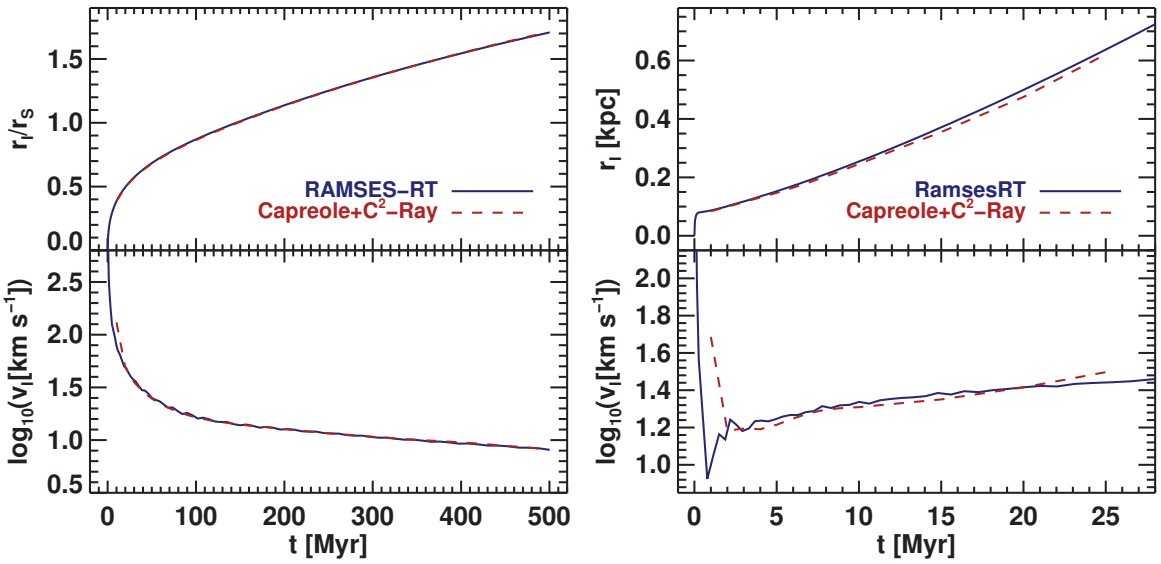


Figure 4.20: ***(a)*** *II09 test 5: Time evolution of the ionization front, compared to the results from the* `Capreole+C²-Ray` *combination. Upper plot shows the radius of the Strömgren sphere in units of 5.4 kpc. The lower plot shows the speed of the front propagation.* ***(b)*** *II09 test 6: Time evolution of the ionization front, compared to the* `Capreole+C²-Ray` *combination.*

4.3.7 II09 test 6: HII region expansion in a r^{-2} density profile

The test mimics a radiative source going off in a dense cloud, e.g. a stellar nursery. The setup is much like that of the preceding test 5, the main difference being that the gas is here non-homogeneous, the box is much smaller, $L_{\rm box} = 0.8$ kpc in width, and the radiative corner source is a hundred times more luminous. It radiates $\dot{N}_\gamma = 5\ 10^{50}$ photons s^{-1}, with a $T = 10^4$ K blackbody spectrum (softer than test 5), which is in `RamsesRT` approximated with three (HI-, HeI-, and HeII-ionizing) photon packages, that travel at $f_c = 1/100$ of the speed of light. As before we use the GLF flux function and don't apply the OTSA. The base resolution is is 64^3 cells, but on-the-fly refinement on $n_{\rm H}$ and $x_{\rm HII}$ gradients ensures the prescribed effective resolution of 128^3 cells at ionization and shock fronts. The initial temperature is 100 K everywhere and the running time is 75 Myr. The dense cloud is centered on the corner source and is set up with a spherically symmetric, steeply decreasing power-law density profile with a small flat central core of gas number density $n_0 = 3.2$ cm^{-3} and radius $r_0 = 91.5$ pc:

$$n_{\rm H}(r) = \begin{cases} n_0 & \text{if } r \le r_0 \\ n_0 (r_0/r)^2 & \text{if } r \ge r_0. \end{cases} \tag{4.5}$$

The Strömgren radius for the core density, given by Eq. 4.3, is $r_S \approx 70$ pc, which lies within the flat core. Thus, the I-front makes an initial transition from R-type to D-type within the core, and then may accelerate back to R-type as it expands into decreasingly dense gas outside the core.

We first compare the evolution of the position and speed of the I-front, which is plotted in Fig. 4.20b for RamsesRT and the Capreole+C^2-Ray combination. The I-front moves very quickly (R-type) to ≈ 70 pc within the first fraction of a Myr, stops for while and then starts to expand again with the flow of the gas. Both the speed and position compare well with C^2-Ray. The initial speed in C^2-Ray has an apparent lag which is due to under-sampling in the front positions: Other code results which are better sampled in Il09 show initial speeds that are virtually identical to the RamsesRT plot. The final front position in RamsesRT is slightly further out than that of C^2-Ray, though very similar to at least three of the codes in Il09 (Flash-HC, Licorice and RSPH). It also appears that the C^2-Ray front is starting to accelerate slightly at the end, whereas the RamsesRT front is about to approach constant speed; RamsesRT also agrees with most other Il09 codes on this point.

Fig. 4.21 shows the overall structure of ionization and the gas at 25 Myr, here with a comparison between RamsesRT (upper two rows) and TVD+C^2-Ray (bottom row). As mentioned earlier, severe instabilities appear in the Capreole+C^2-Ray version of this test, so we compare to the more stable and symmetric test with the TVD solver. In addition to the default RamsesRT run with on-the fly AMR refinement, we show here in the middle row results from an identical RamsesRT run with the base resolution set to 128^3 cells and AMR refinement turned off. There are slight spherical asymmetries appearing in the top row maps, in particular the x_{HII}, T and Mach maps, and the middle row maps are presented here to show that (the first) two of these are purely artifacts of on-the-fly AMR refinement. The slightly square shape of the inner region in the Mach map however does not seem to be due to refinement and is likely rather a grid artifact which is amplified by the radially decreasing density. It should also be noted that the other plots produced for this test (I-front, Fig. 4.20b and radial profiles, Fig. 4.22) are absolutely identical regardless of whether on-the-fly refinement is used or the full resolution applied everywhere, suggesting that AMR refinement produces very robust results.

As usual the I-front is considerably wider in RamsesRT than in the C^2-Ray results, though we don't find the same discrepancy as in the previous test between the photoheating intensity close to the source (also, there is no such discrepancy here between C^2-Ray and the other codes in Il09). The two maps furthest to the right, of density and Mach number, show the expanding shell of dense gas due to photoheating. Here the shell appears considerably thinner in RamsesRT than in TVD+C^2-Ray, and indeed TVD+C^2-Ray appears to have the thickest density shell of any of the codes in Il09 (Capreole+C^2-Ray included, but here there are also severe asymmetries). The RamsesRT maps compare well with the C^2-Ray ones, and to most of the maps in Il09, and don't show any I-front instabilities that seem to have a tendency to come up in this test (and Il09 do show that these are numerical and not physical instabilities).

Fig. 4.22 shows a comparison between RamsesRT and TVD+C^2-Ray for radially averaged profiles at 3, 10 and 25 Myr of the ionization state, pressure, temperature, density and Mach number. The comparison is generally very good. The I-front (and corresponding density shock) lag a little behind in C^2-Ray, but it actually lags a little behind all but one code in this test in Il09, and RamsesRT is spot-on compared with those others in every respect.

All in all, RamsesRT thus performs well on this test, and no problems appear that are worth mentioning.

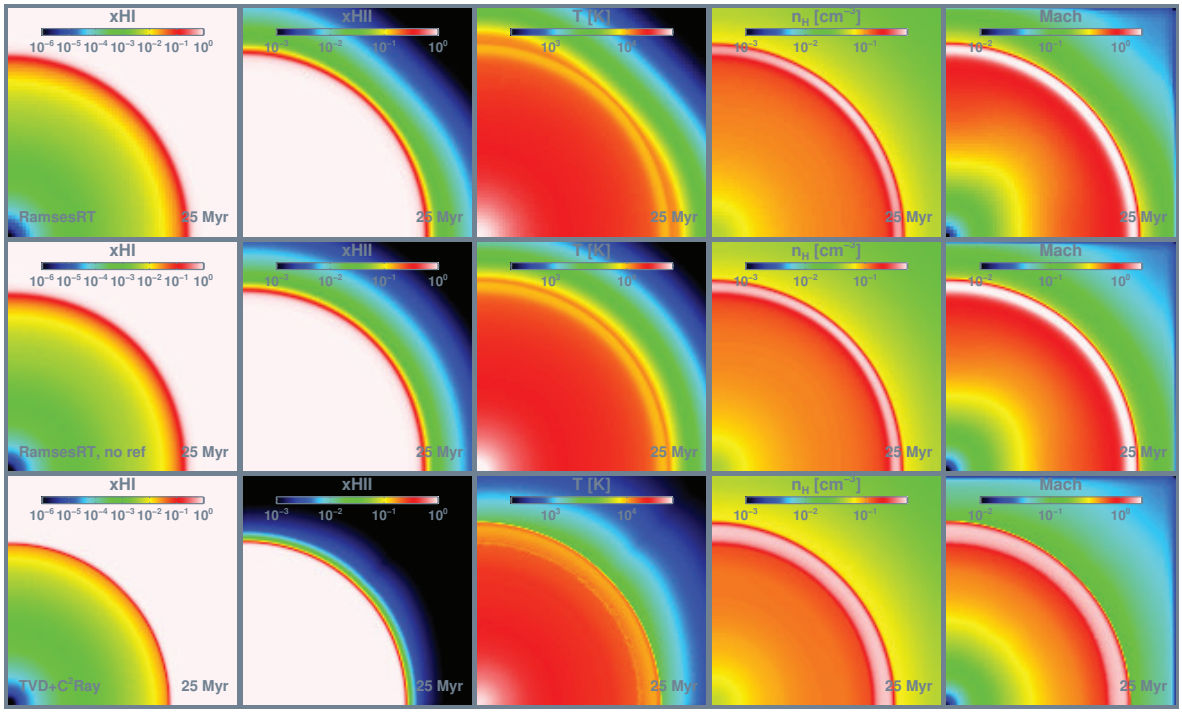


Figure 4.21: *II09 test 6. Maps showing slices at i_z=1/128 of various quantities at 25 Myrs. The top row shows the* `RamsesRT` *results with adaptive refinement. The middle row shows results also from* `RamsesRT`*, but with a fully refined box and adaptive refinement turned off. The bottom row shows the* `TVD+C`^{2}`-Ray` *results for comparison.*

4.3.8 II09 test 7: Photo-evaporation of a dense clump

The setup of this test is identical to test 3 in Il06, where UV radiation is cast on a gas cloud, creating a shadow behind it and a slowly-moving I-front inside it. Here however, since the hydrodynamics are turned on, photo-heating causes the cloud to expand outwards and simultaneously contract at the center. We recap the setup:

The box is $L_{\rm box} = 6.6$ kpc in width. A spherical cloud of gas with radius $r_{\rm cloud} = 0.8$ kpc is placed at $(x_c, y_c, z_c) = (5, 3.3, 3.3)$ kpc from the box corner. The density and temperature are $n_{\rm H}^{\rm out} = 2\ 10^{-4}\,{\rm cm}^{-3}$ and $T^{\rm out} = 8000$ K outside the cloud and $n_{\rm H}^{\rm cloud} = 200\,n_{\rm H}^{\rm out} = 4\ 10^{-2}\,{\rm cm}^{-3}$ and $T^{\rm cloud} = 40$ K inside it. From the $x = 0$ boundary a constant ionizing flux of $F = 10^6$ photons ${\rm s}^{-1}\,{\rm cm}^{-2}$ is emitted towards the cloud. The spectrum is a blackbody with $T_{\rm eff} = 10^5$ K, which is approximated with three (HI−, HeI−, HeII−ionizing) photon packages, that travel at $f_c = 1/100$ of the speed of light. The simulation time is 50 Myr, considerably longer than the 15 Myr in the corresponding pure RT test. The base resolution is 64^3 cells, but on-the-fly refinement on $n_{\rm H}$, $x_{\rm HI}$ and $x_{\rm HII}$ gradients ensures the prescribed effective resolution of 128^3 cells at ionization and shock fronts. In order to best capture the formation of a shadow behind the cloud, we focus on a `RamsesRT` run with the HLL solver, but we also show some results with the usual GLF solver. The OTSA is applied in this test.

Fig. 4.23 shows slices in the xy-plane through the middle of the box of various quantities at 10 and 50 Myr, for the `RamsesRT` result and `C`^{2}`-Ray` for comparison. As in the corresponding

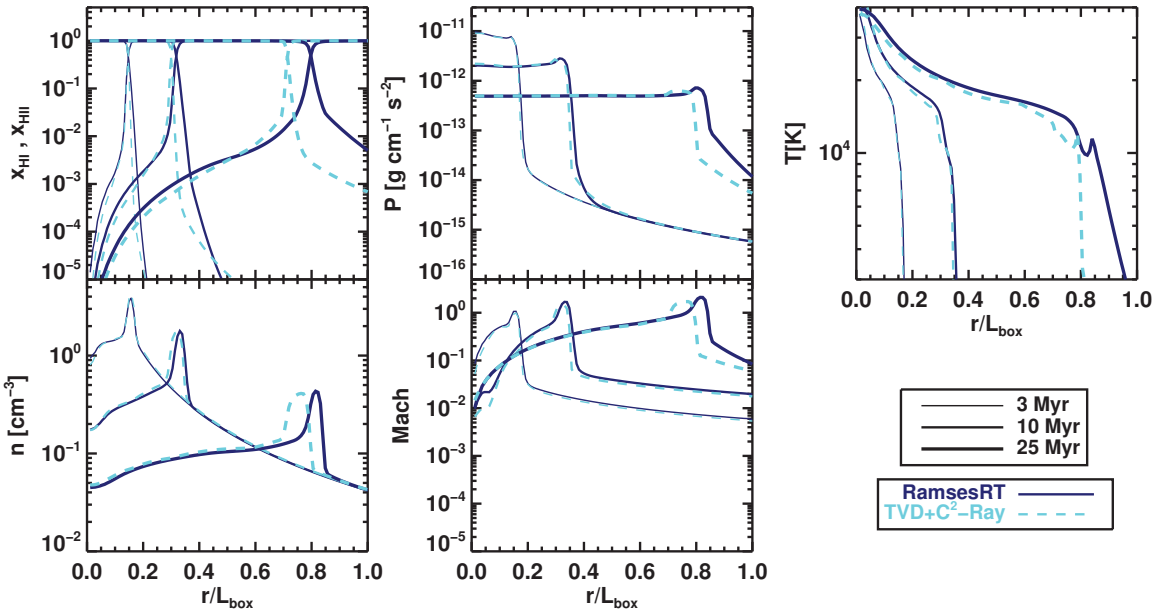


Figure 4.22: *Il09 test 6. Radial profiles at 3, 10 and 25 Myrs, compared to the* TVD+C^2-Ray *results. Clockwise from top left: Ionization fractions, pressure, temperature, Mach number, atom number density.*

pure RT test, it can be seen from the $x_{\mathrm{H_I}}$ maps that the shadow behind the cloud less conserved with RamsesRT than with C^2-Ray, though the HLL solver does a much better job though than GLF. However, the diffusion of photons doesn't have a large impact on the resulting dynamics, or even the propagation of the I-front along the axis of symmetry. The shadow becomes thinner towards the end of the run with all codes in Il09, though it is thinner than most in RamsesRT+HLL, and it pretty much disappears in RamsesRT+GLF. The shadow thickness in RamsesRT+HLL is still comparable at 50 Myrs to the results of RSPH, Zeus-MP and Licorice in Il09. The pressure maps of RamsesRT+HLL, C^2-Ray and other codes in Il09 are very similar both at 10 and 50 Myrs, though C^2-Ray, and also to some extent Flash-HC and Licorice have a fork-like shape inside what remains of the shadow at 50 Myr. The other codes have the same shape as RamsesRT+HLL in this region. The temperature maps are similar as well, though there appears here to be slightly more heating in RamsesRT both inside the cloud and in the expanding density shell around it. This is because for some reason the initial density, both inside and outside the cloud, is slightly too high in the C^2-Ray run, which also seems to result in slightly too low initial temperatures. The shell expands in a very similar way for the two codes, as can be seen in the density and Mach slices. The expansion goes a bit further, though, in RamsesRT. Also, the expanding cloud seems to develop a slightly hexagonal shape in RamsesRT, an effect which is not apparent in any of the codes in this test in Il09 (though there is a hint of it in the Flash-HC result). It can only be speculated that this is a grid artifact. To be sure it doesn't have to do with the on-the-fly refinement we ran an identical experiment with a base resolution of 128^3 cells and no refinement in RamsesRT+HLL. The RamsesRT+HLL maps and plots presented here are virtually identical to this non-refinement run, except of course for graininess in the

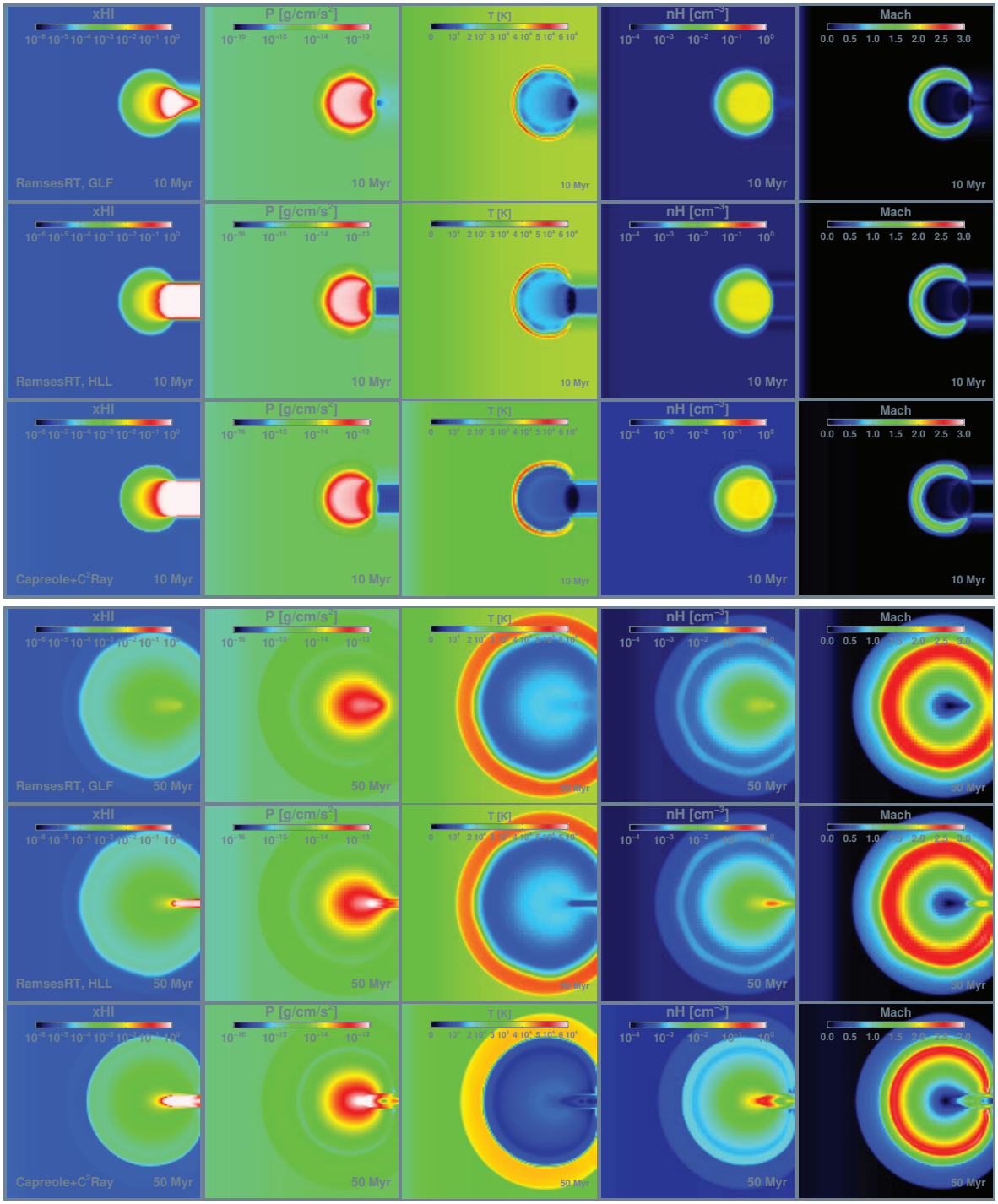


Figure 4.23: *II09 test 7. Maps showing slices at i_z=64/128 of various quantities at 10 Myrs (top panel) and 50 Myrs (lower panel). In each panel, the top row shows the* `RamsesRT+HLL` *results, the middle row shows* `RamsesRT+GLF` *and the bottom row shows the* `Capreole+C`^{2}`-Ray` *results.*

slice maps. None of these discussed effects (hexagons and a slightly over-extended I-front compared to other codes) are thus due to on-the-fly refinement.

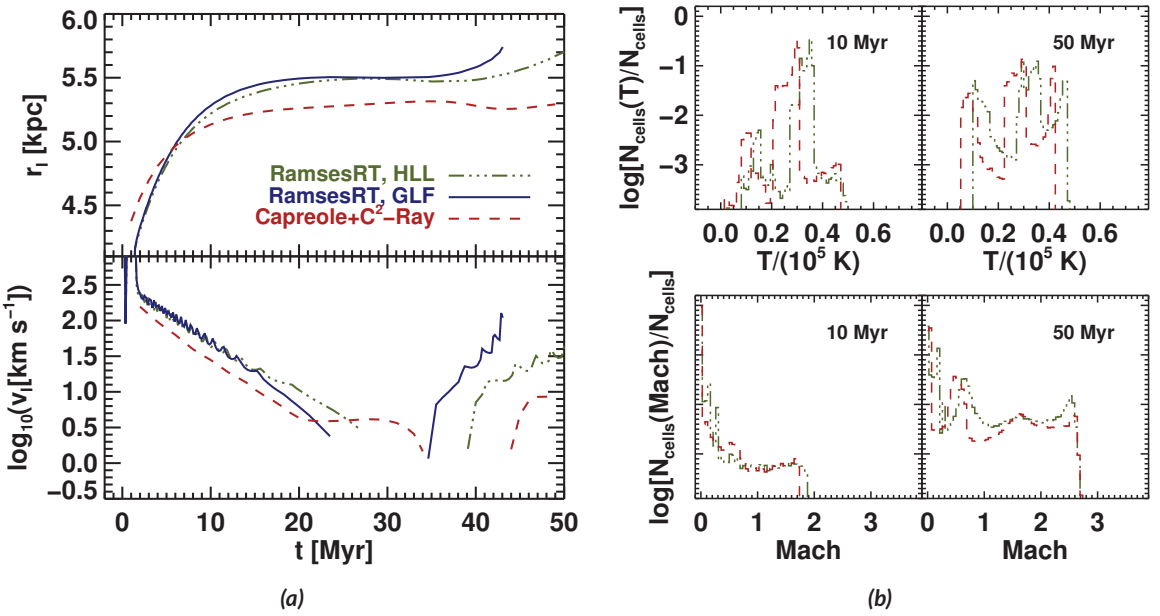


Figure 4.24: *II09 test 7.* ***(a)*** *Time evolution of the position (top) and speed (bottom) of the ionization front along the x-axis of symmetry through the center of the box.* ***(b)*** *Histograms of the gas temperature (upper panel) and flow Mach number (lower panel) at 10 and 50 Myr for* `RamsesRT` *and* `Capreole+C`^{2}`–Ray`.

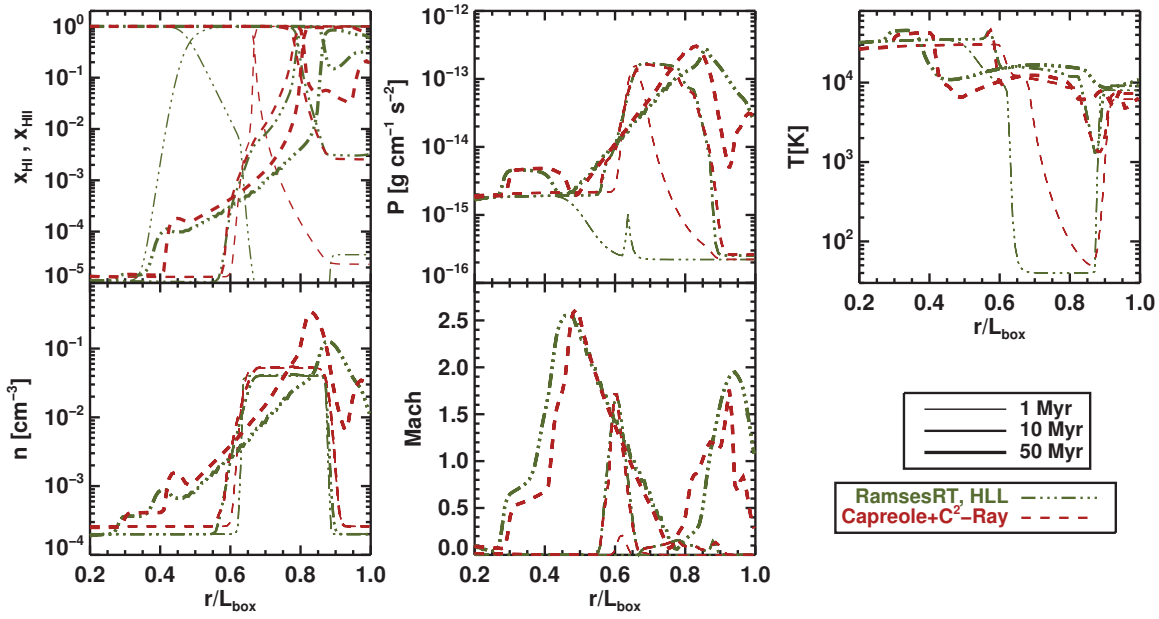


Figure 4.25: *II09 test 7. Profiles along the x-axis of symmetry through the center of the box, at 1, 10 and 50 Myr for the* `RamsesRT` *and* `Capreole+C`^{2}`–Ray` *results. Clockwise from top left: Ionization fractions, pressure, temperature, Mach number, atom number density.*

Next we turn our attention to the evolution of the position and speed of the I-front along the x-axis of symmetry through the box. This is presented for the RamsesRT (HLL and GLF) and C^2-Ray runs in Fig. 4.24a. The I-front propagation is considerably different between RamsesRT and C^2-Ray, but actually C^2-Ray considerably stands out from other codes in this test in Il09. For the first 7 Myrs or so, the RamsesRT front lags behind that of C^2-Ray and in fact all the codes in Il09. This is due to the reduced speed of light: Before hitting the cloud, the photons have to travel from the left edge of the box through a very diffuse medium – so diffuse that here the I-front speed apparently is approaching the speed of light, or is at least considerably faster than the one-one-hundredth of the light speed which is used in the RamsesRT run. However, once the I-front in the RamsesRT run has caught up, the reduced light speed should have a negligible effect on the results. After roughly 7 Myr, the RamsesRT I-front overtakes C^2-Ray front, and stays ahead of it for the remainder of the run. This however is also the case for most of the codes in Il09; their I-front is ahead of the C^2-Ray front, and four out of six codes end up with the I-front at ~ 5.6 kpc. The RamsesRT+HLL front ends up at ~ 5.7 kpc, so slightly ahead of what is typically found in Il09. Using the GLF solver instead of HLL has the effect that the I-front disappears soon after 40 Myrs, which is due to diffusive photons eating into the shadow from it's edges, but up to that point the I-front evolution is much the same. RamsesRT also reproduces the retreat of the I-front between roughly 30 and 40 Myrs, which is seen in all runs in Il09. This momentary negative speed is due to the expansion of the cloud and the D-type movement of the I-front with the gas.

Fig. 4.24b shows histograms of the gas temperature and Mach number at 10 and 50 Myr in the RamsesRT+HLL and C^2-Ray runs. The shapes of the histograms are very similar between the two codes (and are also very similar to RamsesRT+GLF, which is not shown). The temperature histograms are a bit offset, but this is also the case when the C^2-Ray histogram is compared with any of the others in Il09, and is probably just due to the slightly off starting temperature in the C^2-Ray initial conditions.

Finally, Fig. 4.25 shows a comparison between RamsesRT and C^2-Ray of profiles along the x-axis of symmetry of the various quantities at 1, 10 and 50 Myrs. The profiles compare badly at 1 Myr, but as already discussed this is simply due to the I-front having not caught up at this early time when using the reduced speed of light. At later times the profiles generally compare well, though we see these effects which have already been discussed, of a further expanding density-front out of the original cloud, a further progressed I-front, and slight differences in density and temperature which are due to slightly wrong initial conditions in the C^2-Ray run. The RamsesRT profile plots show a staircase effect which is most obvious in the 50 Myrs plot at the radial interval $0.45 \lesssim r/L_{\mathrm{box}} \lesssim 0.75$: This is simply due to the grid being unrefined at this x-interval along the axis of symmetry, i.e. at the effective base resolution of 64^3 cells per box width. The run with the full resolution and no AMR refinement shows no staircases, but otherwise the results are identical to those shown here.

We have made an alternative run with RamsesRT+HLL with the speed of light fraction set to $f_c = 1/10$ rather than the default $1/100$, and here the initial evolution of the I-front position and radial profiles at 1 Myr are almost identical to those of C^2-Ray. At later times the results are very much in line with those where $f_c = 1/100$, except the I-front position is

slightly more advanced at 50 Myr, or at 5.78 kpc rather than at 5.71 kpc.

In summary, RamsesRT performs well on this test with no apparent problems. The reduced light speed ($f_c = 1/100$) has very little effect on the results and on-the-fly refinement gives results which are identical to the fully refined simulation with a homogeneous 128^3 cells grid. Even using the diffusive GLF solver retains much of the results (I-front development, cloud expansion), except that the I-front disappears a bit prematurely.

4.3.9 Benchmark test conclusions

RamsesRT performs very well on all the tests from Il06 and Il09, with no discrepancies to speak of from expected results or those from other codes. The most notable discrepancies clearly result from the reduced speed of light approximation used in RamsesRT, though it is not always clear whether this gives results that are 'less correct' than when using infinite speed of light, as is the case for most of the codes compared in the RT comparison project. Our shadows are also considerably shorter lived with the GLF intercell flux function than those of the RT comparison project codes (most of which use ray-tracing schemes). This can be fixed for problems involving shadows and idealized geometries by using the HLL flux function instead, but as we showed in §3.2.1 the sacrifice is that isotropic sources become not so isotropic. Many codes in the RT comparison project show various instabilities and asymmetries in ionization fronts; no such features are manifested in the RamsesRT results.

5

Extended Lyα emission from cold accretion streams

As discussed in the introduction, computational UV radiative transfer has many astronomical applications. One of those involves accurately estimating the Lyman-alpha (Lyα) emissivity of extended gas structures, which is very sensitive to the thermal and ionization state of the gas, which in turn is very sensitive to the local UV radiation.

Extended Lyα radiation has been a matter of great interest over the last decade or so, largely due to the discoveries of extended and luminous Lyα blobs (LABs) at high redshifts. These blobs could be explained by a few different mechanisms, and the work described in this section, which is the first scientific application of our RHD implementation, is about studying one of these; that LABs are powered by cooling radiation in gas accretion streams entering massive halos. Aside from LABs, the aim of this work is also to study the detectability of large-scale accretion streams and estimate how close we are to observing them directly and unambiguously with upcoming instruments.

We first present in this chapter the published paper that has come out of this work (Rosdahl and Blaizot, 2012). We then elaborate on some topics that only merit a brief discussion in the paper: First of these are convergence tests on resolution and some parameter choices, made to verify that our results and conclusions can be trusted. Other topics are the option of replacing RT with a much simpler and cheaper approximation in these particular simulations; the effects of operator splitting, both in the context of the radiative transfer and the hydrodynamics; and the disruption of gas streams, which we verify to be a physical and not numerical effect.

5.1 The paper

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 5 March 2012 (MN LATEX style file v2.2)

Extended Lyman-alpha emission from cold accretion streams

J. Rosdahl⋆ † and J. Blaizot
Université de Lyon, Lyon, F-69003, France ;
Université Lyon 1, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval, F-69230, France ;
CNRS, UMR 5574, Centre de Recherche Astrophysique de Lyon.

5 March 2012

ABSTRACT
We investigate the observability of cold accretion streams at redshift 3 via Lyman-alpha (Lyα) emission and the feasibility of cold accretion as the main driver of Lyα blobs (LABs). We run cosmological zoom simulations focusing on 3 halos spanning almost two orders of magnitude in mass, roughly from 10^{11} to 10^{13} solar masses. We use a version of the `Ramses` code that includes radiative transfer of ultraviolet (UV) photons, and we employ a refinement strategy that allows us to resolve accretion streams in their natural environment to an unprecedented level. For the first time in a simulation, we self-consistently model self-shielding in the cold streams from the cosmological UV background, which enables us to predict their temperatures, ionization states and Lyα luminosities with improved accuracy. We find the efficiency of gravitational heating in cold streams in a $\sim 10^{11}$ solar mass halo to be around 10-20% throughout most of the halo but reaching much higher values close to the center. As a result most of the Lyα luminosity comes from gas which is concentrated at the central 20% of the halo radius, leading to Lyα emission which is not extended. In more massive halos, of $\gtrsim 10^{12}$ solar masses, cold accretion is complex and disrupted, and gravitational heating does not happen as a steady process. Ignoring the factors of Lyα scattering, local UV enhancement, and SNe feedback, the cold 'messy' accretion alone in these massive halos can produce LABs that largely agree with observations in terms of morphology, extent, and luminosity. Our simulations slightly and systematically over-predict LAB abundances, perhaps hinting that the interplay of these ignored factors may have a negative net effect on extent and luminosity. We predict that a factor of a few increase in sensitivity from current observational limits should unambiguously reveal continuum-free accretion streams around massive galaxies at $z = 3$.

Key words: cosmology: theory, diffuse radiation, large-scale structure of Universe, methods: numerical, radiative transfer

1 INTRODUCTION

The last decade has seen a shift in the way galaxies are thought to have assembled. In the classic theory (Rees & Ostriker 1977; Silk 1977; White & Rees 1978), galaxies collect their baryons via so-called hot mode accretion where diffuse gas symmetrically falls into dark matter (DM) halos and is shock-heated as it hits the gas residing in them. Depending on the mass of the halo, the gas may or may not eventually settle into the galaxy. However, it has become increasingly apparent through theoretical work and simulations that at high redshift ($z \gtrsim 2$), galaxies get their baryons primarily via accretion of relatively dense, cold (10^4 K) and pristine gas which penetrates in the form of *streams* through the diffuse shock-heated medium (Fardal et al. 2001; Birnboim & Dekel 2003; Kereš et al. 2005; Dekel & Birnboim 2006; Birnboim et al. 2007; Ocvirk et al. 2008; Dekel et al. 2009; Brooks et al. 2009; van de Voort et al. 2011b; Faucher-Giguère et al. 2011; van de Voort & Schaye 2011). Simulations consistently show these streams to exist and peak in activity around redshift 3, though it appears that their widths are still dictated mostly by resolution.

The problem is that cold accretion streams have never been directly observed, though we are starting to see some hints, both in emission (Rauch et al. 2011) and absorption (Ribaudo et al. 2011).

⋆ E-mail: joakim.rosdahl@univ-lyon1.fr
† Animations of our simulations can be found at http://www-obs.univ-lyon1.fr/labo/perso/joakim.rosdahl/LABs

Is this lack of observational evidence consistent with the existence of cold accretion streams? Do we not observe them because they're not easily observable or simply because they don't exist?

Faucher-Giguère & Kereš (2011) showed that the streams are hard to detect directly via absorption due to their small covering factor and surrounding galactic winds that overwhelm their signature. Kimm et al. (2011) came to the same conclusion, adding that the low metallicity in streams ($\leqslant 10^{-3}$ solar) further inhibits their detection via metal line absorption. Even so, Fumagalli et al. (2011) and van de Voort et al. (2011a) have argued that a large fraction of observed metal-poor Lyman-limit systems (LLSs) make up for indirect detections of cold streams. Furthermore, we may possibly have been directly observing the tips of these streams during the last decade in the form of Lyman-alpha blobs (LABs).

LABs are extremely bright ($\gtrsim 10^{43}$ erg s^{-1}) and extended ($\gtrsim$ 30 kpc in diameter) Lyα nebulae (e.g. Francis et al. 1996; Keel et al. 1999; Steidel et al. 2000; Matsuda et al. 2004; Palunas et al. 2004; Nilsson et al. 2006; Smith & Jarvis 2007; Prescott et al. 2009; Yang et al. 2010; Erb et al. 2011). They have a slight tendency to be filamentary in structure (Matsuda et al. 2011, hereafter M11), and often have short limbs protruding from the main body. They often coincide with galactic sources that give hints about their physical origin but the mechanism by which the emission becomes so strong and extended is a matter of debate. A subset of LABs however have no apparent coinciding galactic sources (e.g. Steidel et al. 2000; Weijmans et al. 2010; Erb et al. 2011). Up until now about two hundred LABs have been discovered, including about fifteen giant ones ($>$ 100 kpc). Smaller extended Lyα emitters exist in large quantities over a continuous range of sizes down to point sources. LABs appear to be specific to the high-redshift Universe (Keel et al. 2009) and most of them have been detected at $2 < z \lesssim 3$.

The physical nature of LABs is still a matter of debate, but by most accounts they are powered by a combination of some or all of the following processes: *(a)* Cold stream accretion is a natural explanation, where the fuel source is the dissipation of gravitational potential, also termed gravitational heating (e.g. Steidel et al. 2000; Haiman et al. 2000; Fardal et al. 2001; Dijkstra et al. 2006; Dijkstra & Loeb 2009). *(b)* Photo-fluorescence by near-lying sources, such as active galactic nuclei (AGN) or starbursts (e.g. Haiman & Rees 2001; Cantalupo et al. 2005; Kollmeier et al. 2010), *(c)* Lyα scattering, also fuelled by neighbouring star-forming regions (e.g. Laursen & Sommer-Larsen 2007; Zheng et al. 2011). *(d)* Cooling radiation in galactic outflows, fuelled by AGN or supernovae (e.g. Taniguchi & Shioya 2000; Ohyama et al. 2003; Mori et al. 2004).

Furlanetto et al. (2005) used cosmological simulations to look at the contributions of each of these processes, and found that star-forming regions can in principle power all but the largest LABs via photo-fluorescence and Lyα scattering, but that cold accretion alone cannot, except under very optimistic assumptions. They however pointed out that the Lyα emissivity of their simulated gas is highly uncertain due to the lack of modelling of self-shielding from UV radiation: The self-shielding state of the gas affects both the temperature and ionization state, which sensitively dictates the Lyα emissivity. They also pointed out that the efficiency of star-formation in powering LABs is very dependent on the presence of dust. As pointed out by Cen (2011), massive galaxies tend to have large dust content which makes them very efficiently transform their UV (and Lyα) output into infrared radiation. Thus it appears problematic to associate the largest and most luminous LABs to star-formation in the most massive halos in the Universe.

1.1 Recent work on gravitationally driven Lyα emission

Notably, two recent simulation papers have studied gravitational heating as the driver of LABs, but have reached conflicting conclusions:

Goerdt et al. (2010, hereafter G10) analyze two suites of cosmological adaptive mesh refinement (AMR) simulations. They assume self-shielding in post-processing from the UV background in accretion streams. Mock observations of halos of $\sim 4\ 10^{11} M_\odot$ at redshift 2.3 look similar to real LABs in morphology and surface brightness profile, though the association of LABs to halos of such low mass implies an unrealistically high LAB abundance. A Lyα luminosity function derived from their results is not far off from a function derived from observations, though they over-predict number densities somewhat, which implies the cooling emission in their simulations is *too* efficient. As pointed out by Faucher-Giguère et al. (2010, hereafter FG10) this overestimate appears to be due to artificial photo-heating of stream gas, which is not on-the-fly self-shielded from the UV background.

FG10 analyze cosmological smoothed particle hydrodynamics simulations to test different approaches and approximations. Based on radiative transfer (RT) post-processing results, they apply on-the-fly self-shielding by excluding UV photoionization from all gas denser than 10^{-2} H atoms per cm^3. Then they apply a Lyα transfer code to their output to model the scattering of Lyα photons towards the observer and obtain realistic mock observations. According to their results, which are in good agreement with Furlanetto et al. (2005), cooling radiation can in *principle* power LABs, provided one includes emission from gas dense enough to be star-forming to some extent. They note that this gas should be under the influence of feedback processes which introduce a large uncertainty to the cooling emission.

Although G10 and FG10 are not in good agreement on their conclusions, they both agree with Furlanetto et al. (2005) on that proper modelling of self-shielding from UV radiation is crucial to the results.

1.2 This work

We have developed a radiation-hydrodynamics (RHD) version of the AMR code `Ramses` (Teyssier 2002), which puts us in a unique position to continue the work of the aforementioned authors, to study the emissivity of accretion streams in their natural environment at high redshift in simulations that accurately and consistently model self-shielding from the UV background. We also extend previous work by simulating halos of larger masses, which are more likely to host

LABs, and by using an original refinement strategy which allows us to describe cold streams with unprecedented resolution. The increased resolution also allows us to accurately track the state of the gas up to higher densities than in the previous works. The main motivations of our work are: **(a)** Investigate whether gravitational heating is capable and sufficient as a driver of observed LABs. **(b)** Predict the observability of gravitationally powered Lyα emission from accretion streams at redshift 3.

The paper is structured as follows: Section 2 describes the simulation code and the setup of our experiments. Section 3 describes the physical properties at redshift 3 of our simulated halos over a range of masses. Section 4 presents our prediction of the Lyα emission from extended gas around galaxies and its observability. We compare with observations of LABs. In section 5 we discuss the efficiency of gravitational heating as a source of extended Lyα emission and the contribution of cosmological UV fluorescence. We discuss other factors that may affect the extended Lyα emission. Finally we conclude and discuss in section 6.

2 SIMULATIONS

2.1 Code details

We run our simulations in `RamsesRT`, a version of the AMR code `Ramses` (Teyssier 2002) which we have modified to include on-the-fly radiation-hydrodynamics describing the propagation in space of UV photons and their interaction with gas via photoionization and heating of hydrogen and helium.

The widely used `Ramses` code simulates the cosmological evolution and interaction of dark matter, stellar populations and baryonic gas, via gravity, hydrodynamics and radiative cooling. The gas evolution is computed using a second order Godunov scheme for the Euler equations, while trajectories of collisionless DM and stellar particles are computed using a Particle-Mesh solver.

The `RamsesRT` implementation and tests will be fully described in Rosdahl et al. (2012, in preparation), and here, we only briefly present the aspects of `RamsesRT` which are most relevant to the present work.

For the radiative transfer we use a moment-based method with the M1 closure relation, as described in Aubert & Teyssier (2008), the essence of which is to turn rays of radiation into a fluid with a direction of flow that corresponds to an average of rays over all angles. In contrast to the usual ray-tracing codes currently on the market, this gives the advantage that the computational load of RT does not scale linearly - in fact hardly scales at all - with the number of radiative sources in the simulation. This is a particular advantage here as we simulate a spatially continuous source of radiation, which is hard to do with a ray-tracing code. `RamsesRT` takes full advantage of the AMR structure of `Ramses` and photons are propagated through the same cells that define the baryonic gas.

Our RT solver is explicit, which means the timestep length for the propagation of photons is limited by the speed of light. This typically makes the RT timestep three orders of magnitude shorter than the hydrodynamical timestep. Since we're forced to apply a global RHD timestep which is the minimum of the hydrodynamical step and the RT one, we're faced with the rather horrifying prospect that `RamsesRT` simulations are slowed down by a factor of order one-thousand compared to non-RT simulations. To get around this, we invoke the *Reduced Speed-of-Light Approximation* (RSLA) proposed by Gnedin & Abel (2001) (see also discussion in Aubert & Teyssier 2008): The speed of light is reduced by a factor f_c, bringing the RHD timestep closer to the normal `Ramses` one and making `RamsesRT` runnable in reasonable time. In the simulations described here we use $f_c \sim 1/100$. To be sure the choice of light speed is not affecting our results, we have run analogs of our H1 simulation (see Table 1) with the light speed changed by a factor of five in either direction, i.e. $f_c = 1/20$ and $f_c = 1/500$. This has an insignificant effect on the results, and we conclude it is an acceptable approximation for our simulations.

In order to self-consistently evolve the UV field we implement non-equilibrium gas cooling that keeps track of the abundances of all ion species of hydrogen and helium. These abundances are stored in the form of three ionization fractions, as passive scalars that are advected with the gas,

$$\begin{aligned} x_{\rm HII} &\equiv n_{\rm HII}/n_{\rm H}, \\ x_{\rm HeII} &\equiv n_{\rm HeII}/n_{\rm He}, \\ x_{\rm HeIII} &\equiv n_{\rm HeIII}/n_{\rm He}, \end{aligned} \tag{1}$$

where n is number density. The non-equilibrium cooling module evolves these ionization fractions along with photon fluxes and temperature on a per-cell basis, with the timestep constraint that none of these quantities changes by more than 10% in a single timestep, using sub-cycles when needed to fill the RHD timestep.

We have tested and verified `RamsesRT` with the benchmark tests of the 'Cosmological radiative transfer comparison project' (Iliev et al. 2006, 2009), and the results will be presented in Rosdahl et al. (2012, in prep.).

2.2 Simulation setup

We run three cosmological zoom simulations, each targeting the evolution until redshift 3 of a single halo and its large-scale environment. The initial conditions are generated using MPGRAFIC (Prunet et al. 2008). We assume a ΛCDM Universe with $\Omega_\Lambda = 0.723$, $\Omega_m = 0.277$, $\Omega_b = 0.0459$, $h \equiv H_0/100 = 0.702$ and $\sigma_8 = 0.817$, consistent with seven-year WMAP results (Komatsu et al. 2011). We assume hydrogen and helium mass fractions $X = 0.76$ and $Y = 0.24$.

The masses of these three halos span almost two orders of magnitude, the least massive halo ($\sim 3\ 10^{11} M_\odot$) roughly corresponding to halos studied in G10 and FG10, and the more massive halo simulations (up to mass $\sim 10^{13} M_\odot$) based on the expectation that LABs are situated in overdense regions of the Universe (Steidel et al. 2000; Prescott et al. 2008; Yang et al. 2010). The parameters for the individual simulations, named H1, H2 and H3 in order of increasing halo mass, are listed in Table 1.

Each simulation has periodic boundaries and nested levels of refinement in a zoom-region around the targeted halo, both in DM and gas.

On-the-fly refinement is enforced inside the zoom regions according to two criteria: The first is the traditional

Table 1. Simulation parameters

Name	Box size[a] [Mpc]	Halo mass[b] [$M_\odot$]	DM res.[c] [$M_\odot$]	Gas res.[d] [pc]	f_c[e]	n_H^{UV} [f] [cm^{-3}]
H1	28.5	2.9 10^{11}	1.4 10^6	217	1/100	1 10^{-4}
H2	28.5	2.9 10^{12}	1.1 10^7	434	1/300	3 10^{-4}
H3	51.2	1.3 10^{13}	6.4 10^7	780	1/300	3 10^{-4}

[a] Co-moving
[b] DM+baryons at $z = 3$, (all the mass within the virial radius)
[c] Optimal resolution
[d] Optimal physical resolution (not co-moving) at $z = 3$
[e] Reduced light-speed fraction, see Sec. 2.1
[f] Threshold for UV-emitting gas, see Sec. 2.2

'quasi-Lagrangian' criterion, where a cell is refined if it contains more than 8 DM particles or the equivalent baryonic mass[1]. This causes concentrations of mass to be refined to the maximum, but will typically leave the resolution of cold flows many times less, which is a problem when one is most interested in these flows. The second refinement criterion, which is unique to this work, is applied on the hydrogen neutral fraction gradient. According to it, two adjacent cells at positions i and $i+1$ are refined (up to the maximum level of refinement), if

$$2\left|\frac{x_{\rm HI}^{i} - x_{\rm HI}^{i+1}}{x_{\rm HI}^{i} + x_{\rm HI}^{i+1} + x_{\rm HI}^{floor}}\right| > \Delta x_{\rm HI}, \quad (2)$$

where $x_{\rm HI} = 1 - x_{\rm HII}$ and $x_{\rm HI}^{floor}$ is a floor on the neutral fraction under which the criterion becomes inactive. In our simulations we typically use $\Delta x_{\rm HI} = 0.75$ and $x_{\rm HI}^{floor} = 10^{-3}$ in order to resolve gas streams, though we tweak these values a bit (even within the same simulation) to tread the fine line of neither under-resolving the streams nor over-resolving uninteresting regions. This enforces maximum refinement in the cold streams, so while the optimal resolution in our simulations is slightly less than in recent works, our resolution in the streams is unprecedented in cosmological simulations.

The cosmological UV background is incorporated into our simulations with an 'outside-in' method, where it is propagated from under-dense and transparent UV-emitting voids. As such, our UV background can be thought of as *quasi-homogeneous*, as opposed to the completely homogeneous and optically thin implementation commonly used in cosmological codes that lack radiative transfer (e.g. Cen 1992; Katz et al. 1996; Rasera & Teyssier 2006). The reasons we chose this model are mainly twofold. First, it is only a single step further than previous work on the subject. This allows us to isolate the effect of self-shielding, and to interpret our results in a well established theoretical framework. Second, the inside-out method would require finely tuned star formation rates and UV escape fractions for simulated galaxies to produce a 'correct' UV background, and this is a subject onto itself (see e.g. Wise & Cen 2009; Aubert & Teyssier 2010). Also, our simulations zoom in on a relatively small volume with no star formation outside, which would lead to a severe lack of external UV background radiation. We thus postpone such a model to a future paper and instead demonstrate in Sec. 5.4 that a local enhancement of UV radiation due to star formation would not significantly change our conclusions.

In practice, we use a 'void' density threshold $n_{\rm H}^{UV}$ such that all gas cells lower in density are UV emitters, and we impose the redshift-dependent UV background model from Faucher-Giguère et al. (2009) onto these cells under the valid assumption that voids are optically thin. The radiative field is then allowed to diffuse out towards denser regions. The idea is to have the void threshold low enough that it doesn't include the potentially shielded cold streams themselves, but high enough that radiation can quickly reach them (a sensitive issue due to the reduced speed of light). We use void thresholds of $n_H^{UV} \gtrsim 10^{-4}$ cm^{-3} in our simulations. Our results are not sensitive to the fine-tuning of this as long as 10^{-4} cm^{-3} $\lesssim n_H^{UV} \lesssim 10^{-2}$ cm^{-3}.

The spectral shape of the UV field is approximately taken into account by using three (HI-, HeI- and HeII-ionizing) packages of photons which are propagated independently (see Appendix A). In this work we adopt the on-the-spot approximation (OTSA), where any UV photon emitted from a recombination is assumed to re-ionize a nearby atom (i.e. within the same grid cell) – in other words, the simulated gas does not emit UV photons due to recombinations and case B recombination rates are used in computing the gas cooling rate.

For the sake of simplicity, our simulations do not include SN feedback or metals. To prevent artificial fragmentation (Truelove et al. 1997) our simulations employ a polytropic equation of state (Dubois & Teyssier 2008) as a subgrid recipe that keeps the mostly unresolved multi-phase interstellar medium (ISM) from collapsing and fragmenting. The recipe sets a density-dependent temperature floor in every gas cell:

$$T_{min} = T_0 \left(\frac{n_{\rm H}}{n_{\rm H,0}}\right)^{\gamma-1}, \quad (3)$$

where we've chosen the values $T_0 = 10^4$ K, $n_{\rm H,0} = 1$ cm^{-3}, and $\gamma = 1.6$. The value of $n_{\rm H,0}$ also corresponds to the limit above which gas is star-forming.

We identify halos in our simulation outputs with the AdaptaHOP algorithm from Aubert et al. (2004) and Tweed et al. (2009), where the virial radius of a halo, R_{vir}, is defined as the radius where the average density is 200 times the critical density of the Universe, and the halo center corresponds to the DM density maximum.

2.3 Numerical issues

We've established through convergence tests that resolution is adequate in our simulations and that the chosen parameters of light speed and UV emission threshold (f_c and $n_{\rm H}^{UV}$) do not affect our results noticeably. Three other issues should be noted:

The gravitational potential in our simulations is usually dominated by DM particles, but it is resolved to the

[1] A cell is refined if it contains a mass of baryons larger than 8 Ω_b/Ω_m m_{DM}, where Ω_b and Ω_m are the cosmological mass fractions of baryons and matter, respectively, and m_{DM} is the mass of the highest-resolution DM particles.

local cell resolution. With our strategy of optimally resolving gas streams comes the danger that we may over-resolve the gravitational potential, with few and far-between DM particles causing discreteness effects in the potential, which may lead to artificial fragmentation and complexity in the streams. This seems particularly ominous since we find in our simulations that the streams indeed become fragmented and complex around massive halos. To make sure this is not caused by an over-resolved gravitational potential we have run analogues to our simulations with smoothed potentials, which still reveal fragmented streams. So while it is hard to tell whether or not this numerical effect is nonexistent in our simulations, we can conclude that it is not dominating our results and that the complex streams are physical in nature.

Operator splitting is a widely-used method of decomposing unwieldy differential equations into separate parts that can be solved independently and in sequence (e.g. Toro 1999; Press et al. 1992). `RamsesRT` employs this method to split the radiative-hydrodynamics equations into *(i)* advection of gas and photons between cells and *(ii)* chemical reactions within the cells (radiative cooling and photo-heating). The advection part is first solved and then cooling, using the advection result as the initial state and subcycling when needed. Gas normally exists in a competition between advective/gravitational heating and radiative cooling, where the temperature 'adjusts' to a value where these processes cancel each other out. However, when the cooling time is shorter than the advection time, operator splitting may artificially give cooling the upper hand, leading to a slight underestimate in the temperature. Normally this is no big deal, but considering how sensitive Lyα emissivity is to gas temperature (see Fig. 6), this can result in a severe underestimate of Lyα emissivity in the gas. We've verified that this is indeed the case in our simulations. To prevent this from affecting our results, we restart the simulations at $z = 3$ with the global timestep reduced by orders of magnitude, to make sure it is everywhere shorter than the local cooling-time, and run until we reach convergence in Lyα luminosity (this takes a few-thousand fine-cell time-steps).

Cell merging: With the bookkeeping on ionization states, and due to the fact that cell de-refinement takes place just before outputs are written in `Ramses`, special care must be taken on cell merging. Applying the traditional method of giving a merged cell a children-averaged ionization state can sometimes result in a combination of temperature and ionized state which causes it to outshine whole galaxies in Lyα emissivity (see discussion in Sec. 4.1). To prevent this we enforce a photoionization equilibrium (PIE) ionization state to merged cells, assuming the children-averaged values of gas density, pressure, and UV flux.

3 PHYSICAL PROPERTIES OF 3 HALOS

In this section, we first review the qualitative properties of our three simulated halos, and define the different phases of the intra-halo gas. We then describe in detail the impact of self-shielding on the ionization and thermal states of cold streams, and discuss the validity of an approximate treatment of self-shielding introduced by FG10.

3.1 Basic halo properties

Gas density maps of the three targeted halos at redshift 3 are shown in Fig. 1, in close-ups of the halos and zoom-outs to show their environments. Also shown are zoom-out maps of temperature. The halos display a tendency with increasing mass towards more intense, complex and fragmented accretion, and larger and hotter domains of shock-heated intergalactic medium (IGM).

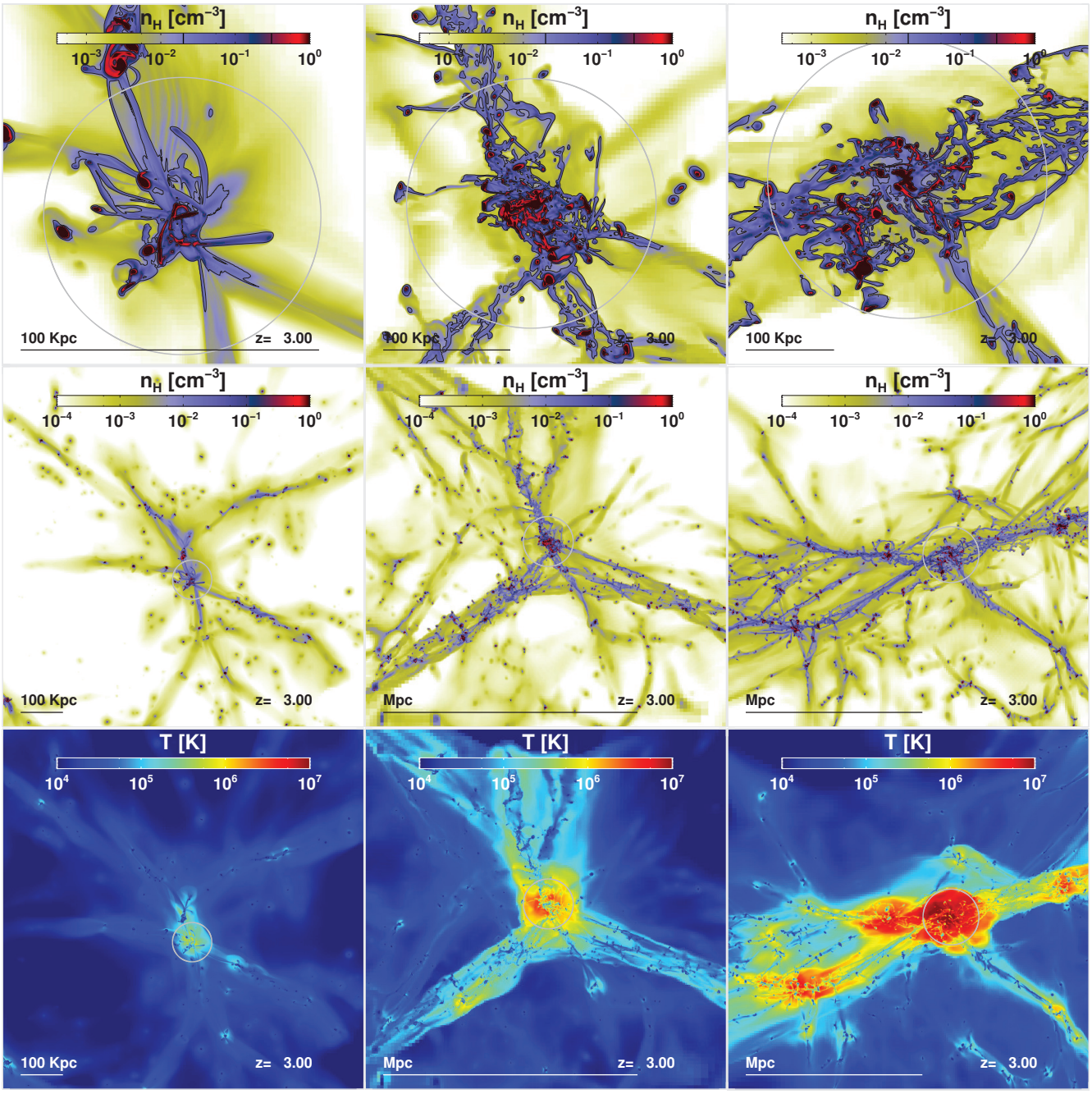
The least massive halo (H1, left) has narrow (down to $\sim$ 1 kpc in diameter) and unperturbed accretion streams and tidal tails stretching from the central galaxy, which can be seen in red at the center of the halo, edge-on and slightly inclined from the horizontal. It is about to undergo a major merger with another halo five times less massive, situated just outside R_{vir} and coming in from the north. Two parallel accretion streams bridge the merging halos. Another accretion stream extends towards a factor 100 smaller merging halo, also at the edge of R_{vir}, but towards the line of sight (LOS), seen as a moon-shaped clump to the south-west. To the south and south-east are two relatively thick and diffuse accretion streams and another one even more diffuse to the west. Other structures in the map are orbiting satellites and tidal tails.

The intermediate mass halo (H2, middle) is more a group of orbiting galaxies than a single galaxy. On the large scale there is a network of filaments mixed with galaxies of varying masses, with at least 6 large scale streams extending towards the central halo. Movies show that the accretion here is notably more spiral than around the H1 halo, with the streams starting to curve around the center of gravity already well outside R_{vir}. Inside the halo we see plenty of streams and tidal tails, but much more disrupted and messy than in the H1 halo, as a result of stronger and more frequent interactions with other streams and galaxies.

This tendency continues with the most massive halo (H3, right), where we find even more disrupted streams, to the point that many of them seem to be completely obliterated close to the halo center. The H3 halo has just undergone a major merger, which makes the accretion activity particularly violent at this point in time.

To facilitate our analysis, we apply the following categorization to divide the gas into phases, as shown in the temperature-density phase diagram in Fig. 2 (note that the categorization is specific to this paper and does not apply in general):

The star-forming ISM is all gas denser than 1 cm^{-3} and as discussed in Sec. 2.2 we apply a temperature floor in the form of a density-dependent polytrope to keep this gas from artificially fragmenting, which manifests itself in the constrained temperature-density relation in the shaded area of the diagram. Our simulations lack the ingredients to accurately model Lyα emission from the ISM (multiphase resolution, dust, Lyα scattering) and our reaction to that is to simply ignore the Lyα emission from there in our analysis. The shaded color of the ISM region in Fig. 2 should remind the reader of this and that this work is about modelling the Lyα emission coming from galactic environments and not the galaxies themselves. The ISM gas density threshold is resolution dependent and reflects the density at which further collapse of gas – i.e. the Jeans length – is no longer



H1 H2 H3

Figure 1. Redshift 3 maps of the three targeted halos in simulations H1, H2 and H3 from left to right (increasing halo mass). Grey circles indicate virial radii of the halos; 46, 98 and 158 kpc for the H1, H2 and H3 halos respectively. **Top row:** Number density maxima along the line of sight, with contours marking 0.02 and 0.3 cm^{-3} as indicated in the color bars, corresponding to our definition of streams. **Middle row:** The same but zoomed out to show the large-scale environment. **Bottom row:** Mass weighted temperature maps, on the same scale as the middle row.

resolved. At our chosen density threshold, assuming minimum temperatures of 10^4 K, the Jeans length is resolved by approximately 10, 5, and 2.5 cell widths in the H1, H2 and H3 simulations respectively. It should be noted that our density threshold is almost an order of magnitude above what has typically been used in recent similar works (e.g. G10, FG10).

The CGM is gas with number densities between 0.3 and 1 cm^{-3}. Ideally these densities form membrane interfaces between the ISM and their more diffuse environment. The lower density limit corresponds to the inner contours in the density maps of Fig. 1 (top row), and from those maps it can be seen that the CGM gas is indeed mostly constrained to galaxies (in red). In the phase diagram we find that most of the CGM gas is cooled down to the temperature floor of $\sim 10^4$ Kelvin where radiative cooling basically stops (metals can cool gas further but we don't include those). Although CGM gas in our simulations is not directly affected by the

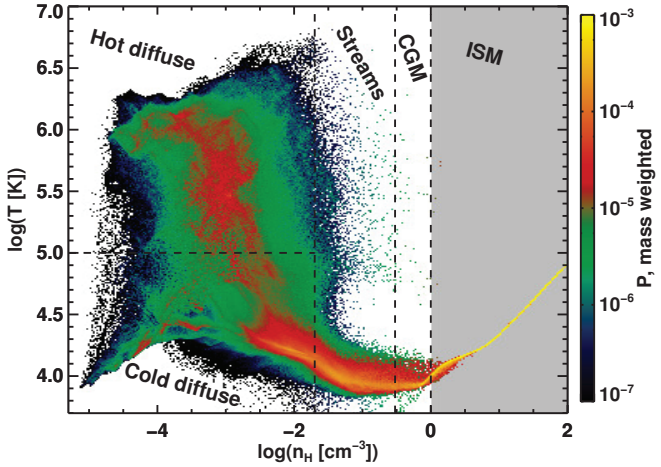


Figure 2. Phase diagram of the H1 halo, showing our density-dependent definitions of streams, CGM and ISM. The ISM is shaded to indicate that we always ignore ISM gas when adding up Lyα emissivities. The color scale represents mass weighted probability per temperature-density bin.

polytropic equation of state, one may expect gas at densities $\gtrsim 0.1$ cm^{-3} to be multiphase and star-forming (Schaye 2004). This cannot happen in our simulations because they don't describe cooling below $\sim 10^4$ K, and this temperature floor provides artificial pressure support for dense gas. This implies a potentially high error in our predictions for the Lyα luminosities of halos, resulting from an overestimated CGM contribution. Thus, while we in general include the CGM gas in our analysis of Lyα luminosities, we also consider at some points the effect of excluding it, to get a grip on how sensitive our results are to the density thresholds applied. In summary we find that GCM gas typically provides a 40% of the Lyα luminosities of our simulated halos, but that in terms of Lyα extent it is less substantial.

Streams are defined in this work as gas with densities between 0.02 and 0.3 cm^{-3}. These limits correspond to the density contours in Fig. 1 (top row) and from those we can see that these densities indeed correspond to thin filamentary structures. Much like the CGM most of the stream gas is found at the bottom of the temperature curve at $\sim 10^4$ K though we do see an increase in temperature in the more diffuse stream gas due to a combination of photo/gravitational heating and inefficient cooling (because of the low densities). Gas at sub-stream densities turns out to be negligible in terms of Lyα emissivity and thus not very important to our results so we crudely split what remains into two categories.

Hot diffuse gas has been shock heated above 10^5 Kelvin. As seen in the temperature maps in Fig. 1 (bottom row) this gas exists in abundance within the virial radii of the halos, but there also seems to be weaker heating around the large-scale accretion streams (and actually not so weak in the large streams around the H3 halo). Shock heating gets decidedly stronger with increasing halo mass, with gas reaching $\sim 2\ 10^7$ K in H2 and $\sim 6\ 10^7$ K in H3. Also, increasingly dense gas exists above 10^5 K in the more massive halos; CGM in H2 and ISM in H3.

Cold diffuse gas is partly gas which is slowly condensing towards the streams and the CGM and partly cosmological gas that has not interacted with the halos at all and is being cooled down by the cosmological expansion.

Table 2. Halo sizes and mass budgets (% of mass within R_{vir}).

Halo	R_{vir}	DM	Stars	Gas			
				ISM	CGM	Streams	Hot
H1	46 kpc	82%	8%	10%			
				73%	3%	8%	8%
H2	98 kpc	81%	7%	12%			
				60%	9%	16%	13%
H3	158 kpc	82%	5%	13%			
				58%	6%	12%	23%

The sizes and mass budgets of our three targeted halos are listed in Table 2. Each of the halo masses consists of roughly 80% dark matter and 20% baryons. The stellar/gas ratio decreases with halo mass, going from roughly one-to-one in H1 to about a one-to-three in H3. The gas mass is primarily in the ISM, going from 73% in H1 to 58% in H3. The hot gas fraction clearly increases with halo mass, going from 8% in H1 to 23% in H3 and correspondingly the cold fraction decreases, going from 8% to less than 1%. Interestingly the stream fraction peaks in the intermediate mass halo at 16%, with half and two-thirds of that in H1 and H3 respectively. The low fraction in H1 can be explained by the smooth accretion that efficiently moves the gas straight into the ISM (hence a high ISM fraction), whereas H3 streams are disrupted to the point of obliteration when they approach the halo center (hence the low ISM fraction and large fraction of hot gas).

3.2 On-the-fly self-shielding

The transfer of UV photons gives us the opportunity to study the extent of self-shielding in gas clumps and streams. We quantify the local UV field intensity in terms of the hydrogen photoionization rate Γ, which expresses the average number of photoionization events per hydrogen atom per unit time (see Appendix B). In the UV model we use, the unattenuated photoionization rate at redshift 3 is $\Gamma = 6.1\ 10^{-13}$ s^{-1} (see Fig. A2), and shielded regions should have $\Gamma \to 0$ s^{-1}.

Fig. 3 shows the UV attenuation in the three targeted halos at redshift 3. The top row contains non-logarithmic maps of projected *minima* of the photoionization rate along the LOS. The light color on the edges of the maps corresponds to the unattenuated value. Towards the centers of the halos the UV field becomes increasingly attenuated due to photo-absorption of the gas and in the densest streams and clumps we see $\sim 100\%$ attenuation. The diffuse streams, with densities $\lesssim 0.02$ cm^{-3}, are not self-shielded. Gas at the centers of the H1 and H2 halos is efficiently shielded but at the center of the H3 halo gas is thermally ionized and thus optically thin.

The bottom row of Fig. 3 shows logarithmic phase diagrams of the hydrogen photoionization rate Γ versus density for the same three halos. The most diffuse gas is UV emitting and has corresponding horizontal lines in the diagrams up to the n_H^{UV}-threshold. Above this threshold there is an immediate spread in the photoionization rate in all

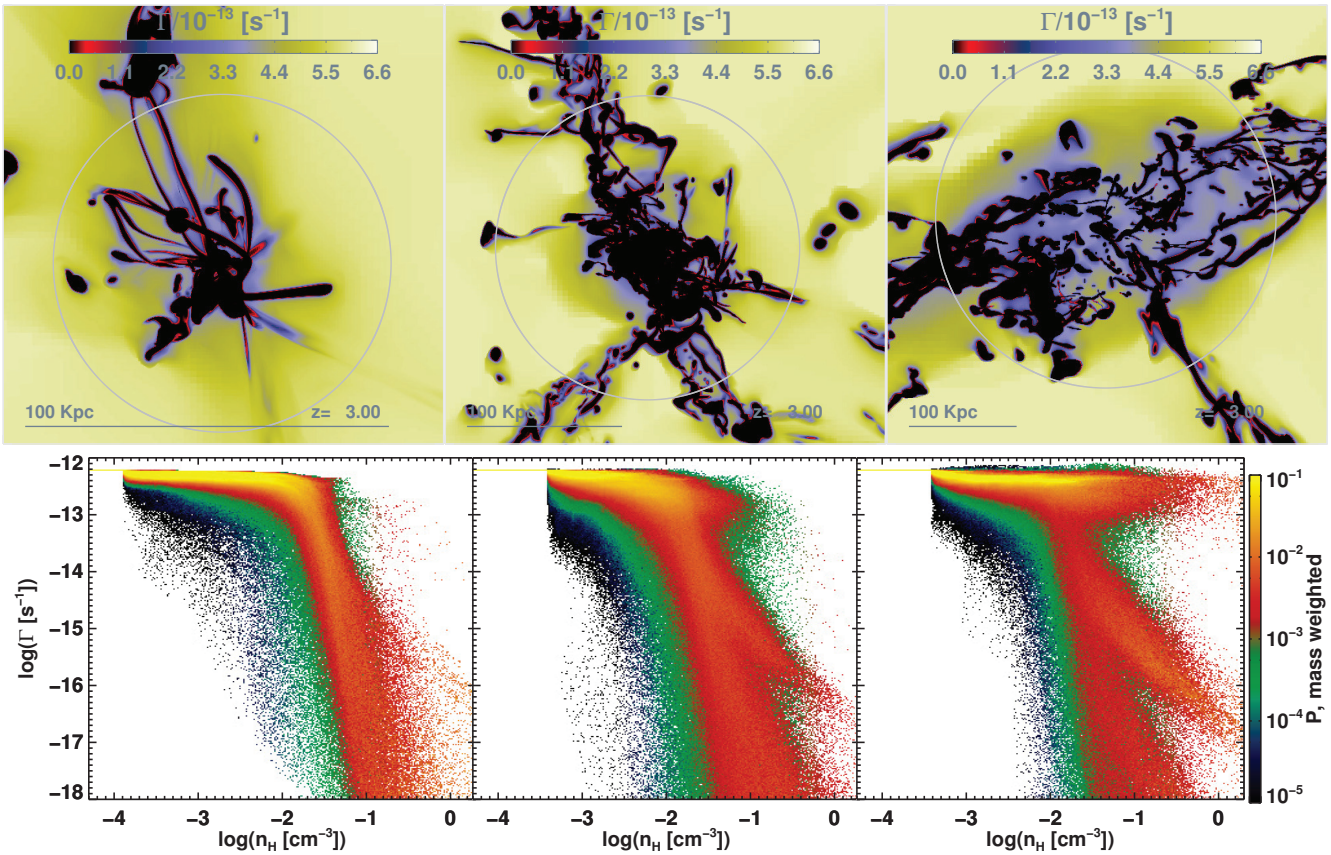


Figure 3. Self-shielding at redshift 3 in the three halos, from left to right, H1, H2 and H3. **Top row:** Maps of projected minima along the LOS of the hydrogen photoionization rate. The scale is non-logarithmic and in units of 10^{-13} s^{-1}. **Bottom row**: Phase diagrams of hydrogen photoionization rate versus density. The color scale represents mass-weighted probability per $\Gamma - n_{\rm H}$ bin over the plotted Γ range.

three halos, ranging from unattenuated UV to about half attenuated. Gas in the H1 halo is mostly self-shielded at $n_H \gtrsim 2\ 10^{-2}$ cm^{-3}. In more massive halos, the advent of thermal ionization in dense gas makes the situation more complex, and gas at $n_H \gtrsim 2\ 10^{-2}$ cm^{-3} exists in two phases, either self-shielded as in H1 or optically thin. The bifurcation in the diagram around $\Gamma \sim 10^{-17} - 10^{-15}$ s^{-1}, $n_H \sim 0.1 - 1$ cm^{-3}, which grows more conspicuous with increasing halo mass is an effect of the ionization fronts becoming under-resolved at high densities, where the mean free path becomes comparable or shorter than the cell sizes. This feature does not affect our results.

Fig. 4 shows maps of HI column densities and phase diagrams of neutral fraction $x_{\rm HI} \equiv n_{\rm HI}/n_{\rm H}$ vs. density. We find the CGM and ISM regions correspond mostly to damped Lyα absorbers (DLAs, $N_{\rm H} > 2\ 10^{20}$ cm^{-2}) and the streams to Lyman limit systems (LLSs, $N_{\rm H} = 1.6\ 10^{17} - 10^{19}$ cm^{-2}) and even super Lyman limit systems (SLLSs, $N_{\rm H} = 10^{19} - 2\ 10^{20}$ cm^{-2}), according to the definitions found in e.g. Fumagalli et al. (2011). The column densities are likely overestimated where they are highest due to lack of locally enhanced UV from star-formation. We see in the phase diagrams an abrupt transition of cold gas from ionized to neutral states, at about $5\ 10^{-2}$ cm^{-3} in all halos. This generic result is consistent with early expectations from Schaye (2001) and recent numerical estimates (e.g. Kollmeier et al. 2010; Faucher-Giguère et al. 2010; Aubert & Teyssier 2010). The dense ($n_H \sim 0.1$ cm^{-3}) and ionized ($x_{\rm HI} \lesssim 10^{-4}$) cells which become increasingly abundant with halo mass correspond to hot shock-heated gas, which is thermally ionized and optically thin.

3.3 A self-shielding approximation

FG10 applied a self-shielding approximation in their simulations, where a UV field is applied homogeneously to gas but with a cutoff at an assumed self-shielding density threshold of 10^{-2} cm^{-3}. We have run an analogue to our H1 simulation using the same self-shielding approximation instead of radiative transfer. Fig. 5 shows the neutral fraction versus density phase diagram at $z = 3$ in this simulation. Apart from a much more discrete jump from ionized to neutral, the diagram is similar to the RHD counterpart (Fig. 4, bottom left), and we find 50% neutral fraction at half the density of the RHD counterpart, or at 0.025 cm^{-3}. In terms of getting right the ionization state of gas at redshift 3 it thus appears that this non-RT self-shielding approximation holds fairly well. One might perhaps consider moving the self-shielding threshold a factor of two towards higher density, but one should be careful not to move it higher than that to avoid over-predicting Lyα emissivities due to photo-heating and photo-fluorescence. The approximation inaccurately de-

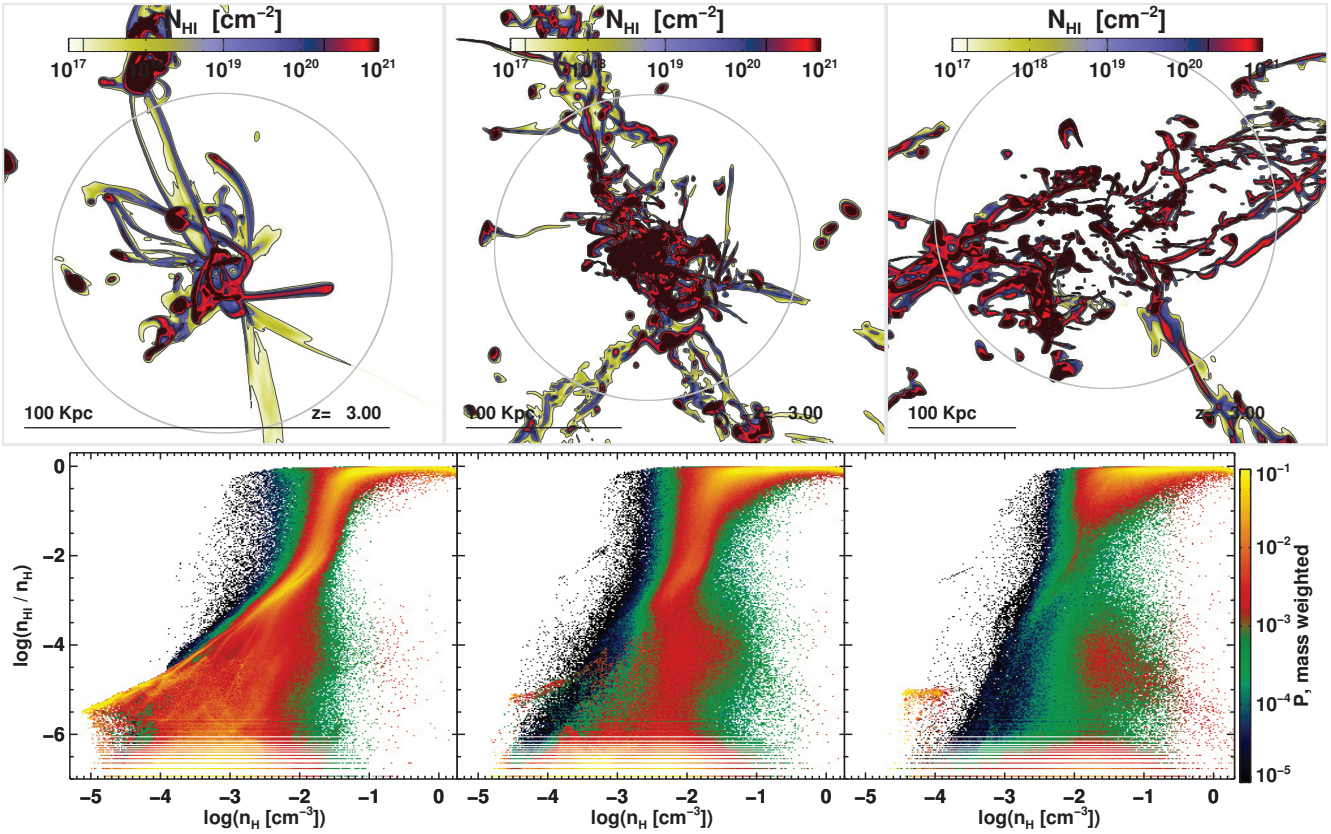


Figure 4. Neutral hydrogen at redshift 3 (H1, H2, H3 halos from left to right). **Top row:** Maps of projected HI column density. Contours correspond to the lower limits for DLAs ($N_{\rm H} = 2\ 10^{20}$ cm^{-2}), SLLSs (10^{19} cm^{-2}) and LLSs ($1.6\ 10^{17}$ cm^{-2}). **Bottom row:** Phase diagrams of neutral hydrogen fraction, $x_{\rm HI} \equiv n_{\rm HI}/n_{\rm H}$, versus density. The color scale represents mass-weighted probability per $x_{\rm HI} - n_{\rm H}$ bin over the plotted $x_{\rm HI}$ range. The quantization-like horizontal lines at the bottom of all diagrams are due to numerical precision of the cell variable $x_{\rm HII}$ in `RamsesRT`, which is roughly seven digits.

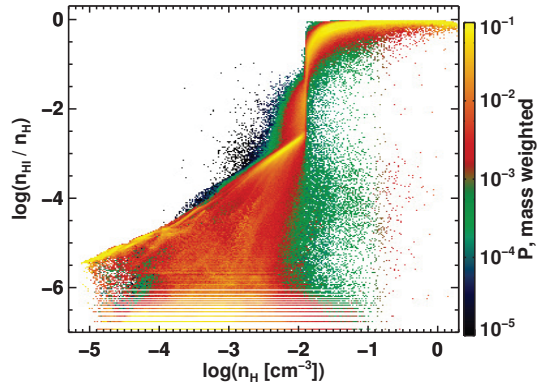


Figure 5. Phase diagram of neutral fraction versus density at redshift 3, in a version of the H1 simulation where a self-shielding approximation is used (instead of RHD) of applying full-strength UV background at densities below 10^{-2} cm^{-3} and zero strength above.

scribes UV attenuation in more massive halos, where much of the gas is thermally ionized and thus UV transparent. This doesn't matter however, since an absence/presence of the UV background in gas which is already so ionized has a negligible effect on its Lyα emissivity (which is dictated by collisional ionization equilibrium (CIE) heating/cooling.

4 PREDICTED Lyα LUMINOSITIES

4.1 Computing the gas Lyα emission

In most astrophysical contexts, an electron in the excited level 2P of the hydrogen atom will practically instantly relax to the ground state (1S) via the emission of a Lyα photon. There are two channels to produce such excited atoms, and hence to produce Lyα radiation[2]:

Collisional: A collision with a free electron excites the H-atom, which may release a Lyα photon when it relaxes back to the ground state. The collisional emissivity is approximated with

[2] To be exhaustive, there is a third one, which is absorption of photons with energies in the range 10.2 - 13.6 eV, which will excite the electron to any level $\geqslant 2$, which will in turn cascade down and sometimes produce a Lyα photon. This process is likely sub-dominant in the regime that we are investigating (Furlanetto et al. 2005; Kollmeier et al. 2010), and requires Lyα radiative transfer, which we postpone to a future paper.

$$\varepsilon_{coll} = C_{\mathrm{Ly}\alpha}(T)\, n_{\mathrm{e}}\, n_{\mathrm{HI}}\, \epsilon_{\mathrm{Ly}\alpha}, \quad (4)$$

where n_{e} and n_{HI} are number densities of electrons and neutral hydrogen, respectively, and $C_{\mathrm{Ly}\alpha}(T)$ is the rate of collisionally induced 1S-to-2P level transitions. An expression for this rate is given by G10, fitting results from Callaway et al. (1987). It is always less than the hydrogen collisional excitation cooling rate, Λ^{HI}_{coll}, used in the code (from Maselli et al. 2003), since cooling also takes into account excitations to atomic states other than 2P (the most likely of which is the non-Lyα releasing 2S state). The ratio of $C_{\mathrm{Ly}\alpha}/\Lambda^{HI}_{coll}$ goes from 71% at 10^4 K to 57% at 5 10^4 K.

Recombinative: A free electron combines with a proton at any level ($\geqslant 2$), and may cascade down to the 2P level. The recombinative Lyα emissivity of this process is given by

$$\varepsilon_{rec} = 0.68\, \alpha^{\mathrm{B}}_{\mathrm{HI}}(T)\, n_{\mathrm{e}}\, n_{\mathrm{HII}}\, \epsilon_{\mathrm{Ly}\alpha}, \quad (5)$$

where the 0.68-factor is the average number of Lyα photons produced per case B recombination (from Osterbrock & Ferland 2006) and $\alpha^{\mathrm{B}}_{\mathrm{HI}}(T)$ is the case B recombinations rate, i.e. counting recombinations to all levels except directly to the ground one. We use the expression from Hui & Gnedin (1997).

Unless otherwise specified, the Lyα emissivities calculated in this paper are:

$$\varepsilon = \varepsilon_{coll} + \varepsilon_{rec}. \quad (6)$$

Figure 6 shows the collisional and recombinative Lyα emissivities of gas at typical stream density, assuming the gas is UV exposed (thick curves) and self-shielded (thin). Also shown in dotted black curves is the neutral fraction of the gas, approximately extracted using a simplified model of hydrogen-only and PIE/CIE equilibrium.

The plot illustrates that it is crucial to be *consistent* in following the gas state of $(T, x_{\mathrm{HI}}, \Gamma)$ in the simulation code, since independently changing one of those factors without considering the effect on the others can have a dramatic effect on ε. If, for example, self-shielding is assumed in post-processing and the neutral fraction changed accordingly without considering the change in temperature, the ε_{coll} estimate can increase by almost an order of magnitude. This is unphysical – what really happens if gas suddenly becomes UV shielded is that the temperature drops somewhat due to lack of photo-heating, the end-result being a slightly lowered value of ε.

The accuracy of the $(T, x_{\mathrm{HI}}, \Gamma)$-state is secondary to consistency, because if the code handles things properly, ε should simply reflect the work put into the gas by the UV background and gas advection. In the limit that the UV energy input is negligible compared to gravitational heating, accurate modelling of the UV background isn't really crucial in the context of Lyα emissivity, and applying e.g. a sensible shielding approximation like the one discussed in Sec. 3.3 should be OK. This breaks down when UV photofluorescence becomes non-negligible.

4.2 Intrinsic luminosities

Fig. 7, top row, shows maps of the rest-frame Lyα surface brightness S of the three targeted halos, which is calculated by integrating the Lyα emissivity (Eq. 6) along the LOS. We don't take absorption or scattering into account: These factors would certainly diminish the brightest spots associated with CGM regions, but we don't expect them to affect the more diffuse streams much (see discussion in Sec. 4.3). The surface brightness is concentrated around CGM regions in all three halos, with $S \approx 10^{40} - 10^{41}$ erg s^{-1} kpc^{-2}, and the brightness in streams is typically lower by one or two orders of magnitude.

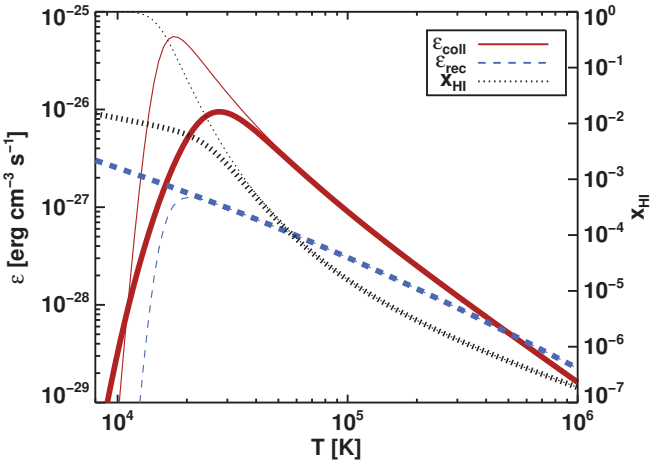


Figure 6. Lyα emissivity of gas at number density $n_{\mathrm{H}} = 3\ 10^{-2}$ cm^{-3}, which is close to the lower limit for accretion streams in our simulations. Thick curves show gas exposed to a UV field with $\Gamma = 6.1\ 10^{-13}$ s^{-1}, corresponding to redshift 3. Thin curves show UV shielded gas ($\Gamma = 0\ s^{-1}$). Blue dashed curves show recombinative Lyα emissivity, red solid curves show collisional Lyα emissivity and black dotted curves show the approximate neutral hydrogen fraction in the gas, assuming equilibrium between photoionization, collisional ionization and recombinations.

The bottom row in Fig. 7 shows radially cumulative Lyα luminosities for the halos, i.e. fraction of the total luminosity within a given radius (black solid curves). Streams and the diffuse medium consistently contribute about 60% of the total luminosity, as indicated by the thin black curves.

It is evident from both the maps and plots that there is a trend of more extended emission with increasing halo mass. In the H1 halo, half of the total luminosity comes from the central 16% of the virial radius (dotted lines), while in H2 this radius is 20% and 33% in H3. Partly this is because the more massive halos consist of increasing quantities of orbiting galaxies so the surface brightness is just following the increased spread of CGM regions, as can be seen from red dots of surface emissivity in the maps and from corresponding steps in cumulative surface brightness in the plots. That is not the whole story though: The streams become more efficient Lyα emitters with increasing halo mass.

As seen from the blue curves in the luminosity plots, electron-hydrogen collisions dominate the total luminosity, and recombinations are borderline negligible, as should be expected outside ISM regions. This dominance increases with halo mass, with recombinations contributing 10% to the total in the H1 halo and only about 5% in H2 and H3. The red curves in Fig. 7 will be discussed in Sec. 5.

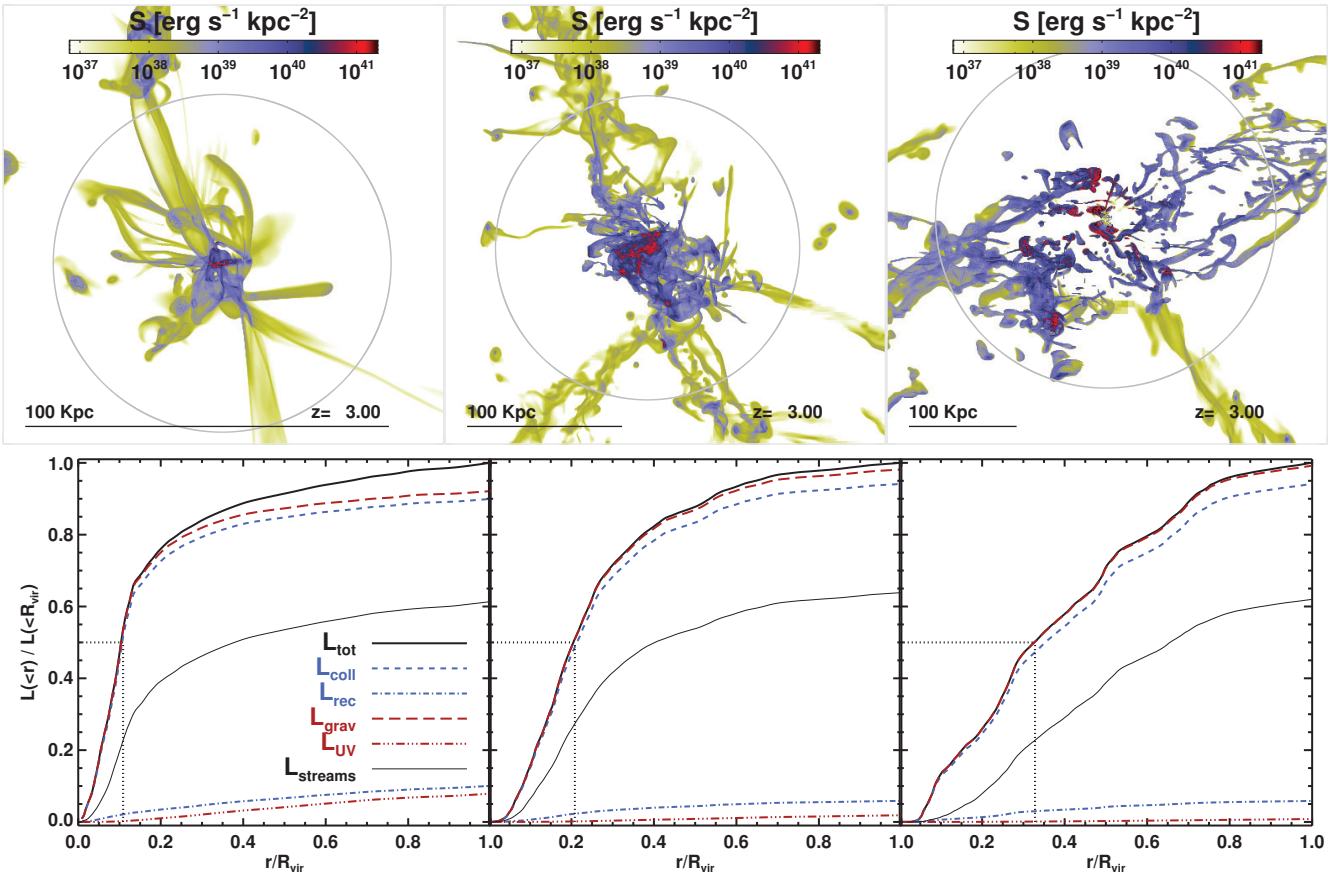


Figure 7. Top row: Rest-frame Lyα surface brightness maps of the targeted halos (H1, H2 and H3 simulations, from left to right). **Bottom row:** Radially (3D) cumulative Lyα luminosities of the same halos. The thick black curves show the total luminosity. The dotted lines denote at what radius 50% of the total luminosity is accounted for. The blue curves show how the total luminosity is split between collisional and recombinative channels, given in Eqs. 4 and 5 respectively. The red curves show how the total is split between the contributions of the UV background and gravitational dissipation, discussed in Sec. 5.2. The thin black curves show the contribution of sub-CGM density gas (i.e. mostly streams) to the total luminosities.

In Fig. 8 we plot total halo luminosities versus halo mass (which is defined, as in Table 1, as the total mass of all dark matter and baryons within the virial radius). From each of the three simulations we extract all halos from within the zoom-in volume and integrate Eq. 6 over their virial radii, excluding ISM gas. The halos roughly line up into a power law indicated by a red solid line, with exponent 1.25. There is a systematic tendency for halos in more massive simulations to be more luminous for a given mass, which is presumably an environment effect since the cosmological over-density of the zoom-in regions increases between the H1, H2 and H3 simulations respectively.

Large thick symbols mark the three main halos targeted in our simulations. For those we also plot luminosities excluding consecutive phases of gas. Blue symbols show luminosities when ignoring the ISM and the CGM, and green symbols show what happens if we also ignore the stream densities. The CGM accounts for about 40% of the total luminosity in all three targeted halos (see also Fig. 7) and the streams account for most of the rest, or $50-60$%, with sub-stream densities accounting for 8%, 4% and 2% in the targeted halos of H1, H2 and H3 respectively.

As can be read directly from Eqs. (4) and (5), the Lyα emissivity of gas in principle scales with density squared, though temperature and ionization state have their influence as well. Fig. 9 shows a luminosity weighted phase diagram of Lyα emissivity of gas versus density in the H1 halo (the H2 and H3 diagrams are similar). Over-plotted on the diagram in dashed grey lines are two power laws that the gas emissivity approximately follows, with a knee between 10^{-2} and 10^{-1} cm^{-3}. The knee roughly corresponds to where the gas becomes self-shielding and the change in slope is caused by the corresponding transition in temperature and ionization state. Below the knee the gas emissivity is split in two ridges with slightly different slopes. The upper one has power index ≈ 2.5 and is dominated by collisional emission (Eq. 4), whereas the lower one has power index ≈ 2.2 and is dominated by recombinations (Eq. 5). The emissivity above the knee is completely dominated by collisions. The stronger than 2 power law below the knee stems from the increasing abundance of neutral atoms with density and a temperature that tends towards peak Lyα emissivity, whereas the less than 2 power law above it results from the decreasing relative abundance of electrons with density.

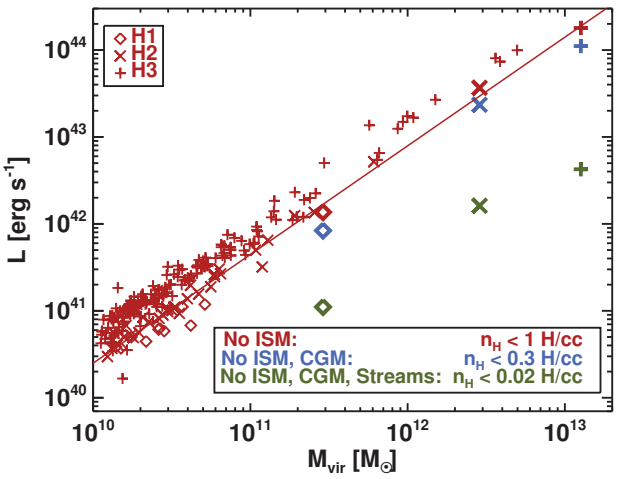


Figure 8. Total halo luminosities versus halo mass for the three simulations. The halos are extracted from the zoom-in volumes, and we exclude sub-halos. The three targeted halos are indicated by large thick symbols. For those halos we also show luminosities excluding different phases of the gas. The red solid line indicates a power-law with exponent 1.25.

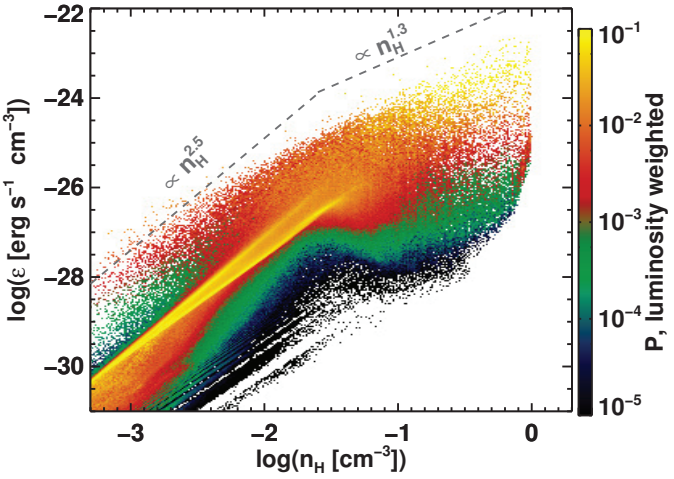


Figure 9. Phase diagram showing Lyα emissivity of gas versus density at z=3 in the H1 halo. Each density bin is Lyα luminosity weighted independently. The grey dashed lines show power laws that approximately fit the data. The color scale represents Lyα luminosity weighted probability per $\varepsilon - n_{\rm H}$ bin over the plotted ε range.

4.3 Observational properties

We now consider mock observations of our simulated halos. To produce those we first convert the rest-frame surface brightness maps in Fig. 7 to *observed* surface brightness I, using

$$I = \frac{S f_\alpha}{4\pi(1+z)^4}, \qquad (7)$$

where f_α is a cosmological transmission factor that accounts for absorption and scattering of Lyα photons on the LOS from the object to the observer. We adopt in this paper a value of $f_\alpha = 0.66$ based on the work of Faucher-Giguère et al. (2008). (f_α is only applied to mock observations and not to the intrinsic emissivity and luminosity, Figures 7 and 8). To the result of Eq. 7, we then apply a Gaussian point spread function (PSF) with a 0.6 arcsec full width at half maximum (FWHM) to mimic atmospheric and instrumental distortion, and assume a camera pixel size of 0.2 arcsec (Fig. 10). This corresponds to very good seeing conditions in state-of-the-art instruments. We present maps made with a PSF about twice as broad in Appendix C. These are directly comparable to the observations of M11.

Unlike FG10, we don't model the scattering of Lyα photons in this work. These authors show that Lyα transfer dominates the spectral shape of extended Lyα emission, but hint that it has little effect on the morphology and extent. Their Fig. 8 shows this to be the case for a halo corresponding in mass to our H1 halo, if only the Lyα emissivity of gas is considered – though their Fig. 9 also illustrates that strong point-like sources can produce extended Lyα structures via scattering. We will assume here that scattering has little overall effect on our predicted morphologies and extents, in the case that these structures are already well extended, though we do expect that it will likely produce subtle changes in observable LAB areas – indeed Fig. 8 in FG10 shows that the inclusion of scattering can make some observable Lyα structures narrow down and others widen out. We will include and investigate the effect of Lyα scattering on spectral shapes, luminosities, and morphologies of our objects in a future paper.

4.3.1 Observed blobiness of cold accretion streams

Mock observations of our three targeted halos are shown in the top row of Fig. 10. The middle contour in the maps is set at $I_{-18} = 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2}, roughly corresponding to current observation limits (e.g. M11, Erb et al. 2011, see Appendix C for a more accurate comparison), and the inner and outer contours correspond to ten times brighter and ten times dimmer, i.e. 10^{-17} and 10^{-19} erg s^{-1} cm^{-2} arcsec^{-2}.

Assuming I_{-18} as our instrumental sensitivity limit, the H1 halo (top left) is a Lyα emitter that is centered on a galaxy, circularly symmetric in shape, about 20 kpc in diameter and doesn't trace streams. Thus the H1 halo is not a LAB. The total observed luminosity, i.e. I integrated over the area within I_{-18}, is $L_{obs} = 6\ 10^{41}$ erg s^{-1}.

The H2 halo observation (top middle) differs dramatically from that of H1. At I_{-18} we do see a borderline giant LAB, asymmetric and about 100 kpc in length, and we can see the end of an accretion stream poking out to the north-west. The observed luminosity integrated above I_{-18} is $L_{obs} = 2\ 10^{43}$ erg s^{-1}. The H3 halo (top right) has observable Lyα emission all over the place, is about 200 kpc in diameter and very asymmetric. Its observable luminosity is $L_{obs} = 10^{44}$ erg s^{-1}.

Provided there is nothing special about these halos, we can conclude that in general the *cooling emission from halos with masses greater than a few times* $10^{12}\,M_\odot$ *can produce giant LABs* ($\gtrsim 100$ *kpc) at redshift 3, assuming current instrument sensitivity limits.* Qualitatively this compares well with Yang et al. (2010), who find that at redshift 2.3, LABs should occupy halos $\gtrsim 10^{13}\,M_\odot$. Qualitatively again, the

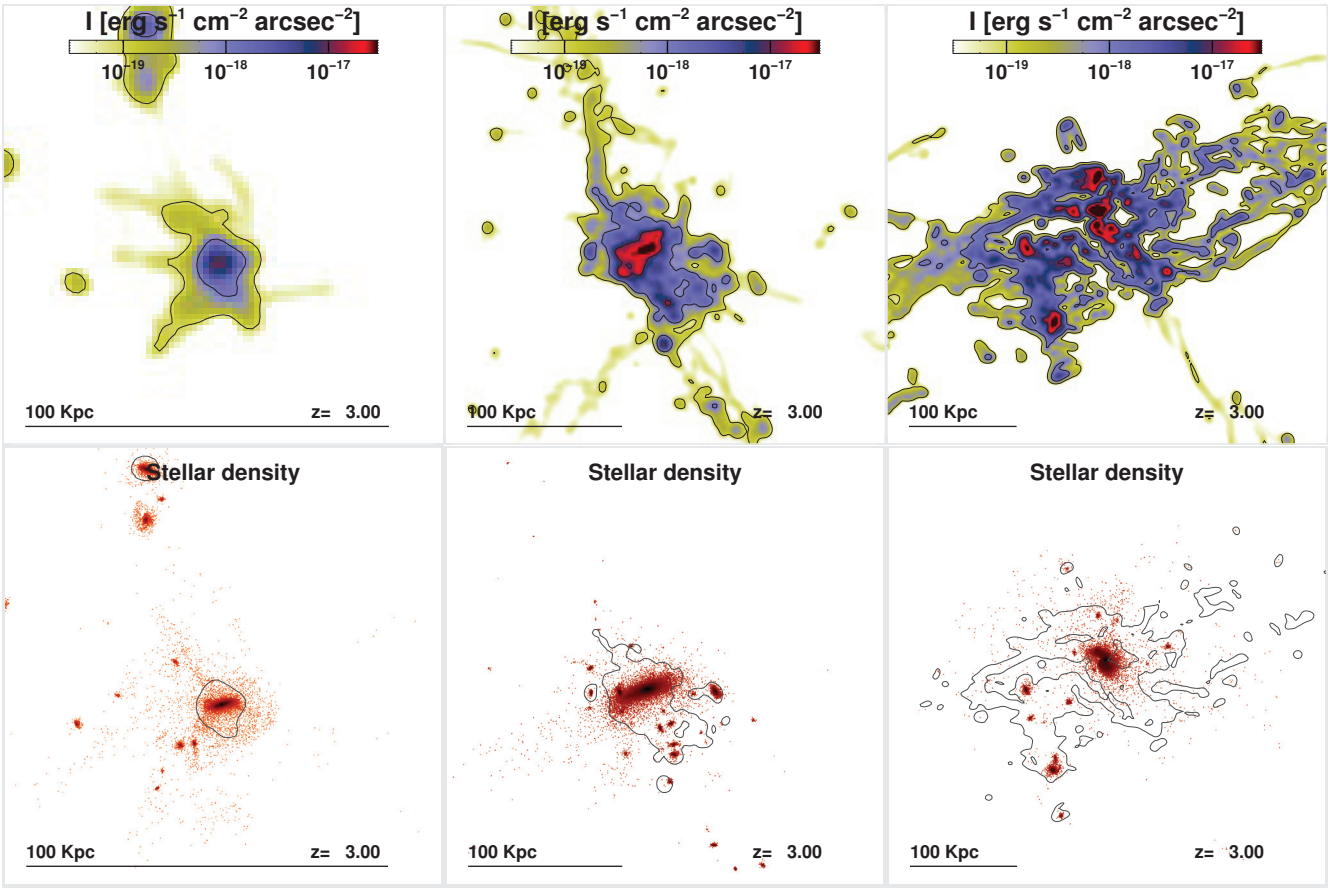


Figure 10. Top row: Mock images showing predicted observed surface brightness in the targeted halos of H1, H2 and H3, from left to right. The contours mark 10^{-17}, 10^{-18} and 10^{-19} erg s^{-1} cm^{-2} arcsec^{-2}. The images where computed using an optimistic PSF of FWHM 0.6 arcsec. **Bottom row:** Stellar density maps for the same halos, illustrating that the bright spots of Lyα emission are centered on galaxies. Over-plotted are contours marking 10^{-18} erg s^{-1} cm^{-2} arcsec^{-2} in observed surface brightness (the middle contours from the upper maps).

maps presented in Sec. C, which mimic the observational conditions of M11, show that the morphologies of our simulated LABs are very similar to those observed.

Interestingly, we note that the LABs produced by cold accretion streams are naturally extended in the direction of the main large-scale filaments that they are connected to. This is particularly visible for H2 and H3 in Fig. 10 (see also Fig. 14), and lends support to the observational findings of Erb et al. (2011).

Another matter are those mysterious LABs which do not seem to be centered on observed galactic counterparts (e.g. Steidel et al. 2000; Weijmans et al. 2010; Prescott et al. 2011). We are not able to reproduce this phenomenon in our simulations. The bottom row of Fig. 10 shows stellar densities in our targeted halos, with the I_{-18} sensitivity contour over-plotted. Clearly all the peaks of Lyα brightness would have continuum counterparts in observations, unless these counterparts would for some reason be hidden from view. Such LABs are rare among rare events, though, and our three simulations have little statistical chance of reproducing such oddities. A larger sample of simulations would be required to investigate this issue further.

4.3.2 Size distribution of simulated LABs

We shall now statistically compare our results with a catalogue of 202 observed LABs from the surveys described in M11 (courtesy of Yuichi Matsuda and team). The aim here is to derive a cumulative LAB area function from our results and see how it compares with real data.

We follow M11 by assuming $z = 3.1$ in Eq. 7, and applying a PSF with FWHM=1.4 arcsec. We calculate the observed LAB area A of each halo within the zoom regions of our simulations by integrating its total area above the surface brightness limit $I = 1.4\ 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2}. We 'observe' each halo in three directions (x, y, and z). In Fig. 11, we plot the LAB areas as a function of halo mass. The large thick symbols correspond to our targeted halos H1, H2, and H3. The observed LAB area is a reasonably well-behaved function of halo mass, with more massive halos producing larger LABs, and the points can be bracketed by a couple of power laws of indexes 0.87 and 1.19 (see Fig. 11).

We now make the assumption that extended Lyα emission is an inherent property of dark matter halos and that the observed LAB properties are direct functions of halo mass. This assumption is substantiated by our results (Figs. 8 and 11). We thus convolve the power laws of Fig. 11 with

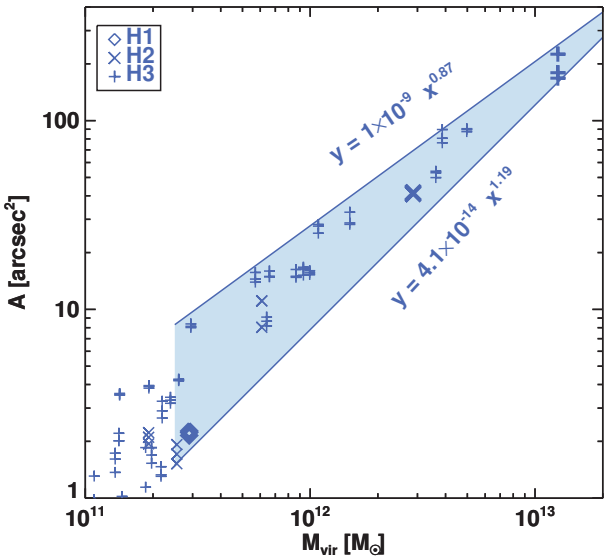


Figure 11. Plot of mock observed areas within contours of 1.4 10^{-18} erg s^{-1} cm^{-2} arcsec^{-2} versus halo mass for zoomed halos in the three simulations. Each halo is observed from three directions. The targeted halos are marked with thick symbols. The points are enveloped by two power-laws shown in the plot.

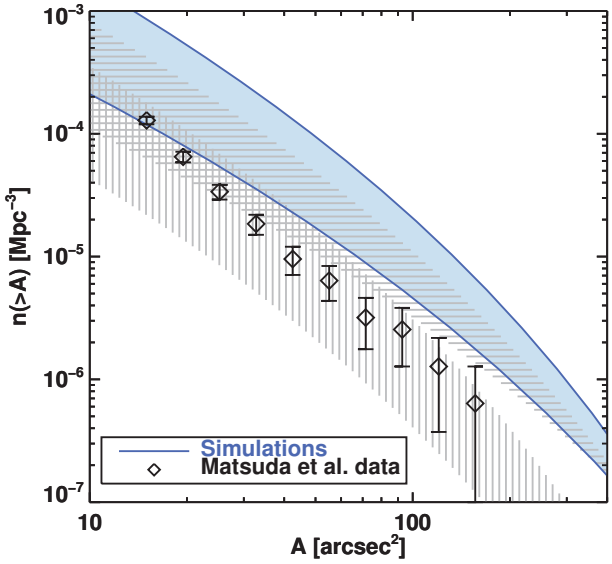


Figure 12. The shaded region represents the boundaries of our predicted LAB area function, derived from the power laws in Fig. 11. The black symbols mark a rough area function derived from a sample of 202 observed LABs from the survey of M11. The horizontally and vertically line-filled regions represent similarly predicted area functions, but with gas densities of $n_{\rm H} \geqslant 0.3$ cm^{-3} and $n_{\rm H} \geqslant 0.1$ cm^{-3} excluded, respectively.

the halo mass function at redshift 3.1 (taken from Sheth & Tormen 1999) in order to produce the cumulative area function envelope shown in Fig. 12. There, the black diamonds represent actual observations for comparison: they are a rough estimate of the area function at redshift 3.1 based on the 202 LABs of the M11 survey, derived by binning the LABs by area and dividing the count by the total survey volume of 1.57 10^6 Mpc^3 (the error bars are Poissonian). The comparison between our predicted area function and the observationally derived one is very satisfactory, although we systematically over-predict the function by a factor of 2-3.

There may be several causes to this over-prediction. First, our derivation of the observed LAB area function is too simplified. For example, we do not take into account the shape of the narrow band filter, or any $1/V_{\rm max}$ corrections. This introduces systematic errors that could well be of about a factor two. Second, our mock observations are also simplified, and do not include noise, which could possibly affect the measured area in a systematic way. Third, perhaps we have overshot in our choice of $f_\alpha = 0.66$. As noted by G10, cosmic extinction may be stronger than average for sources that reside in over-dense regions, as LABs tend to do. Fourth, our prediction is based on only a few objects and to a lesser degree the same applies to the observation-derived function. Fifth, the predicted LAB areas are sensitive to the applied PSF smoothing, which may not be entirely consistent in all the 202 observed LABs. And finally, we may lack physics in our simulations that would drive down the LAB areas. For example, Lyα scattering, if applied, could induce a slight spread in the predicted rest-frame Lyα surface brightness, which could in some cases bring down both the observed area and luminosity within sensitivity ordained brightness contours. Also, metal-line cooling may drive down the Lyα emissivity of gas by cooling it below 10^4 K. Furthermore, as shown by van de Voort et al. (2011b), Faucher-Giguère et al. (2011), and van de Voort & Schaye (2011), feedback driven winds can destroy cold accretion streams in the vicinity of galaxies, hence terminating their Lyα emissivities.

Our predicted area function is not very sensitive to the density threshold of gas applied throughout this paper, where we have excluded ISM densities ($n_{\rm H} > 1$ cm^{-3}) in our analysis. To illustrate this, Fig. 12 also shows, with line filled regions, predicted area function envelopes that have been derived from our simulations via a convolution with the Sheth-Thormen halo mass function, but including only more diffuse gas, $n_{\rm H} < 0.3$ cm^{-3} (i.e. sub CGM densities) and $n_{\rm H} < 0.1$ cm^{-3}, for the horizontal and vertical line-fillings respectively. The prediction using the sub-CGM densities is close to the original prediction, which can be expected since these densities account for $\sim 60\%$ of the total luminosities of all three targeted halos (see Figs. 7 and 8). Even using $n_{\rm H} < 0.1$ cm^{-3} gas only (which is comparable to the more conservative prescriptions used in FG10) still produces giant LABs hosted by massive halos and gives an area function that is compatible to the observational data. This confirms that the extent of our simulated LABs is largely driven by low density cold streams.

We have also compared our results to observations via a LAB luminosity function (rather than the area function just discussed). However, since LAB emissivity typically peaks around compact sources, and since we neither model the emission nor absorption coming from the compact ISM regions, such a comparison is less robust than using the area function which should be more or less dictated by the state of more diffuse gas on much larger scales. The luminosity

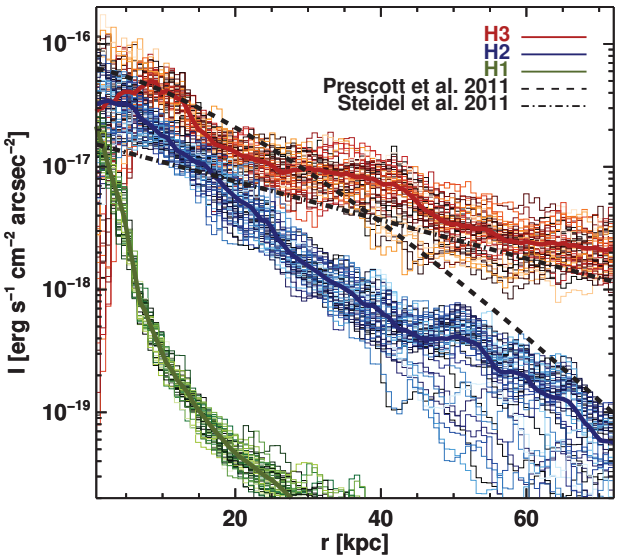


Figure 13. Surface brightness profiles of our three targeted halos (solid) compared to the profile of LABd05 from Prescott et al. (2011) (dashed) and an average surface brightness profile of 11 LABs from Steidel et al. (2011) (dot-dashed). The thin solid lines represent different random orientations (50 for each halo) and the thick solid ones are averages of the thin ones.

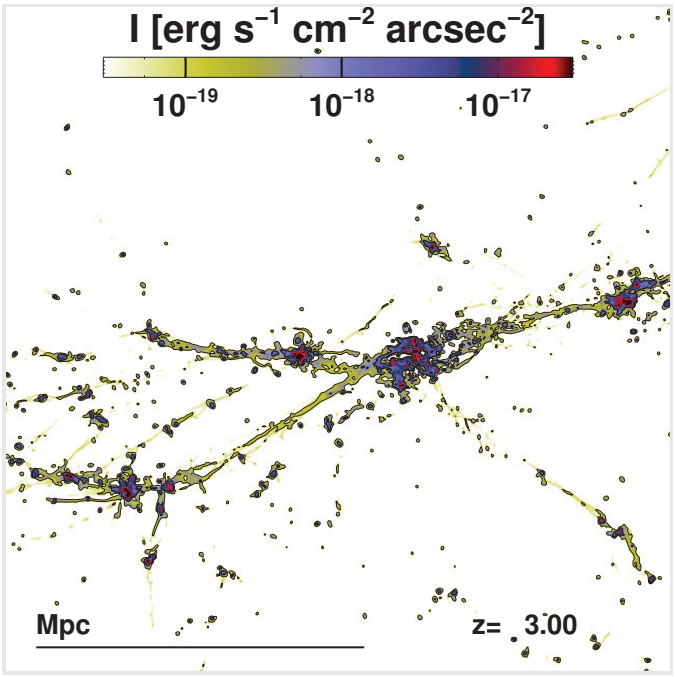


Figure 14. Same as Fig. 10, but showing the large-scale Lyα map of the H3 halo and its environment. Thick inner (thin outer) contours mark $I = 10^{-18}$ (10^{-19}) erg s^{-1} cm^{-2} arcsec^{-2}. Large-scale streams connecting massive halos and extending over several Mpc would be visible at 10^{-19} erg s^{-1} cm^{-2} arcsec^{-2}.

comparison, which is discussed in detail in Appendix D, is actually surprisingly good, but it is problematic to draw any conclusions from it because of the lack of modelling of compact regions.

4.3.3 Surface brightness profiles

Fig. 13 shows observed surface brightness profiles of our three targeted halos. Profiles are taken for each halo in 50 planes of random orientation, represented by thin coloured lines, and then these are averaged into the thick coloured lines. The profiles are transformed from rest-frame to observed surface brightness via Eq. 7, assuming redshift 3, but no smoothing or cosmic extinction is applied, and as before Lyα scattering is neglected. Over-plotted are surface brightness profiles from observations: The black dashed curve represents a Sersic fit to the observed profile of the giant LABd05 at redshift 2.656 from Prescott et al. (2011), which we have scaled to z=3. Notably, LABd05 doesn't have a galactic counterpart at, or even close to the peak of Lyα emission, though it has 17 small galaxies substantially offset from the peak (by $\gtrsim 20$ kpc). The black dot-dashed curve is an exponential disk fit to an average of 11 LAB profiles observed at $z \approx 2-3$, reported in Steidel et al. (2011), with no scaling applied.

The profiles of H2 and H3 are similar in shape and magnitude to the observed profiles. Interestingly, each of those compares favourably to different observations, with the H2 profile being similar to LABd05 and the H3 profile similar to the 11 LABs from Steidel et al. (2011). The comparison indicates that these observations fit well within the model of cold accretion powered LABs, but due to the very limited statistics of our simulations (i.e. one halo per mass bin of three), and different redshifts of the observed LABs, it is problematic to make quantitative deductions, e.g. about masses of the host halos of observed LABs. Rather than representing different halo masses, the different profile shapes (and to some degree their magnitudes) may just as well reflect the different morphologies one may find in galactic groups and clusters.

Prescott et al. (2011) compare the LABd05 surface brightness profile with simulated profiles from G10 and FG10, and find that the simulations appear to fit very badly with reality, with the G10 profile being both too peaky at the center and too shallow at large radius, and the FG10 profile being too weak and steep. Our H1 profile actually agrees with the simulated profile from FG10 (their model 7, see Fig. 9 in Prescott et al. 2011), and it seems to us that FG10 in fact don't pose any mismatch with the LABd05 observation: The fault lies in Prescott et al. (2011) assuming that the surface brightness profile scales linearly with halo mass, which is not at all the case judging from our simulations (and to be fair, these authors admit that their assumption is probably not accurate).

We admittedly don't provide large statistics here, but we can conclude that the surface brightness profiles produced by our simulations do not disagree with LAB observations, and at the same time we can argue that neither do the simulations of FG10.

4.3.4 Implications for future observations

Having demonstrated reasonable agreement between our simulations and LAB observations, we now wish to highlight a prediction from our work which is particularly relevant in

the context of direct searches for IGM emission at high redshifts. The outermost contours in the upper row of Fig. 10 mark Lyα brightness at 10^{-19} erg s^{-1} cm^{-2} arcsec^{-2}. At this limit, accretion streams start to show up even in the least massive halo, and in the more massive halos we would detect them unambiguously. The deepest observations to date are not quite there yet, but almost, and this is an exciting perspective. Perhaps even more exciting is the map shown on Fig. 14, where the thin (resp. thick) contours again mark the limit at 10^{-19} (resp. 10^{-18}) erg s^{-1} cm^{-2} arcsec^{-2}. This zoomed out view of our H3 halo shows that deep Lyα observations around massive halos may even reveal the large-scale filamentary structure of the IGM on scales of a few Mpc !

Although such observations are extremely challenging (if not plain impossible) today, upcoming instruments, such as MUSE (Bacon et al. 2006) or K-CWI (Martin et al. 2010), should greatly increase our chances of observing directly this source term of galaxy formation in a very near future.

5 WHAT DRIVES THE Lyα EMISSION?

Since Lyα scattering and stellar feedback are not included in our simulations, the only possible power sources of Lyα emission in our results are gravity and the UV background. We now look at how gravitational heating contributes to the Lyα emission along streams and attempt to quantify its efficiency. We consider the contribution of UV fluorescence and show how it is sub-dominant for typical values of the UV background. We conclude this section by discussing to what extent locally enhanced UV fluxes could boost Lyα emission from cold streams.

5.1 Gravitational efficiency

Gravitational heating is generally viewed as a progressive release of gravitational potential energy that heats the gas along the cold streams (Dijkstra & Loeb 2009). As long as this heating is not too fast it can be balanced by radiative cooling, and as long as the gas remains metal-poor and at temperatures $\sim 10^4$ K, Lyα emission is the dominant cooling mechanism, meaning that the thermal energy is efficiently converted into Lyα photons.

Gravitational heating in cold streams can be parametrized by the *gravitational efficiency* f_{grav}, the fraction of the change in gravitational potential that dissipates into thermal energy during in-fall. The rest is converted into bulk kinetic energy, increasing the speed of the gas. A value of $f_{grav} = 1$ thus means perfect conversion of potential into heating, implying constant in-fall speed, and $f_{grav} = 0$ means that there is no conversion into thermal energy and the stream should be in free-fall. Dijkstra & Loeb (2009) derive an analytic model in which $f_{grav} \gtrsim 20\%$ is required if Lyα blobs are to be driven by gravitational heating in cool flows.

For each stream we can distinguish in our simulated halos in a given output, we can extract the stream speed profile $v_{str}(r)$ by following its core from end to end. Using this and a corresponding free-fall profile $v_{ff}(r)$ for a body starting at a position and speed identical to the outer end of the stream, we can estimate f_{grav} with

$$f_{grav}(r) = \frac{v_{ff}^2(r) - v_{str}^2(r)}{v_{ff}^2(r) - v_{init}^2}, \quad (8)$$

where v_{init} is the speed at the outer starting position.

We calculate an approximate free-fall profile for the stream by assuming static state and spherical symmetry, integrating the free-fall speed from the starting position towards the halo center using

$$dv_{ff}(r) = \frac{1}{v_{ff}(r)} \frac{GM(<r)}{r^2}\, dr, \quad (9)$$

where r is radius, G is the gravitational constant and $M(<r)$ is the total halo mass within r.

In practice we divide the halo mass into radial bins r_i, where increasing i corresponds to decreasing radius, and solve Eq. 9 by recursively computing

$$v_{ff}(r_{i+1}) = v_{ff}(r_i) + \frac{1}{v(r_i)} \frac{GM(<r_{i+1})}{r_{i+1}^2}(r_{i+1} - r_i), \quad (10)$$

where $M(<r_{i+1})$ is the mass measured within r_{i+1} in the simulation, and the initial condition is the stream speed at the outer end, $v_{ff}(r_0) = v_{init}$.

In Fig. 15 we show phase diagrams for the three targeted halos of gas speed versus radius (normalized to R_{vir}), where we exclude all but gas at stream densities, so that the streams can stand out more clearly. In the smallest halo (H1) the streams pop out nicely, smooth and undisturbed basically over the whole radius range, though they do dilute a bit at the central 10% of R_{vir}. In H2 we can still see streams, but they are much more disrupted, and not distinguishable within the central 20% of R_{vir}. In H3 only a few streams can be distinguished in the outer 40% of R_{vir}, and in the central $\sim 20\%$ they are completely destroyed.

For those streams we can clearly distinguish in the diagrams, we have plotted in yellow the corresponding free-fall profiles, using Eq. 10, which show approximately the speeds that the streams would follow were they in free-fall. Qualitatively it can be seen that the streams are close to free-fall, though usually they lag a little behind the free-fall profile, and conversely on some occasions we even see streams that seem to accelerate faster than free-fall (due to sub-halos and/or the inaccuracy of assuming static state and spherical symmetry in our free-fall calculation).

We plot our estimates of f_{grav} using Eq. 8 directly above each phase diagram. For two of the three streams we have extracted in the H1 halo we get a fairly consistent estimate of $f_{grav} \sim 0.1$ from the halo outskirts towards the central $\sim 15\%$ of R_{vir}, whereas for the third (and more diffuse) stream we get a value which is two to three times higher. In the H2 halo things are much messier, and for those fragments of streams that we can extract we find a large scatter in f_{grav}, going from negative values to about 0.3 (the initial large values are a numerical noise due to resolution in the phase-space). Finally, in the H3 halo, we can only extract two streams at the outer edges of the halo, one of them showing $f_{grav} \sim 0.1$ and the other accelerating faster than our free-fall approximation.

It appears that gravitational heating is inefficient if seen only as a smooth and steady process along unperturbed streams as in H1. However, heating and subsequent release of Lyα photons seems to be more efficient when it involves disrupted and wiggly streams. This also appears reasonable, since the gas at the core of a straight and unperturbed

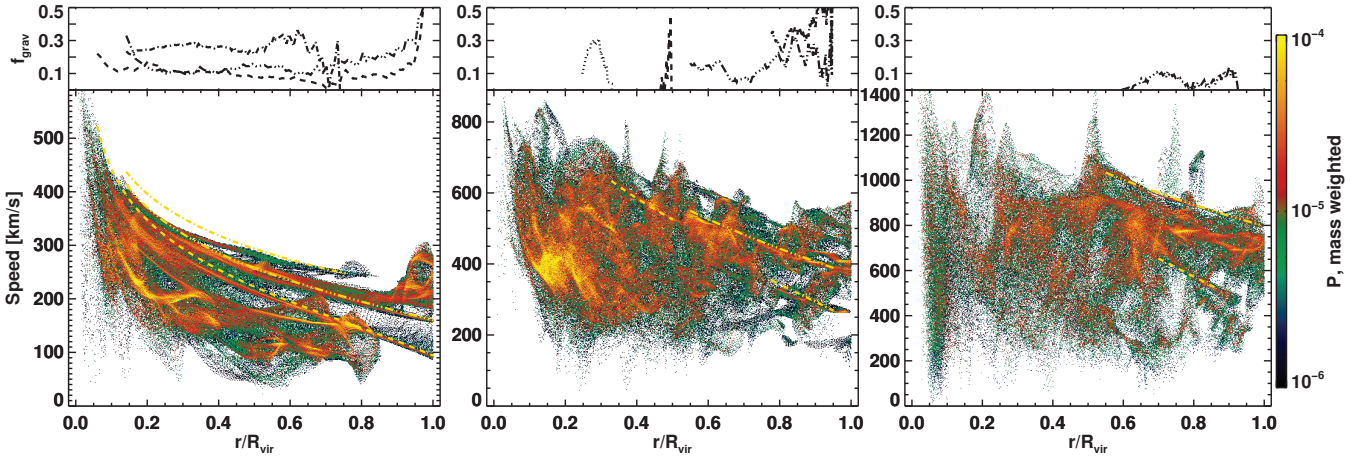


Figure 15. Phase diagrams showing speed of stream gas as a function of radius within targeted halos H1, H2 and H3 from left to right. The color scale represents mass weighted probability per speed-radius bin. Only stream densities are included here so that the streams stand out in the diagrams and are not drowned in the more diffuse gas. Over-plotted in yellow are free-falling speed profiles. Plots above each phase diagram show gravitational heating efficiency in the clearest streams, compared to the free-fall profiles.

stream can flow virtually unopposed towards the central galaxy whereas if the streams are wiggly and disrupted there should be greater opposition from the surrounding hot and diffuse gas.

It remains to be seen how much photo-fluorescence from the UV background is contributing the Lyα emissivities compared to gravitational processes, both smooth and messy. Before comparing these factors we describe how they're derived from the simulation output.

5.2 Computing the Lyα contributions

Since we store the photon flux in each cell we can easily keep track of the photo-heating and photoionization rates in the gas. If we assume that every photoionization leads to a recombination and that all the energy provided by photo-heating is released via collisional excitations,[3] we can estimate the UV contribution to the Lyα emissivity in each cell as:

$$\varepsilon_{UV} = 0.7\,\mathcal{H}_\gamma + 0.68\,\Gamma\,n_{\rm HI}\,\epsilon_{{\rm Ly}\alpha}, \qquad (11)$$

where $\mathcal{H}_\gamma$ is the photo-heating rate, the 0.7-factor is the conversion efficiency of cooling into Lyα photons[4], and the second term on the right is akin to Eq. 5. We refer to Appendix B for how to calculate the photo-heating rate. The UV contribution tends to be overestimated and can in fact be estimated higher than ε in hot regions where collisional excitation is not the dominant cooling channel, but since these regions are Lyα dim anyway this isn't a concern.

[3] The timescale for recombinations in streams is on the order of 10^5 - 10^6 years, which is short compared to the timescale for the in-fall of streams in these halos, $\sim$ 100 million years. The cooling timescale in streams is typically on the order of 10^4 to 10^5 years.
[4] This factor (roughly) represents the ratio of $C_{Ly\alpha}(T)$ to the hydrogen collisional excitation cooling rate discussed in Sec. 4.1. It means that we assume 70% of the energy dissipated via cooling to go into Lyα photons.

The only other driver of Lyα emission in our simulations are hydrodynamical processes which we can coin gravitational heating. Thus the approximate gravitational contribution to Lyα emissivity can be calculated in each gas cell as:

$$\varepsilon_{grav} = \max\,(0,\ \varepsilon - \varepsilon_{\rm UV}). \qquad (12)$$

5.3 Gravitational heating vs. UV fluorescence

Applying Eq. 11 we calculate the UV contribution to Lyα luminosity in each gas cell. We find that the relative UV contribution becomes weaker with increasing halo mass, with the ratio to the total halo Lyα luminosity going from 8% in H1 to 2% in H2 and 1% in H3 (see bottom row of Fig. 7). However, the relative UV contribution is generally stronger on the edges of the halos than near their centers.

In the top row of Fig. 16 we map the fractional UV contribution to the Lyα emissivity in streams and CGM gas, and in the bottom row of the same figure we plot the density distribution of the total luminosity, split into the UV (red) and gravitational (blue) contributions. As the maps and histograms show, the UV contribution is negligible everywhere except for the smooth and diffuse streams with $n_H \lesssim 0.05$ cm^{-3} in the H1 halo and on the outskirts of the H2 halo. A comparison with the mock observations in Fig. 10 reveals that these diffuse streams where the UV background contribution is non-negligible are nowhere close to being observable and all the observable emission is completely dominated by the gravitational contribution. The UV background contribution to extended Lyα emission can thus safely be ignored, at least until the observational sensitivity increases by two orders of magnitude or so.

Inclusion of local stellar UV radiation in our simulations may boost the UV contribution, and thus both the total luminosity of the halos and the extent of observable emission. Alternatively, the presence of a luminous quasar nearby may also significantly enhance the Lyα luminosity through fluorescence, as demonstrated by Cantalupo et al. (2005) and Kollmeier et al. (2010).

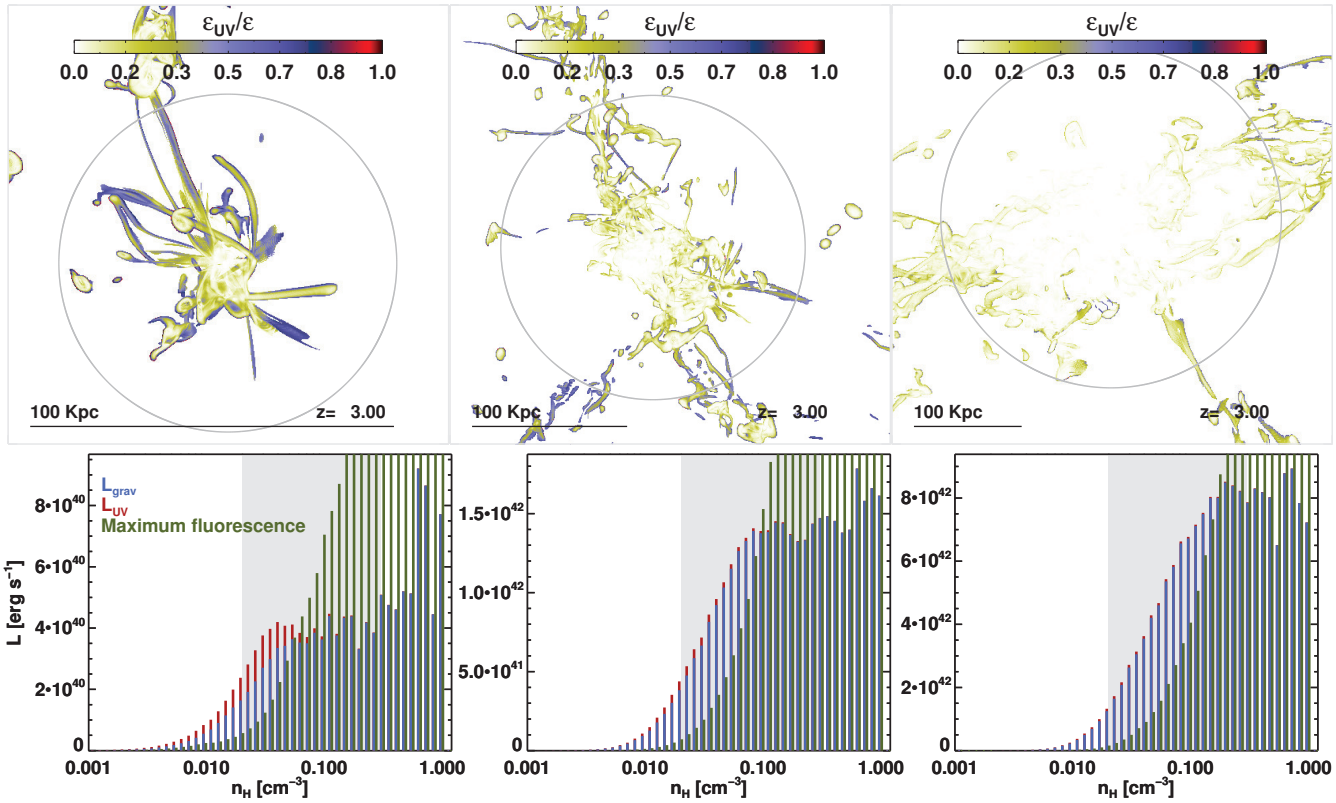


Figure 16. Top row: Fractional UV background contribution to the gas Lyα emissivity in the three main halos, H1, H2 and H3, from left to right. Shown are mass weighted averages along the LOS, and everything below stream densities (0.02 cm^{-3}) is ignored. **Bottom row:** Density distribution of the total Lyα luminosity of the same halos, split into the UV (red) and the gravitational (blue) contributions. Green columns denote the Lyα luminosity from maximum fluorescence. The shaded area represents stream densities. Note that the histograms have different scales on the y-axes.

5.4 Maximum fluorescence

Even though we lack in this work the inclusion of local sources of UV radiation, we can still evaluate the upper limit to the fluorescent Lyα luminosity that we can get from our simulated structures. This gives us a idea of the relative luminosity increase a local UV enhancement would provide, both in terms of the global luminosities of our halos and in terms of where the Lyα emissivity is boosted and where it is dimmed, compared to the gravitationally driven emission we have calculated.

To show the maximum fluorescent Lyα luminosities we can obtain, we re-calculate the Lyα emissivity of the simulated gas in the limit that $x_{\rm HII} = 1$ everywhere, corresponding to an infinite flux of UV photons. Note that in this limit ε_{coll} is zero everywhere and the Lyα emissivity is purely recombinative.

The green columns in the histograms in the bottom row of Fig. 16 show the Lyα luminosity of gas in different density bins in this limit. In all three halos at $n_{\rm H} \gtrsim 0.1{\rm cm}^{-3}$, maximum fluorescence outshines the normal gas due to the increase in HII abundance, while at lower densities it is dimmer than what we predict with gravitational heating, because collisional emission goes to zero.

In Fig. 17, we show the effect of maximum fluorescence on the Lyα emissivity of the gas within our three simulated halos. This is displayed as the ratio of $S_{\rm maxUV}$ – the luminosity computed assuming an infinite amount of ionizing photons everywhere –, to S – the luminosity used in the rest of the paper, which assumes a standard (though inhomogeneous) background value. Clearly, a strongly enhanced UV fluorescence will boost the Lyα emission in a significant fraction of the gas (the blue part), and may contribute significantly to LABs, as demonstrated by Cantalupo et al. (2005) and Kollmeier et al. (2010). However, from the perspective of observing accretion streams, the price to pay is the strong dimming of lower density structures (red).

This maximum fluorescence scenario is obviously optimistic, and only a tiny volume fraction of the Universe will likely come close to it, in the vicinity of rare and bright quasars in over-dense regions. Most of the IGM will more likely be in a regime comparable to our fiducial description, and its Lyα luminosity will be powered by collisional excitation.

6 SUMMARY AND CONCLUSIONS

We have in this work addressed the questions of whether gravitational heating may be the main driver of LABs, and how close we are to making direct and unambiguous detections of cold accretion streams via their Lyα emission.

To this purpose we have run and analyzed cosmological RHD simulations specifically tailored to accurately predict Lyα emission from extended structures. These simulations are idealized in the sense that the effects of stellar feedback

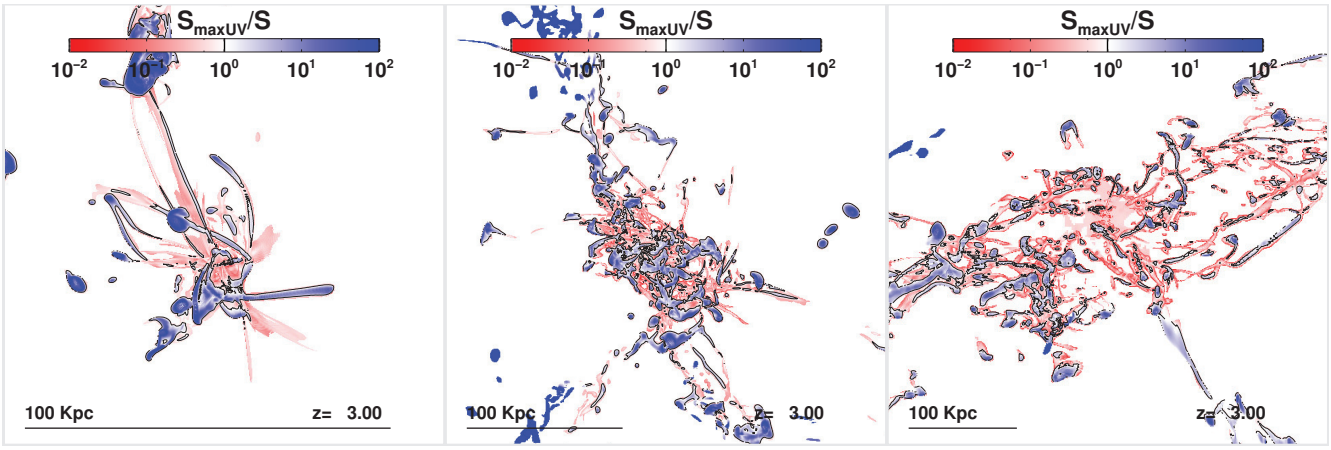


Figure 17. Ratio of the rest-frame Lyα surface brightness assuming maximum fluorescence (i.e. infinite UV flux, $S_{\rm maxUV}$) to our fiducial model, for halos H1, H2 and H3, from left to right. Fluorescence considerably boosts the densest clumps at the price of dimming the most diffuse streams.

and Lyα scattering are ignored such as to isolate the efficiency of gravitational heating in generating Lyα photons. Our analysis is focused on redshift 3, which corresponds to most LAB observations.

Our approach improves upon previous works in the following ways: (a) Using `RamsesRT`, our newly developed RHD version of the `Ramses` code, we include on-the-fly propagation of UV photons, which allows us to consistently and accurately model the self-shielding state in accretion streams and their resulting temperatures and ionization fractions, which are all very important to accurately predict their Lyα emissivity. (b) We apply a novel refinement strategy that allows us to optimally resolve accretion streams to an unprecedented degree and on much larger scales than previously. This allows us to spatially resolve the competition between gravitational heating and radiative cooling in those streams. (c) We post-process our simulation outputs with very small timesteps to ensure we also resolve said competition temporally. Failing to do this leads to a dramatic underestimate of Lyα emissivity of gas due to the commonly utilized numerical method of operator splitting, and previous works may have been marked by this problem. (d) We simulate more massive halos than hitherto done, based on the growing consensus that LABs are hosted by the most massive halos in the Universe.

There are nontheless issues regarding uncertainties that potentially affect our results. One is the likely presence of artificial overcooling in shocks. As pointed out by Creasey et al. (2011), shocks – or the mean-free paths of particles inside them – are almost exclusively under-resolved in cosmological simulations. The artificially broadened shocks can prevent the creation of hot and diffuse gas phase and instead allow it to efficiently cool and remain at temperatures where Lyα emission is the most effective cooling channel. We may thus over-predict the Lyα emissivities of shock regions in our simulations. However, this effect should be most severe in regions where gas is shocking on to galactic disks, and should thus be mainly constrained to CGM regions, and to the densest gas under consideration, i.e. $n_H \sim 1$ cm^{-3}. Weaker shocks may also exist at the boundaries of the disrupted streams in and around our more massive halos, but it seems unlikely that numerical overcooling is a big issue here, due to the high resolution, large volumes, low densities, and the fact that the Lyα emissivity is not particularly concentrated at the stream boundaries.

It is an unavoidable fact that the denser the gas in our simulations, or in any simulations for that matter, the larger the uncertainty in its Lyα emissivity. In particular, at densities $\gtrsim 0.1$ cm^{-3}, gas may cool down to $<< 10^4$ K via molecular or metal-line cooling, neither of which is included in our simulations. The densest gas is also in general the most Lyα luminous in our simulations: What we term CGM gas ($n_H > 0.3$ cm^{-3}) consistently accounts for 40% of the total Lyα luminosities of our halos, so we can estimate the total Lyα luminosities to be uncertain by (very) roughly 50%, and even more if we exclude still more diffuse gas than the CGM. We have however shown that our results and conclusions regarding LAB areas are not sensitive to the density threshold applied (i.e. above which densities we ignore Lyα emissivity).

Our main results are the following:

- Cold accretion streams in halos more massive than $\sim 10^{12}$ M$_\odot$ produces extended and luminous Lyα nebulae which are by large compatible with LABs observed at $z \sim 3$, in terms of morphology, luminosity and extent. Gravity alone provides most of the energy, and we find that extra sources such as UV fluorescence, Lyα scattering or superwinds are not necessary. This clearly doesn't rule out these other processes though, as they are likely all significant in the case of LABs, and further work is needed to study their complex interplay.
- In our simulations, LAB area and luminosity are reasonably well-behaved functions of halo mass. We use these relations to compute the cumulative luminosity and area distributions, and find that they are in reasonable agreement with observations given the relatively large uncertainties. This comparison however suggests that the combined effects of SN feedback, Lyα scattering and an enhanced local UV field may possibly have a negative impact on the luminosity and extent of simulated LABs, when conjoined with cold accretion.

- The model of gravitational heating as a driver of extended Lyα emission works according to our results, but we need to alter our notion of *how* it works: It is inefficient in the classic sense where gas accretion is smooth. Rather the accretion is messy and disrupted in massive halos and probably involves some mass loss to the surrounding hot diffuse medium.
- Our examination of maximum photofluorescence hints that in extreme cases local UV enhancement, e.g. near quasars, can boost the Lyα luminosity of LABs and to a lesser degree their extent. As demonstrated by Cantalupo et al. (2005) and Kollmeier et al. (2010), this means that large accretion flows may be more easily observed in the proximity of quasars than elsewhere.
- We find that cold accretion streams should be unambiguously observable via direct Lyα emission for the first time in the near future, on upcoming instruments such as MUSE and K-CWI which will allow to probe emission at surface brightnesses as low as $\sim 10^{-19}$erg s^{-1} cm^{-2} arcsec^{-2}.

Although we have significantly improved on previous work, a large number of theoretical issues remain to be addressed. In forthcoming papers, we plan to investigate the effects of Lyα scattering SNe-driven winds and local UV enhancement from star formation.

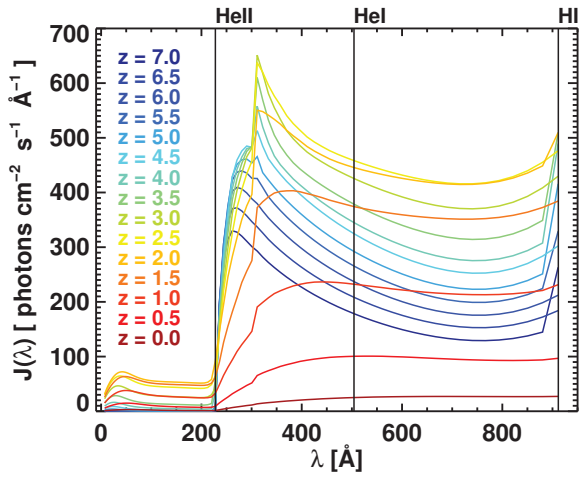


Figure A1. Evolution with redshift of the Lyman-continuum part of the UV spectrum of Faucher-Giguère et al. (2009) which we use in our simulations. The plot shows photon flux versus wavelength (λ) for selected redshifts. The vertical lines indicate how we split the spectrum into three (HI-, HeI-, HeII-ionizing) photon packages.

ACKNOWLEDGEMENTS

We thank Dominique Aubert and Romain Teyssier for helping us implement radiative transfer in `Ramses`. We are grateful to Yuichi Matsuda for kindly providing observational data, and we acknowledge valuable help and discussion on this work from Sebastiano Cantalupo, Stephanie Courty, Julien Devriendt, Tobias Goerdt, Joop Schaye and Romain Teyssier. Last but not least, we thank the referee, Claude-André Faucher-Giguère, for his careful and constructive review of this paper.

This work was funded in part by the Marie Curie Initial Training Network ELIXIR of the European Commission under contract PITN-GA-2008-214227. The simulations were performed using the HPC resources of CINES under the allocation 2011-c2011046642 made by GENCI (Grand Equipement National de Calcul Intensif). We also acknowledge computing resources at the CC-IN2P3 Computing Center (Lyon/Villeurbanne - France), a partnership between CNRS/IN2P3 and CEA/DSM/Irfu. JB acknowledges support from the ANR BINGO project (ANR-08-BLAN-0316-01).

APPENDIX A: THE QUASI-HOMOGENEOUS UV BACKGROUND

We use the UV background model of Faucher-Giguère et al. (2009), which is available on the web, and consists of the redshift-evolving spectrum shown in Fig. A1. As indicated by vertical lines in the plot, we discretize the spectrum into three photon *packages*; HI ionizing with frequencies in the range $(\nu_{\rm HI}, \nu_{\rm HeI})$; HeI ionizing in the range $(\nu_{\rm HeI}, \nu_{\rm HeII})$; and HeII ionizing in the range $(\nu_{\rm HeII}, \infty)$. All photons belonging to a package i share the common properties of flux J_i, average cross sections $\bar{\sigma}_{ij}$, where j stands for the three ionizable primordial species HI, HeI and HeII, and average energies $\bar{\epsilon}_{ij}$ per photoionization (again versus the three species). These properties are integrated from the redshift-dependent UV spectrum and updated every coarse timestep in the simulation. They are derived in the following way:

For each package i that is defined for the frequency interval (ν_{i0}, ν_{i1}) and given the UV spectrum $J(\nu)$ [photons cm^{-2} s^{-1} Hz^{-1}] (Fig. A1), we assign an average photoionization cross section against each species j (HI, HeI, HeII) as

$$\bar{\sigma}_{ij} = \frac{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_j(\nu) J(\nu)\, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} J(\nu)\, d\nu}, \tag{A1}$$

where we use the expressions for $\sigma_j(\nu)$ from Hui & Gnedin (1997). Similarly, we assign to each photon package average photon energies per photoionization event against each species:

$$\bar{\epsilon}_{ij} = \frac{\int_{\nu_{i0}}^{\nu_{i1}} h\nu\, \sigma_j(\nu) J(\nu)\, d\nu}{\int_{\nu_{i0}}^{\nu_{i1}} \sigma_j(\nu) J(\nu)\, d\nu}, \tag{A2}$$

where h is Planck's constant. The flux injected isotropically into each diffuse gas cell is derived for each package as

$$J_i = \int_{\nu_{i0}}^{\nu_{i1}} J(\nu)\, d\nu. \tag{A3}$$

When injected this way, the photons flow into adjacent cells which are above the UV density threshold and thus evolve into *local* photon fluxes F_i^γ representing the quasi-homogeneous UV field. Fig. A2 shows how the package properties evolve with redshift.

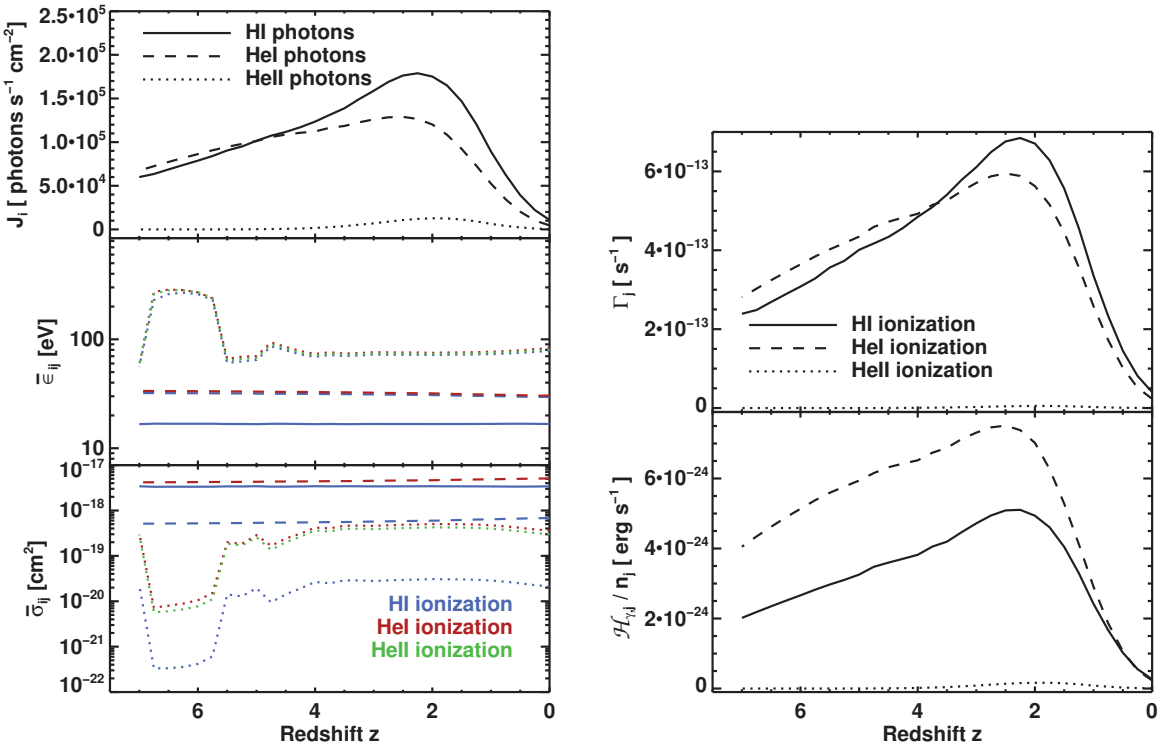


Figure A2. Left. Redshift-dependent properties of the three photon packages, integrated from the spectra in Fig. A1. **Right:** The redshift evolution of the photoionization rates against each species (upper plot) and per-species photo-heating rates (lower).

APPENDIX B: CALCULATING THE PHOTOIONIZATION AND PHOTO-HEATING RATES

The hydrogen photoionization rate Γ for the Lyman-continuum, in units of ionization events per hydrogen atom per unit time, is given by

$$\Gamma = \int_0^\infty \sigma_{\rm HI}(\nu) J(\nu)\, d\nu, \tag{B1}$$

where $\sigma_{\rm HI}$ is the hydrogen ionization cross section and J is the local photon flux, integrated over all directions. Since the UV spectrum in our RHD simulations is discretized into three photon packages, the photoionization rate is extracted from each gas cell in the simulation output as

$$\Gamma = \sum_{i=1}^{3} \bar{\sigma}_{i\rm HI} F_i^\gamma, \tag{B2}$$

where $\bar{\sigma}_{i\rm HI}$ is the average hydrogen ionization cross section for package i and F_i^γ is the *local* flux of package i photons (see Appendix A).

The photo-heating rate $\mathcal{H}_\gamma$ for the Lyman-continuum, in units of energy per time per volume, is given by

$$\mathcal{H}_\gamma = \sum_j^{\rm HI,HeI,HeII} n_j \int_0^\infty \sigma_j(\nu) F^\gamma(\nu)\,[h\nu - \epsilon_j]\, d\nu, \tag{B3}$$

where we sum the photo-heating rates over the primordial ion species HI, HeI and HeII. Here n_j and is the number density of ion species j, $\sigma_j(\nu)$ is the species' cross-section, $F^\gamma(\nu)$ is local photon flux, h is Planck's constant and ϵ_j the photoionization-threshold energies for species j.

With the discretization of the UV spectrum into three photon packages (see Appendix A) the integral in eq. (B3) becomes a sum:

$$\mathcal{H}_\gamma = \sum_j^{\rm HI,HeI,HeII} n_j \sum_{i=1}^{3} \bar{\sigma}_{ij} F_i^\gamma\,(\bar{\epsilon}_{ij} - \epsilon_j)\,, \tag{B4}$$

where F_i^γ is the local flux of photons in package i, and $\bar{\sigma}_{ji}$ and $\bar{\epsilon}_{ij}$ are the photon package properties defined in appendix A. We plot the redshift evolution of the per-species ionization- and heating rates in Fig. A2 (right).

APPENDIX C: MOCK LAB MAPS

Fig. C1 shows mock observation thumbnails of the most luminous halos in the H2 and H3 simulations, produced in the same way as the ones in Fig. 10, but applying observational parameters to match the surveys of M11 for direct comparison (these images can also be compared with thumbnails in e.g. Yang et al. 2010 and Erb et al. 2011). In practice, this means that we assume our objects are at redshift 3.1, smooth the images with a PSF with FWHM=1.4 arcsec, and put the thick inner surface brightness contours at $I = 1.4\ 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} (the thin outer ones are at $I = 10^{-19}$ erg s^{-1} cm^{-2} arcsec^{-2}). Morphologically our mocks resemble real LABs, asymmetric with a slight tendency to be filamentary and often having short sub-filaments

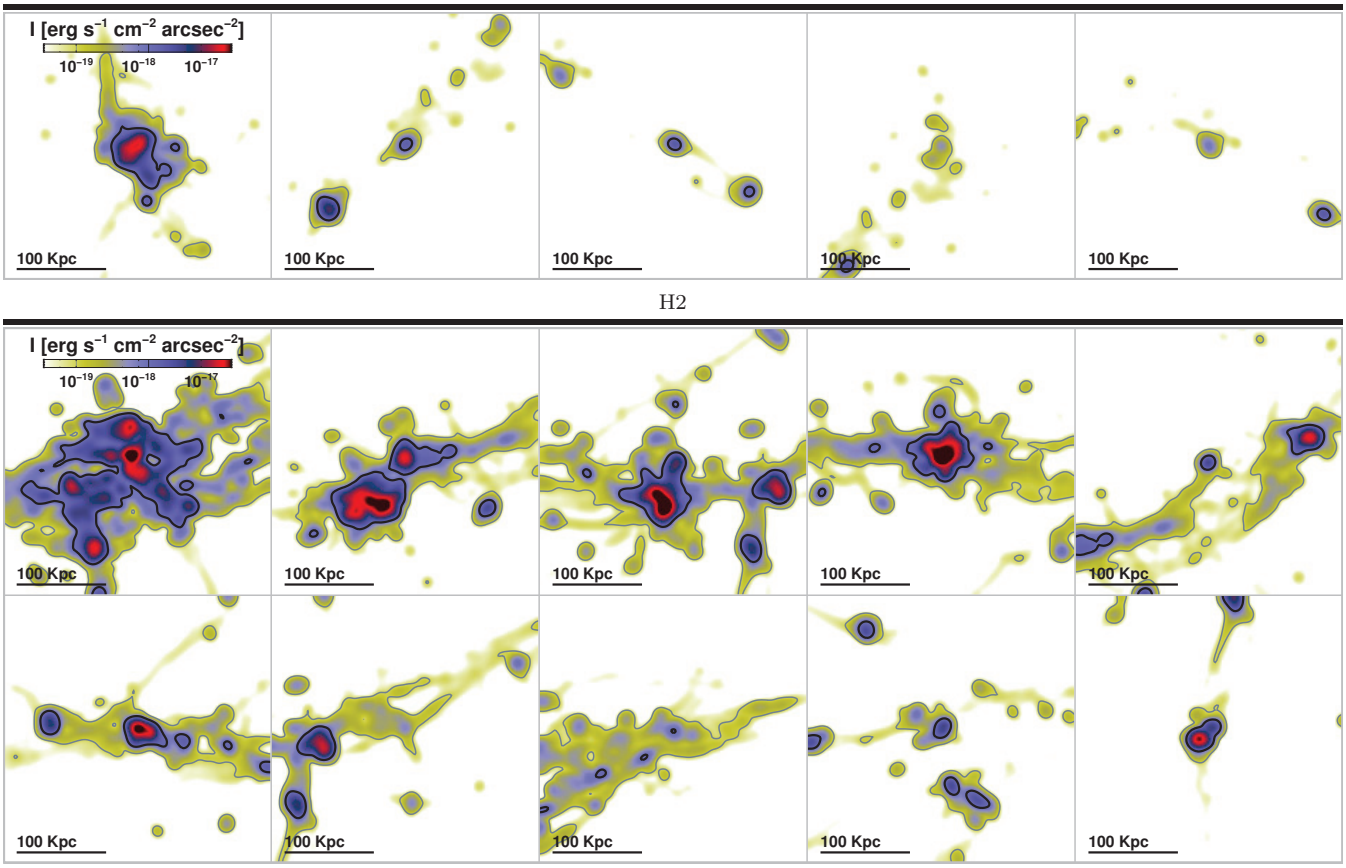


Figure C1. Mock observations showing I for the largest objects in the H2 and H3 simulations at redshift 3, smoothed with a Gaussian PSF of FWHM=1.4 arcsec to match recent observations. The physical widths of these thumbnail squares are all identical at 300 kpc ($\approx$ 40 arcsec). The thick inner contours correspond to $I = 1.4\ 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} and the thin outer ones to 10^{-19} erg s^{-1} cm^{-2} arcsec^{-2}.

that poke out of the main structure. In other words, our simulated LABs look like real ones.

APPENDIX D: COMPARING THE LAB LUMINOSITY FUNCTION WITH OBSERVATIONS

In Sec. 4.3.2, we derive a cumulative area function from our mock LAB observations and compare to observations. We have done the same comparison for a cumulative luminosity function, though our prediction should be less robust than the area function due to our lack of ISM modelling. Following M11, we assume $z = 3.1$ and apply a PSF with FWHM=1.4. We calculate the observed Lyα luminosities L_{obs} of all halos within the zoom regions of our simulations by integrating the surface brightness within $I = 1.4\ 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} contours. These luminosities are plotted against halo mass in Fig. D1 (left). Note the difference between L_{obs} in this plot and L in Fig. 8: There we plot intrinsic luminosities of halos whereas here we plot observable luminosities, assuming instrument sensitivity and cosmological extinction ($f_\alpha = 0.66$). The plot points are enveloped by a shaded region bordered by two power laws, with indexes 1 and 1.45 as indicated in the plot. These power laws are convolved with a Sheth-Tormen halo mass function (Sheth & Tormen 1999) at redshift 3.1 in order to produce the cumulative luminosity function envelope in Fig. D1 (right).

The black diamonds in Fig. D1 (right) represent the observations from M11, derived by binning the LABs by luminosity and dividing the count by the total survey volume, with Poissonian error bars. The comparison of our results to the observations is good, somewhat surprisingly so considering the lack of modelling of the emission and absorption in the compact peaks of Lyα emission, which contribute substantially to the total luminosity. As with the area function, we over-predict LAB abundances, though the prediction here is slightly closer to observations than in the case of areas.

The plot also shows, with line filled regions, predicted luminosity function envelopes, where gas of densities $n_{\rm H} \geqslant 0.3$ cm^{-3} and $n_{\rm H} \geqslant 0.1$ cm^{-3} is excluded from the analysis, for the horizontal and vertical line-fillings respectively. Much as with the area function (Fig. 12), excluding CGM densities and above ($n_{\rm H} \geqslant 0.3$ cm^{-3}) has relatively small impact on the luminosity function. However using the lower density threshold of excluding gas with $n_{\rm H} \geqslant 0.1$ cm^{-3} gives

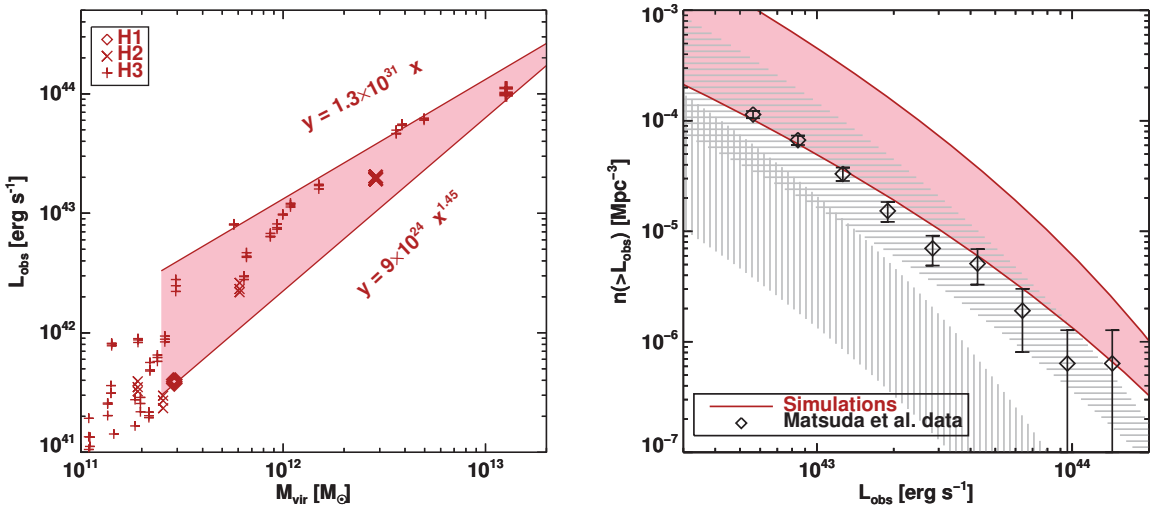


Figure D1. Left: Mock observed Lyα luminosities within contours of 1.4 10^{-18} erg s^{-1} cm^{-2} arcsec^{-2} versus halo mass for all zoomed halos in the three simulations. Each halo is observed along three simulation box axes and thus is represented by three points in the plot. The targeted halos are marked with thick symbols. The points are enveloped by two power laws shown in the plot. **Right:** Predicted luminosity function of LABs (shaded region), according to the power-law envelope from the left plot. The black symbols mark a rough luminosity function derived from a sample of 202 observed LABs from the survey of M11. The horizontally and vertically line-filled regions represent predicted luminosity functions with gas densities of $n_{\rm H} \geqslant 0.3$ cm^{-3} and $n_{\rm H} \geqslant 0.1$ cm^{-3} excluded, respectively.

a very abrupt change in the function, illustrating that the simulated LAB luminosities are more sensitive to the applied density cut than their areas.

APPENDIX E: COMPARISON TO PREVIOUS THEORETICAL WORK

Our work is similar in nature to the work of FG10 and G10 (see Sec. 1), and for comparison to those we have chosen the mass of the H1 halo to be similar to the halos on which they focus their analysis.

Fig. 12 in G10 shows a plot of halo luminosities versus virial masses at redshift 3.1, much like our Fig. 8. Their mass-luminosity power-law exponent is ~ 0.8, which is considerably shallower than our value of $\sim$ 1.25. Their less massive halos are more luminous than ours, with our ones catching up around 10^{12}M$_\odot$. Their targeted halos of masses $\approx 4\ 10^{11}$ M$_\odot$ are typically a few times more luminous than ours, and much more extended in Lyα emission. Their Fig. 10 shows mock observations of two of their targeted halos. Also using I_{-18} as a sensitivity limit, they have observable Lyα emission which is very asymmetric, clearly traces accretion streams and extends to about 100 kpc in length.

Their prediction is probably a bit over the top, since a giant LAB in a halo of this size implies that LABs should be very common, and this contradicts the generally accepted view that they are uncommon and associated with unusually over-dense regions in the Universe (e.g. Steidel et al. 2000; Prescott et al. 2008; Yang et al. 2010). The cause of their over-prediction can probably be traced to an overestimate in gas temperatures due to their self-shielding approximation being applied in post-processing, as pointed out by FG10 (see also our Sec. 4.1).

Fig. 2 in FG10 shows a plot of halo luminosities versus halo masses at redshift 3 for the various numerical approaches. Their estimate which is most comparable to ours (their prescription 7, that sums all gas) has a power law with exponent ~ 1.1, steeper than that of G10, but still a bit shallower than our exponent, and their halos are slightly Lyα dimmer than ours, typically around half the luminosity for a given mass, though this varies quite a lot due to scatter. The luminosity difference may be partly explained by the overdensity of our simulated regions, which tends to increase the brighness of halos of similar mass, going from the least massive to the most massive simulation. Their more conservative prescription excludes star-forming gas from their analysis, which in their simulations is gas with $n_{\rm H} > 0.13$ cm^{-3}, likely a more realistic threshold than our $n_{\rm H} > 1$ cm^{-3}. On this exclusion, the luminosity drops by 1-2 orders of magnitude. This is a bit more dimming than we find in our results: If the left plot of Fig. D1 is considered, where the region filled with vertical grey lines corresponds to our luminosities where gas with $n_{\rm H} > 0.1$ cm^{-3} is excluded from our analysis, it can be seen that the luminosity drops by $\lesssim 1$ order of magnitude compared to our normal prescription of including all gas with $n_{\rm H} > 1$ cm^{-3}.

They also show mock observations of a 2.5 10^{11} M$_\odot$ halo at redshift 3, that includes Lyα scattering. Their Fig. 7, middle left, can be compared to ours (again, their prescription 7, that sums all gas). A contour at I_{-18} marks a very circular source centered on a galaxy, about 15 kpc in diameter, which is similar to our H1 halo observation. The Lyα luminosity of their halo is 8 10^{41} erg s^{-1}, close to the 'observed' luminosity of our H1 halo of 6 10^{41} erg s^{-1}.

In terms of the emission coming from $\sim 10^{11}$ M$_\odot$ halos we thus seem to be in fair agreement with FG10, though the LABs produced by our simulations appear to be somewhat

Table F1. Table of symbols

A	Area
$\alpha_{\rm HI}^{\rm B}(T)$	Case-B recombination rate for hydrogen
$C_{\rm Ly\alpha}(T)$	Rate of collisional excitations
ε	Lyα emissivity
$\epsilon_{\rm Ly\alpha}$	Energy of a Lyα photon (10.2 eV)
f_α	Cosmological transmission factor for mock observations (we use $f_\alpha = 0.66$).
f_c	Light-speed fraction
f_{grav}	Gravitational efficiency
G	Gravitational constant
Γ	Hydrogen photoionization rate
$\mathcal{H}_\gamma$	Photo-heating rate
I	Observed Lyα surface brightness
I_{-18}	Fiducial observational sensitivity limit, set to 10^{-18} erg s^{-1} cm^{-2} arcsec^{-2}
L	Lyα luminosity
M_{vir}	Virial mass
$n_{\rm H}^{UV}$	Density threshold for UV background emitting gas
n_i	Number density of species i
N_i	Column density of species i
r	Radius from halo center
R_{vir}	Virial radius
S	Rest-frame Lyα surface brightness
T	Temperature
v	Speed
x_i	Ionization fraction of ion species i
z	Cosmological redshift

more luminous, even when matching their more conservative prescriptions. In terms of LAB extent it is harder to tell, since they don't consider mock observations of halos more massive than 3 10^{11}M$_\odot$, and it is thus hard to tell whether or not their massive halos produce LAB-like objects.

APPENDIX F: TABLE OF SYMBOLS

REFERENCES

Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Aubert D., Teyssier R., 2008, MNRAS, 387, 295
Aubert D., Teyssier R., 2010, ApJ, 724, 244
Bacon R. et al., 2006, The Messenger, 124, 5
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Birnboim Y., Dekel A., Neistein E., 2007, MNRAS, 380, 339
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Callaway J., Unnikrishnan K., Oza D. H., 1987, Phys. Rev. A, 36, 2576
Cantalupo S., Porciani C., Lilly S. J., Miniati F., 2005, ApJ, 628, 61
Cen R., 1992, ApJS, 78, 341
Cen R., 2011, AJ, 742, L33
Creasey P., Theuns T., Bower R. G., Lacey C. G., 2011, MNRAS, 415, 3706
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A. et al., 2009, Nature, 457, 451
Dijkstra M., Haiman Z., Spaans M., 2006, ApJ, 649, 37
Dijkstra M., Loeb A., 2009, MNRAS, 400, 1109
Dubois Y., Teyssier R., 2008, A&A, 477, 79
Erb D. K., Bogosavljević M., Steidel C. C., 2011, AJ, 740, L31
Fardal M. A., Katz N., Gardner J. P., Hernquist L., Weinberg D. H., Davé R., 2001, ApJ, 562, 605
Faucher-Giguère C.-A., Kereš D., 2011, MNRAS, 412, L118
Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, ApJ, 725, 633
Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982
Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416
Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 681, 831
Francis P. J. et al., 1996, ApJ, 457, 490
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, eprint arXiv, 1103, 2130
Furlanetto S. R., Schaye J., Springel V., Hernquist L., 2005, ApJ, 622, 7
Gnedin N. Y., Abel T., 2001, New Astronomy, 6, 437
Goerdt T., Dekel A., Sternberg A., Ceverino D., Teyssier R., Primack J. R., 2010, MNRAS, 407, 613
Haiman Z., Rees M. J., 2001, ApJ, 556, 87
Haiman Z., Spaans M., Quataert E., 2000, ApJ, 537, L5
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Iliev I. T. et al., 2006, MNRAS, 371, 1057
Iliev I. T. et al., 2009, MNRAS, 400, 1283
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Keel W. C., Cohen S. H., Windhorst R. A., Waddington I., 1999, AJ, 118, 2547
Keel W. C., White R. E., Chapman S., Windhorst R. A., 2009, AJ, 138, 986
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kimm T., Slyz A., Devriendt J., Pichon C., 2011, MNRAS, 413, L51
Kollmeier J. A., Zheng Z., Davé R., Gould A., Katz N., Miralda-Escudé J., Weinberg D. H., 2010, ApJ, 708, 1048
Komatsu E., Smith K. M., Dunkley J., Bennett C. L., et al., 2011, ApJS, 192, 18
Laursen P., Sommer-Larsen J., 2007, ApJ, 657, L69
Martin C., Moore A., Morrissey P., Matuszewski M., Rahman S., Adkins S., Epps H., 2010, Ground-based and Airborne Instrumentation for Astronomy III. Edited by McLean, 7735, 21, (c) 2010: American Institute of Physics
Maselli A., Ferrara A., Ciardi B., 2003, MNRAS, 345, 379
Matsuda Y. et al., 2004, AJ, 128, 569
Matsuda Y. et al., 2011, MNRAS, 410, L13
Mori M., Umemura M., Ferrara A., 2004, ApJ, 613, L97
Nilsson K. K., Fynbo J. P. U., Møller P., Sommer-Larsen J., Ledoux C., 2006, A&A, 452, L23
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Ohyama Y. et al., 2003, ApJ, 591, L9
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei
Palunas P., Teplitz H. I., Francis P. J., Williger G. M., Woodgate B. E., 2004, ApJ, 602, 545
Prescott M. K. M. et al., 2011, eprint arXiv, 1111, 630
Prescott M. K. M., Dey A., Jannuzi B. T., 2009, ApJ, 702, 554
Prescott M. K. M., Kashikawa N., Dey A., Matsuda Y., 2008, ApJ, 678, L77
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of

scientific computing
Prunet S., Pichon C., Aubert D., Pogosyan D., Teyssier R., Gottloeber S., 2008, AJS, 178, 179
Rasera Y., Teyssier R., 2006, A&A, 445, 1
Rauch M., Becker G. D., Haehnelt M. G., Gauthier J.-R., Ravindranath S., Sargent W. L. W., 2011, MNRAS, 418, 1115
Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541
Ribaudo J., Lehner N., Howk J. C., Werk J. K., Tripp T. M., Prochaska J. X., Meiring J. D., Tumlinson J., 2011, ApJ, 743, 207
Schaye J., 2001, ApJL, 562, L95
Schaye J., 2004, ApJ, 609, 667
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Silk J., 1977, ApJ, 211, 638
Smith D. J. B., Jarvis M. J., 2007, MNRAS, 378, L49
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2000, ApJ, 532, 170
Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, ApJ, 736, 160
Taniguchi Y., Shioya Y., 2000, ApJ, 532, L13
Teyssier R., 2002, A&A, 385, 337
Toro E. F., 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction
Truelove J. K., Klein R. I., McKee C. F., Holliman J. H., Howell L. H., Greenough J. A., 1997, ApJ, 489, L179
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
van de Voort F., Schaye J., 2011, eprint arXiv, 1111, 5039
van de Voort F., Schaye J., Altay G., Theuns T., 2011a, eprint arXiv, 1109, 5700
van de Voort F., Schaye J., Booth C. M., Haas M. R., Vecchia C. D., 2011b, MNRAS, 414, 2458
Weijmans A.-M., Bower R. G., Geach J. E., Swinbank A. M., Wilman R. J., de Zeeuw P. T., Morris S. L., 2010, MNRAS, 402, 2245
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wise J. H., Cen R., 2009, ApJ, 693, 984
Yang Y., Zabludoff A., Eisenstein D., Davé R., 2010, ApJ, 719, 1654
Zheng Z., Cen R., Weinberg D., Trac H., Miralda-Escudé J., 2011, ApJ, 739, 62

5.2 Notes on extended Lyα emission

The following sections elaborate upon some remarks made in the Rosdahl and Blaizot (2012) paper. Mostly this concerns convergence tests made to show that the tweaking of some parameters doesn't affect our results and conclusions, specifically the resolution, light speed, and the UV emission density threshold. We elaborate on the effect that operator splitting has on the Lyα emissivity of gas and show how the problem is fixed. We discuss further the validity of using a simple self-shielding approximation that makes the simulations much cheaper. Then we show how the disruption of streams is a physical and not a numerical effect, and finally, we discuss the effect of using the smooth RT trick, described in §3.2.5.

We will refer here to the simulations presented in the paper as production simulations, to differentiate from alternative simulations we run for various checks and comparisons. As in the paper we refer to our three zoom simulations as H1, H2, and H3, in order of increasing mass of the targeted halo. We refer to the Rosdahl and Blaizot (2012) paper as RB12.

5.2.1 Resolution convergence

In RB12 we state that we've established the resolution in our simulations to be adequate for our purposes. Partly this concerns resolving the collapse of structures via the Jeans length – this point is discussed fully in the paper – but also one wants to verify via simple convergence tests that main results are not resolution dependent. This is typically done by varying the resolution: If it has the effect of changing the results notably, one is forced to conclude that the resolution is inadequate.

For this we have run a lower resolution equivalent of the H1 simulation, where the resolution is degraded by one level compared to the production run: The initial conditions thus consist of three nested grids with effective resolution of 128^3, 256^3 and 512^3 resolution elements (gas/DM) rather than 256^3, 512^3 and 1024^3, and the maximum AMR level reached at $z = 3$ is 14 rather than 15. We compare density and Lyα emissivity maps from these simulations are redshift 3 in Fig. 5.1, where we also include for comparison a degraded simulation where stream refinement has furthermore been turned off (far right). Corresponding structures can clearly be recognized between the simulations, though they do notably smear out with lower resolution.

In Fig. 5.2 we plot one of the quantities of main importance to this work, the halo luminosities, and compare between production and degraded resolution. The Lyα luminosities are not substantially changed, though the degraded resolution seems to have the systematic effect of slightly increasing them.

Although the resolution in the H1 simulation is higher than in H2 and H3 in terms of physical mass and length (by factor half and a quarter, respectively), it is actually lower in terms of the number of resolution elements that represent the halos, which scales approximately with mass and radius cubed. We thus should expect that if the H1 resolution is sufficient, the H2 and H3 resolutions should also be. To be sure, we have done a similar check on the H3 simulation, here going in the other direction, i.e. increasing the resolution from the production run rather than decreasing it. This is made possible by turning off RT in the high-resolution equivalent simulation and instead employing the density-threshold self-shielding approximation described in the paper, where the UV background is applied

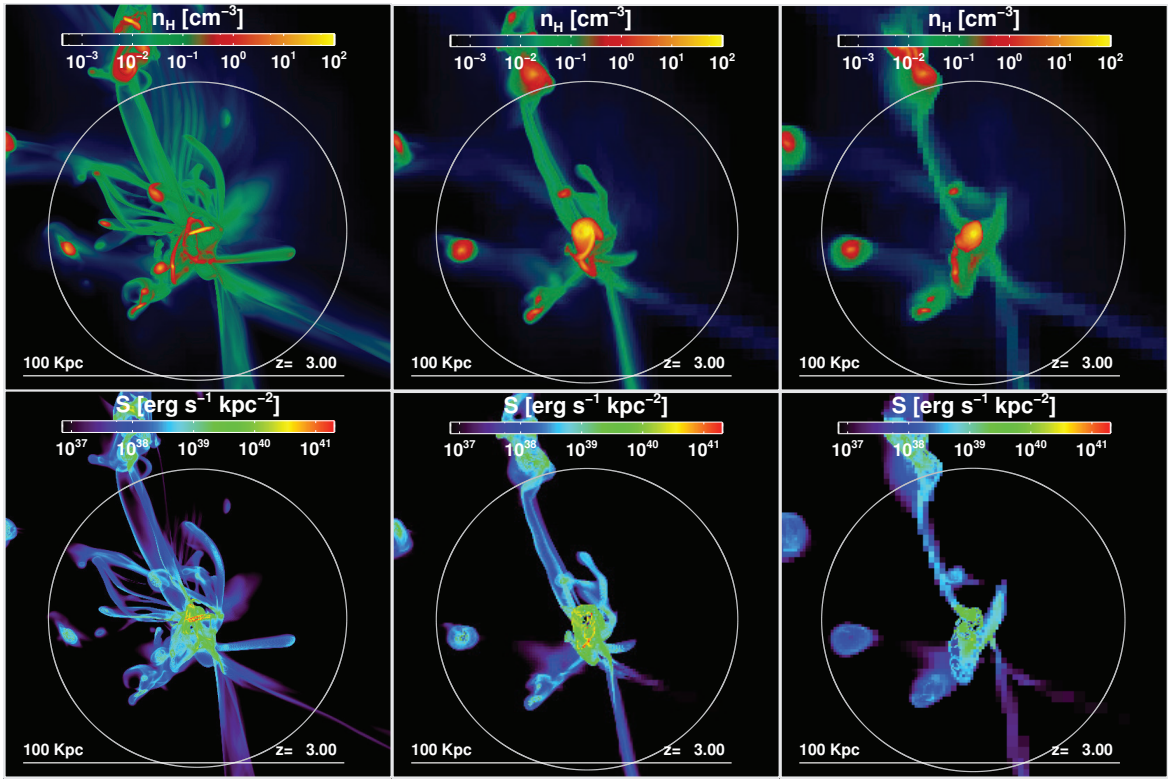


Figure 5.1: *Resolution check: Comparison of varying resolution in the H1 simulation at redshift 3. The top row shows gas density maps and the bottom row shows Lyα surface emissivity.* ***Left:*** *The production resolution.* ***Middle:*** *One-level degraded resolution, i.e. AMR maxlevel=14 instead of 15, and the mass of each DM particle is increased by a factor of eight.* ***Right:*** *One-level degraded resolution, and no refinement criteria on streams (also, the polytropic limit is at $n_{\rm H} = 0.1\,{\rm cm}^{-3}$ here, which reduces the Lyα luminosity substantially).*

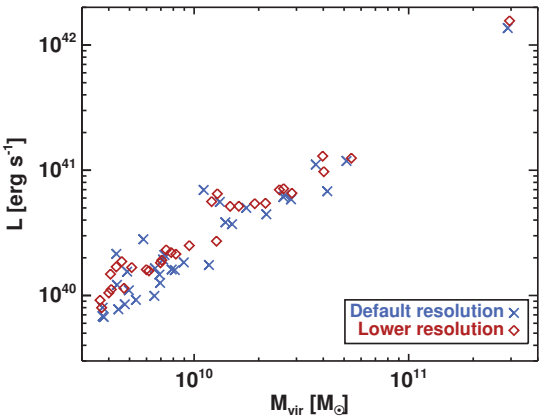


Figure 5.2: *Resolution check: Comparison of halo luminosities at redshift 3 for different resolutions of the H1 simulation. Blue x's are from the production H1 simulation and the red diamonds are from a simulation where the initial conditions and the maximum reachable resolution have been degraded by one AMR level (i.e. particles are 8 times heavier and cells are doubled in width).*

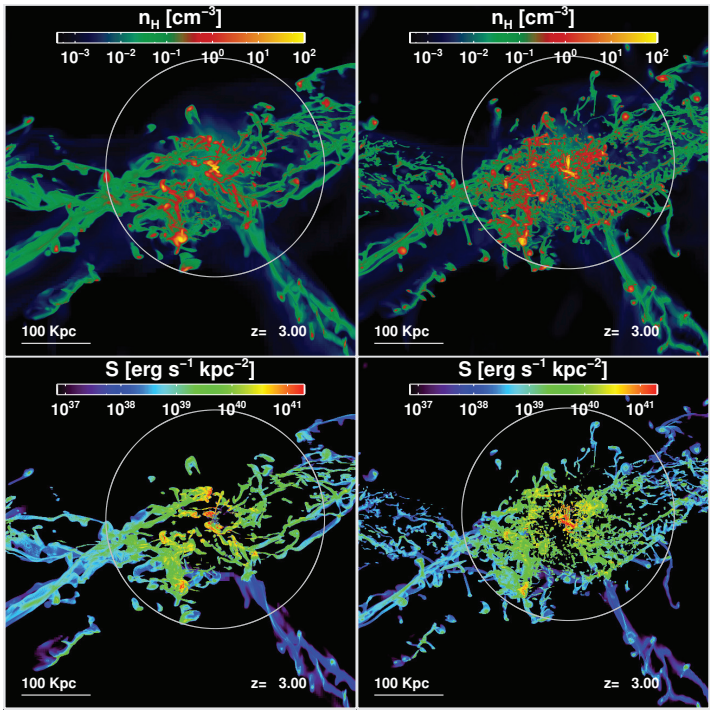


Figure 5.3: *Resolution check: Comparison of varying resolution in the H3 simulation at redshift 3. The top row shows gas density maps and the bottom row shows Lyα surface emissivity.* ***Left:*** *The production resolution.* ***Right:*** *One-level upgraded resolution, i.e. AMR maxlevel=15 instead of 14, and the mass of each DM particle is decreased by a factor of eight. Also, note this is a non-RT simulation, using instead a simple density-threshold self-shielding approximation – but as we discuss in the paper, and quantify in §5.2.5, this should produce the same Lyα emissivities and luminosities in our simulations.*

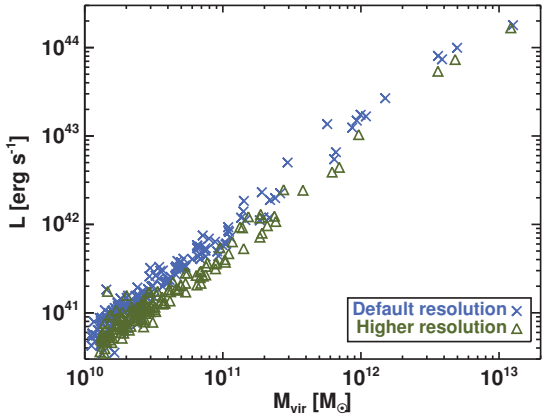


Figure 5.4: *Resolution check: Comparison of halo luminosities at redshift 3 for different resolutions of the H3 simulation. Blue x's are from the production H3 simulation and the green triangles are from an equivalent non-RT simulation where the initial conditions and the maximum reachable resolution have been upgraded by one AMR level.*

homogeneously in every cell, except it is turned off in cells with gas denser than 10^{-2} cm^{-3}. As we conclude in the paper, and show more quantitatively here in §5.2.5, employing this approximation gives Lyα emissivity results that are just as robust as using full RT, with the caveat that the UV contribution is negligible. Fig. 5.3 shows maps comparing gas structures and Lyα emissivities in the H3 production run and the higher-resolution equivalent. As with the H3 comparison, we see more structures appearing with higher resolution, but the overall morphology and Lyα emissivities stay very similar. The differences are largely found at scales small enough that they would be 'smeared' away with camera resolution and, more importantly, seeing.

Fig. 5.4 compares the halo luminosities at those two resolutions. Note the production resolution shows more halos because a slightly wider volume around the targeted halo is allowed there for the halo finder. We see the same trend here as for the different H1 resolutions, namely that going to higher resolution systematically reduces the total Lyα luminosities by a small fraction.

We conclude from this evidence that the resolution is adequate in all production simulations, though there are hints our Lyα luminosities may be overestimated by some fraction due to limited resolution. There is admittedly a somewhat hand-waving quality to this argument, but this is how the game goes. In the end, our resolution is not a product of choice, but rather just the highest we can attain with the resources at hand, and the best we can do is justify it after the fact.

5.2.2 Light speed convergence

In RB12 we state that we use a light speed fraction of $f_c \sim 10^{-2}$ in our simulations, and that using this rather than a full light speed doesn't affect our results. We substantiate this by stating that we have run simulation tests with the light speed tweaked by a factor of few in both directions, i.e. up and down, without notable effect on our results. We did this running three lower resolution (by one level) analogues of the H1 simulation, with stream refinement furthermore turned off. Note also that in these simulations, the polytropic equation of state sets the temperature in gas denser than 0.1 cm^{-3}, which is a tenth of the limit we use in the production runs.

Fig. 5.5 shows maps at $z = 3$ for these three low-resolution simulations. The upper map shows gas density and the lower ones the photoionization rate. From left to right are the light speed fractions 0.2%, 1% and 5%. The middle value is what is used in the H1 production run and the low one is close to what is used in the H2 and H2 runs (0.33%). From the upper maps it can be seen that the diffuse gas structures are identical the different light speed runs, but the compact structures are slightly different, which is likely an effect of the different size time-steps which are used in the simulations because of the different Courant constraints. If we consider the ionization rate maps, it appears that we have very similar self-shielding morphologies in the simulations.

The important thing to check, as far as we are concerned, is whether the different light speeds have an effect on the Lyα luminosities. Fig. 5.6 shows a plot of the Lyα luminosities of all halos within the zoom region in the three light speed simulations. The luminosities are lower than for the H1 production run, which is just due to the different polytropic limit: Because of it, we are forced to ignore any gas denser than 0.1 cm^{-3} here. It is clear though

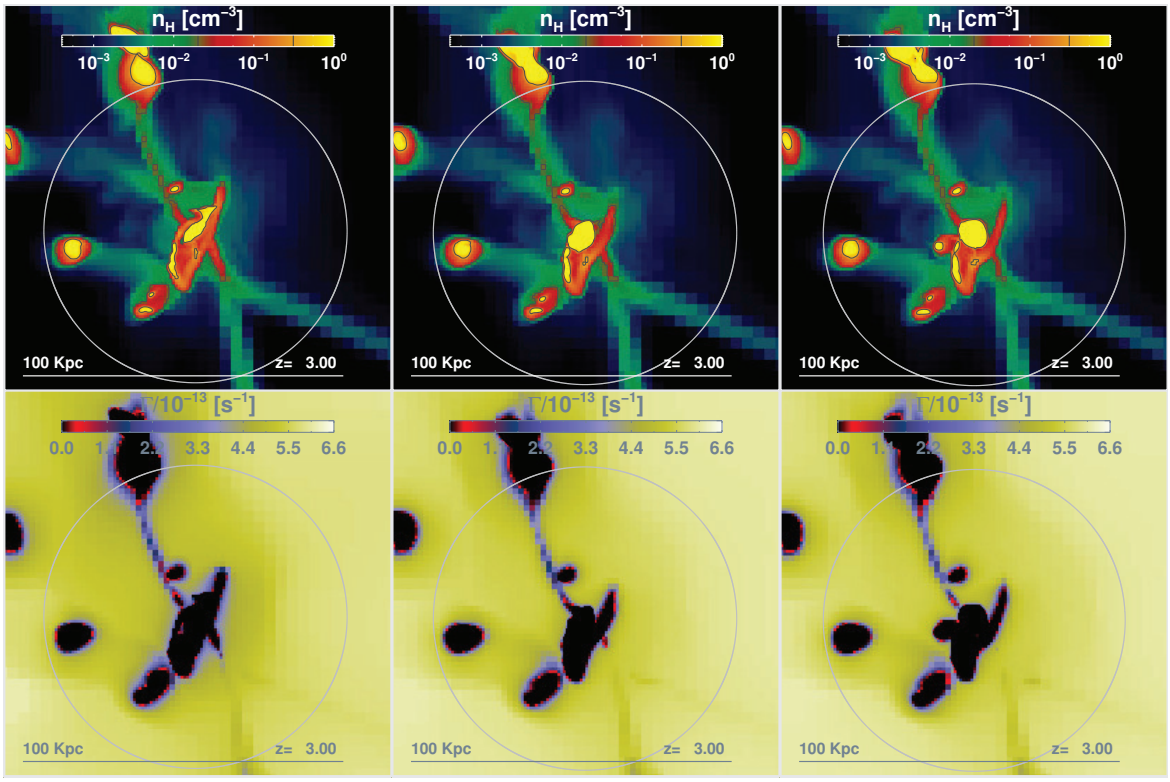


Figure 5.5: *Light speed check: Comparison of different light speeds in low-resolution equivalents of the H1 simulation at redshift 3.* ***Top*** *row shows gas density maxima along the LOS and* ***bottom*** *row shows H photoionization rate minima along the LOS.* ***Left:*** $f_c = 0.2\%$. ***Middle:*** $f_c = 1\%$. ***Right:*** $f_c = 5\%$.

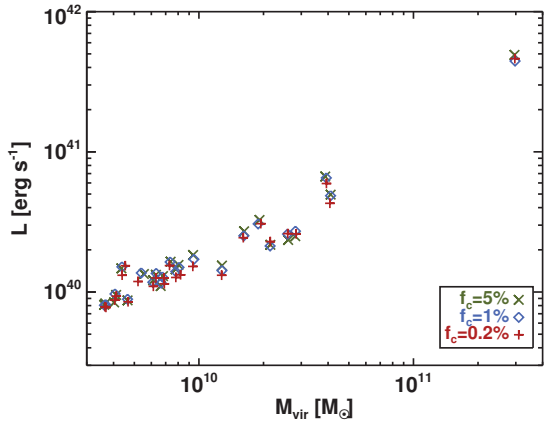


Figure 5.6: *Light speed check: Comparison of halo luminosities at redshift 3 in low resolution equivalents of the H1 simulation that are identical except different light speeds are used.*

from the plot that the different light speeds have a negligible effect on changing the halo luminosities, and any small differences seem to be random – there is no trend that favours one light speed to give brighter luminosities in general than others. And since tweaking it from 1% to 5% doesn't have any effect on the luminosities, we shouldn't expect any further

luminosity differences in going to full light speed.

We should thus be in the clear in using light speed fractions of $\sim 1\%$, *in low-resolution equivalents of the H1 simulation, with a shifted density limit for the polytrope*. Does this allow us to conclude that we are also in the clear with similar light speed fractions in the production runs? The answer is surely yes: There is no reason to think changing the resolution would change the effect of different light speeds on the Lyα emissivities, not even in gas denser than the polytropic limit.

It is also fully expected: What matters here is that the absorbing regions are seeing the 'correct' UV background, i.e. that they don't see a background which was emitted a cosmologically substantial time ago, when the background intensity was dramatically different. And as long as the UV sources are close by, as we ensure with the $n_{\rm H}^{UV}$ threshold, we are indeed in the clear. Furthermore, one of our conclusions is that the UV contribution is actually negligible in the observable Lyα emissivities, so any delay of UV radiation, or even a complete lack of it for that matter, should have a negligible effect on our results.

5.2.3 UV emission density threshold convergence

In RB12 we emit the UV background from under-dense regions, which are defined by a gas density threshold, $n_{\rm H}^{UV}$, which is set to 10^{-4} cm^{-3} in the H1 production run, and a higher value of 3 10^{-4} cm^{-3} in the H2 and H3 runs to compensate for lower light speed. We made a convergence check on an analogue of the H1 simulation with one-level degraded resolution, to see if any tweaking of this limit affects our conclusions. Fig. 5.7 shows the results of this, in the form of halo Lyα luminosities. As with the light speeds, we find here that the difference is negligible. One should be careful though with the density threshold: Using the lowest value presented here, of $n_{\rm H}^{UV} = 10^{-5}$ cm^{-3} actually results in a substantially delayed UV background reaching the main halo center, as it makes the UV sources rather distant from the halo. It also actually delays the UV background turning on in the simulation: We

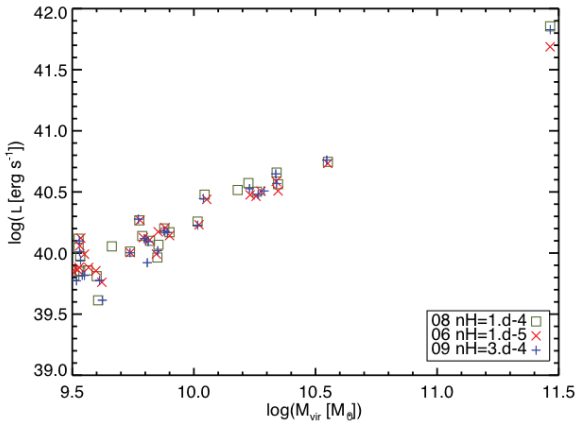


Figure 5.7: *UV emission threshold check: Comparison of halo luminosities at redshift 3 in low resolution equivalents of the H1 simulation that are identical except different density limits are used for the UV emission threshold. Green boxes correspond to $n_{\rm H}^{UV} = 10^{-4}\,{\rm cm}^{-3}$ (used in the production H1 simulation), red X's to $n_{\rm H}^{UV} = 10^{-5}\,{\rm cm}^{-3}$ and blue crosses to $n_{\rm H}^{UV} = 3\,10^{-4}\,{\rm cm}^{-3}$ (used in H2 and H3).*

set it to turn on at redshift 8, but it turns out that there is little if any gas that exists with densities as low as 10^{-5} cm^{-3} – which makes for little if any sources of UV at this early redshift. As with the light speed this still has little effect on the Lyα emissivities though, but has more to do with accurately estimating the densities at which the gas is self-shielding: If the UV background is non-existent or greatly delayed, we have little chance of getting the self-shielding right.

5.2.4 Operator splitting and underestimated Lyα emissivities

In section 2.3 of RB12 we mention the issue of operator splitting (OS) in `Ramses/RamsesRT`, which has the effect that the Lyα emissivity of gas may be greatly underestimated, and that we fix this problem by post-processing our production simulations with very small time-steps. We will here elaborate upon this issue.

We have already in §3.2 of this thesis discussed the use of OS in the context of the RT implementation inside `Ramses`, the main purpose of which is to *(i)* reduce the complexity of the RT equations that are solved and *(ii)* reduce the time-step length of the RT transport step. We also described (in §3.2.5) a problem that arises with OS in the RT implementation, namely that the discreteness that is inherent to OS sometimes causes long and unnecessary sub-cycling in the thermochemistry step, and a strong depletion of photons. The problem here, with OS and Lyα emissivities, is similar in nature.

Much like the splitting of the RT equations into transport and thermochemistry, so is the Euler equation of energy conservation (2.3) also split into advection and thermochemistry in `Ramses` and `RamsesRT` alike. The advection is solved over a time-step which is constrained by Courant-type conditions, and the thermochemistry is sub-cycled over time-steps that are often much smaller. The result of the advection step, i.e. the thermal energy at the end of it, goes as an initial condition into the first step in the thermochemistry cycle, as is shown schematically in Fig. 5.8. In the case that the cooling time is shorter than the advection time, and the radiative cooling is sub-cycled, the temperature at the end of the hydrodynamical time-step will be slightly *underestimated* compared to a more realistic scenario where the two processes of advective heating and radiative cooling are allowed to happen simultaneously. This is basically because cooling is allowed to happen without the direct competition of advective heating, allowing it, in a sense, to get the upper hand. The degree of this underestimation largely depends on the amount of sub-cycling, which depends on the cooling time of the gas, which in turn scales inversely with the gas density squared. This is normally not a big issue in simulations: The temperature underestimate is only a slight one, and it should have no significant effect on the dynamics of the simulation.

However, in the context of gravitationally powered Lyα emission from gas, the effect becomes very important. Here, operator splitting is separating the cause and effect we're trying to simulate: Gravitational heating is an advective process and Lyα radiation is a thermochemical one. According to the model of gravitationally powered Lyα emission, cold accretion streams should exist in a semi-equilibrium state where the rate of gravitational heating is matched by the rate of Lyα cooling. Furthermore, as we show in RB12 (Fig. 6), the Lyα emissivity of the gas is very sensitive to the temperature in those streams, precisely because Lyα cooling is very effective at the stream temperatures (just above 10^4 K). Because OS doesn't allow advective heating and Lyα cooling to happen in direct competition, and

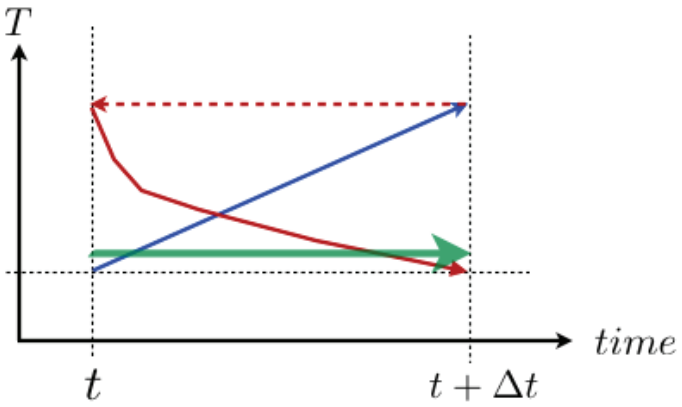


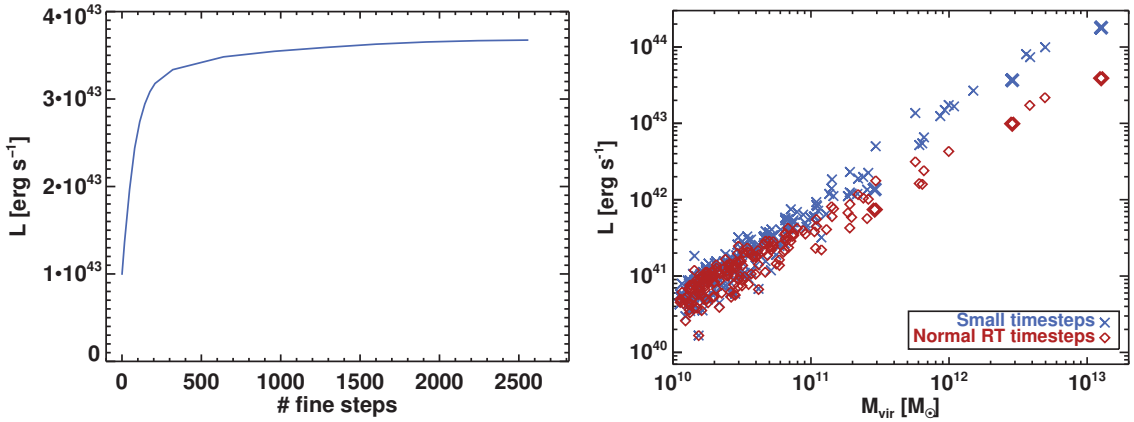
Figure 5.8: *Operator splitting in* `Ramses`*: Schematic showing the evolution over a hydrodynamical step Δt of a cell temperature in* `Ramses/RamsesRT`*. Starting at time t, the advection is first solved over the whole time-step (blue), resulting in an increased temperature at time $t + \Delta t$, e.g resulting from gravitational heating. This temperature is then used as an initial condition in solving the thermochemistry over the same time-step, which may happen in subcycles (red). The thick green line represents a more realistic scenario where operator splitting is not used and the processes of advective heating and radiative cooling (here in equilibrium) are allowed to take place simultaneously, with a final temperature that is different – typically higher – than with operator splitting.*

because the Lyα cooling typically happens in sub-cycles, the temperature in those streams is usually underestimated to some degree, resulting in underestimated Lyα emissivities.

We have dug ourselves out of this hole with a simple solution, namely to re-start the production simulations at the moment of interest, in this case redshift 3, with hydrodynamic time-steps that are forced to be shorter than the shortest cooling time in the box, in order to get rid of any thermochemistry sub-cycling, and thus correctly model the competition between heating and cooling. By checking the largest number of thermochemistry sub-cycles in the box, we find the necessary time-step length to be on the order of one-thousandth of the Courant-condition hydrodynamical time-step, so we post-process with a hydrodynamical Courant-factor of about 10^{-3}. We let this run for a few thousand fine-scale time-steps, and verify after the fact that the total Lyα emissivity converges to a value, which we infer is the 'correct' one.

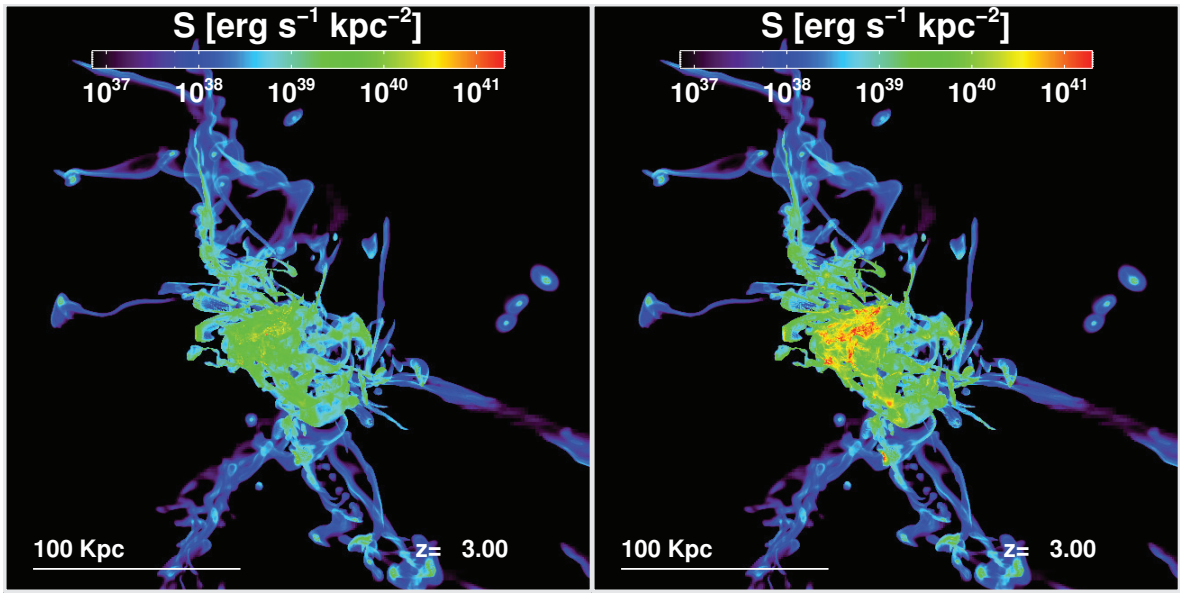
Fig. 5.9a shows how the Lyα luminosity of the targeted halo in the H2 production simulation rises and converges over about 2500 shortened fine-scale time-steps during post-processing. The total luminosity rises by almost a factor of four, so OS indeed has a non-negligible effect on the Lyα luminosity. In Fig. 5.9b we plot the Lyα luminosities of all zoom halos in the production simulations, before and after small-timestep-post-processing, with red diamonds and blue X's, respectively.

In Fig. 5.10 we show where the Lyα emissivity is most affected by OS: We show Lyα emissivity maps of the targeted halo in the H2 production simulation, before (left) and after (right) post-processing. If these maps are compared with a density map of the same halo (Fig. 1 in RB12) it can be seen that the regions where the Lyα emissivity changes most correspond to the densest gas. This makes sense, as the cooling time scales approximately inversely with the density squared. Note that the densest regions also are the most Lyα luminous ones, hence the Lyα brightest regions are those most affected by the operator splitting effect.



(a) *(b)*

Figure 5.9: *Operator splitting and Lyα emissivities:* ***(a)*** *Increase in the Lyα luminosity during small time-step post-processing of the targeted halo in the H2 production simulation (halo mass of $3\,10^{12}\,M_\odot$ at redshift 3).* ***(b)*** *Halo luminosities in the three production simulations at redshift 3 before the small time-step post-processing (red diamonds) and after (blue X's).*



(a) *(b)*

Figure 5.10: *Operator splitting and Lyα emissivities: Comparison of rest-frame Lyα surface emissivities in the H2 production simulation, before* ***(left)*** *and after* ***(right)*** *the small time-step post-processing. The largest increase in emissivity is in luminous and compact regions, where the cooling time is the shortest.*

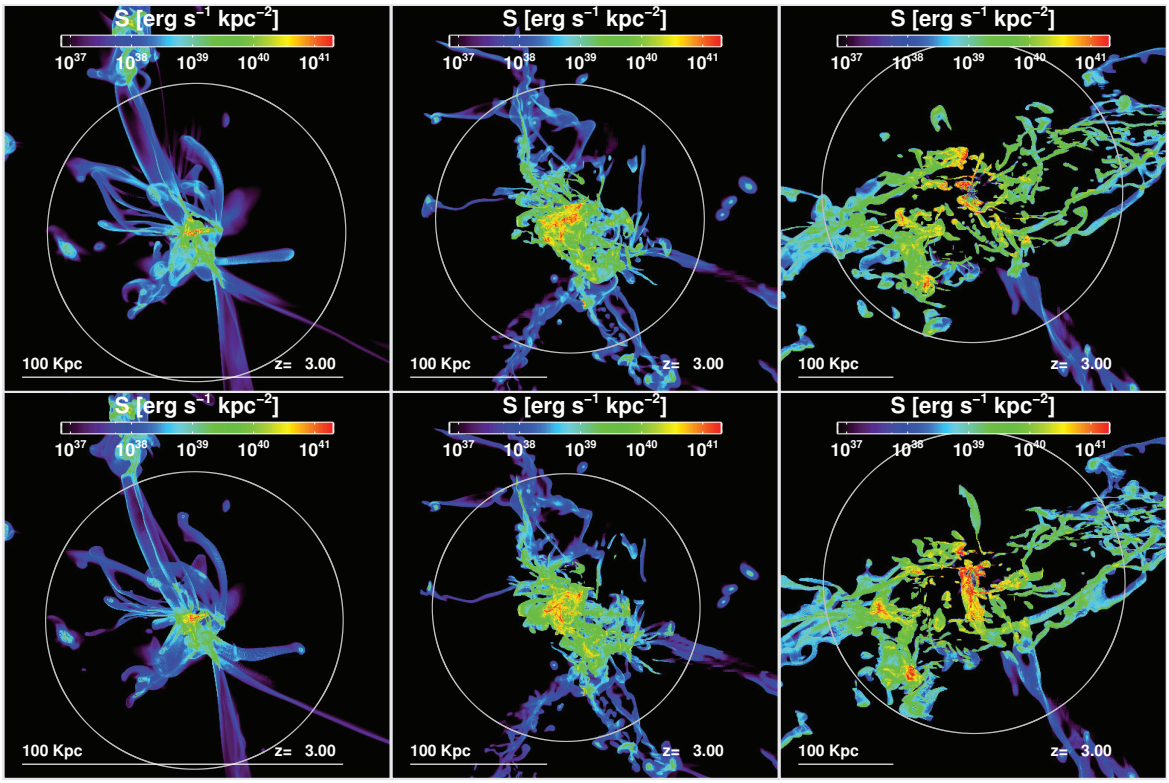


Figure 5.11: *Comparison of RT and a simple self-shielding approximation with a density threshold. The maps shows rest-frame* Lyα *surface emissivity in the H1, H2, and H2 simulations, from left to right. The* ***top*** *row shows the production simulations and the* ***bottom*** *row shows corresponding simulations with a self-shielding approximation instead of RT.*

5.2.5 Does a simple self-shielding approximation suffice?

In RB12 we discuss the validity of replacing RHD with a simple self-shielding approximation in hydrodynamics – the approach used by Faucher-Giguère et al. (2010) – where UV photoionization and heating are turned off at a density limit of $\sim 10^{-2}$ cm^{-3}. If this gives results in terms of Lyα emissivities that are as robust as RHD, there is much to gain in computing time.

In the paper we considered the densities at which gas becomes self-shielding at $z = 3$ in RHD and found it nicely matches the Faucher-Giguère et al. (2010) limit. We also argued that since the UV contribution to observable Lyα emissivities is negligible in our simulations, the simple self-shielding approximation should be quite adequate for our purposes, though this would likely break down in the limit where UV fluorescence and photo-heating becomes important, i.e. if one were to add stellar UV feedback. We present here a simple verification of this, which is just to show that we get the same Lyα emissivities and luminosities from analogues of our production simulations, by skipping RT altogether and instead applying the self-shielding approximation.

Fig. 5.11 shows Lyα emissivity maps at redshift 3 from these two sets of runs, with the H1, H2 and H3 simulations from left to right, respectively. The top row shows the

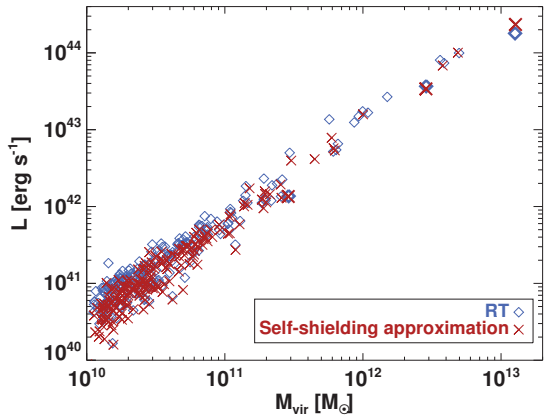


Figure 5.12: *Comparison of halo Lyα luminosities using RT (blue diamonds) and a simple self-shielding approximation (red X's).*

production runs, with RT and self-consistently modelled self-shielding, while the bottom row shows analogues where the propagation of UV photons is turned off and photoionization and heating is instead applied in all gas cells with $n_{\rm H} \leq 10^{-2}$ cm^{-3}, using the same UV background model (Faucher-Giguère et al., 2009). Since these analogue simulations are free of RT, they are cheaper by a factor 5-10 than the production runs. It is clear from these maps that the resulting morphologies and Lyα emissivities alike are very similar between the two versions.

Fig. 5.12 presents more quantitative evidence, in the form of halo luminosities. Here again, we see that there is little difference between using RT and the self-shielding approximation. We can conclude that the production runs presented in our paper are overkill – RHD is not really necessary to predict the Lyα emissivities of isolated gravitational processes in and around massive halos! This is an important result in itself, though: It is a verification that RHD is not necessary in this context, and allows us to cut back on computing time in future simulations of similar purpose.

Alas, this doesn't mean that we won't be using RT in future work on this subject. One of the main upcoming prospectives is to study the effect of stellar UV feedback on Lyα emissivity through photo-fluorescence and photo-heating, and here a simple self-shielding approximation won't do, unless we find (somewhat unexpectedly) that stellar UV feedback has a negligible effect on extended Lyα emission.

5.2.6 Are disrupted streams a numerical effect of over-resolved gravity?

To resolve accretion streams to the maximum available resolution in our simulations, we have included an unconventional refinement criterion that works on inter-cell gradients in the hydrogen neutral fraction. Usually, AMR refinement is only set to 'follow the mass' in simulations, but with our added criterion there exist maximally refined regions that nontheless contain very little mass. There is a danger associated with this, which is that we inadvertently over-resolve the gravitational potential in these highly resolved yet under-dense regions, because by default, the gravitational potential in `Ramses` is solved on the

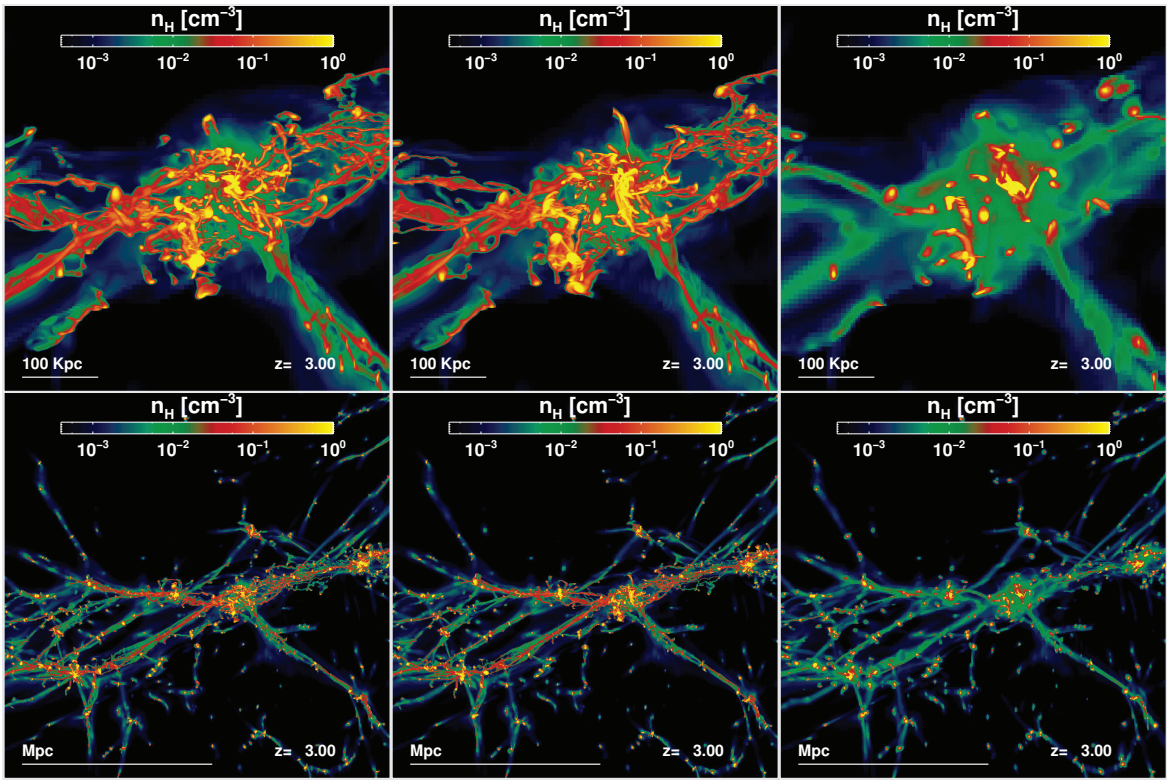


Figure 5.13: *Stream disruption in simulations. Gas density maps of three versions of the H3 simulation at redshift 3. The lower row shows wider zooms of the upper row maps.* ***Left:*** *The production H3 simulation, i.e. with stream refinement.* ***Middle:*** *An equivalent simulation, except that the gravitational force is smoothed to a maximum AMR level of 12 out of 14. (Also, instead of RT, a self-shielding approximation is used with $n_{\rm H}^{UV} = 10^{-2}\,{\rm cm}^{-3}$.)* ***Right:*** *A normal* `Ramses` *simulation started from the same initial conditions and run with the same AMR levels as the production run. This is thus a run with equilibrium thermochemistry, without RT, without self-shielding, and without stream refinement.*

existing AMR grid. This may result in artificial fragmentation in these over-resolved regions, where the grid resolution is finer than the density spatial sampling of particles. As our simulations show very fragmented streams around massive halos, this is cause for concern. Are the fragmented streams a physical effect that arises in our simulations because the resolution allows it, or are they a numerical effect caused by over-resolved gravity?

It is quite hard to tell, save for going for a full increase in the DM resolution and turning off the new refinement criterion – which is unaffordable. Instead we have gone around the problem, by keeping the new criterion, but forcing down the resolution of the gravitational solver. This can easily be done in `Ramses` with a parameter that states the maximum AMR level for gravity. We have run an analogue to the H3 simulation with this level set to 12 – two levels lower than the optimal refinement are redshift 3, so the potential is calculated on cells that are at least four times wider than the highest AMR resolution allows. The idea is that this should to some extent prevent the artificial fragmentation of streams, but if the fragmentation is physical it should stay much the same.

The results of this are presented in Fig. 5.13, in the form of gas density maps of the main

halo (top row) and its surroundings (bottom row). The leftmost maps show the production run and the middle row the analogue simulation with smoothed gravitational potential. On comparison with the production run, the streams are almost exactly the same, which makes us inclined to think the disrupted streams are physical and not numerical. On the far right we have another analogue simulation where the stream refinement criterion is altogether turned off. Here we see that the streams become much smoother and a lot of the structure inside and close to the halo disappears due to the lack of resolution. On the large scale, however, parts of the streams can be seen to remain disrupted, e.g. to the east of the main halo.

It is an unexpected and exciting bonus feature of our simulations that we see all this structure in accretion streams and large scale filaments that has gone undetected so far in simulations due to lack of resolution, and we intend to pursue an investigation of the nature of these structures in future work.

5.2.7 Smooth vs non-smooth RT

We wrap up the discussion of the RB12 paper by showing the effect of applying the smooth RT trick, discussed in §3.2.5, the idea of which is to reduce the thermochemistry-sucbycling and thus the computational load by 'fixing' the discreteness which is introduced by operator-splitting the RT steps of transport, injection, and thermochemistry. We consider here *(a)* the gain in computational speed that we get by applying this trick and *(b)* whether it has a physical effect on the dynamics and results.

For this purpose, we have made two low-resolution (by one level) analogue runs of the H1 simulation; one with the smoothing trick applied – which is the default, and the option used in the production runs – and the other without smoothing. We plot the results in Fig. 5.14 in the form of phase diagrams of photon flux (top row) and neutral hydrogen fraction (bottom) as a function of gas density. The neutral fraction diagrams are pretty much identical, but there is a clear difference in the photon fluxes: The smoothing has the effect of allowing photons to propagate further through dense gas, and the photon extinction is more gradual than when there is no smoothing. This makes perfect sense: Introducing the photon densities and fluxes as source terms in the thermochemistry solver, that are added gradually in each subcycle, has the effect that their depletion is less catastrophic than if the whole bunch is added in the first subcycling step, the result being that there is more likely a little bit left when the last step is finished.

However, considering that the y-axis of the phase diagram is logarithmic, the difference in the shielding is actually a very minor one. Where the photon flux is cut off by more than ten orders of magnitude in the non-smooth case, it is cut off by two or three orders of magnitude in the smooth case: One could approximately say that they are equivalent and the photon fluxes are completely cut off in both cases. Indeed, Fig. 5.15 shows that there is not notable difference in the Lyα luminosities of halos between using smooth and non-smooth RT.

And there is much to gain: The run-time of the smooth simulation is only about a quarter of that of the non-smooth simulation.

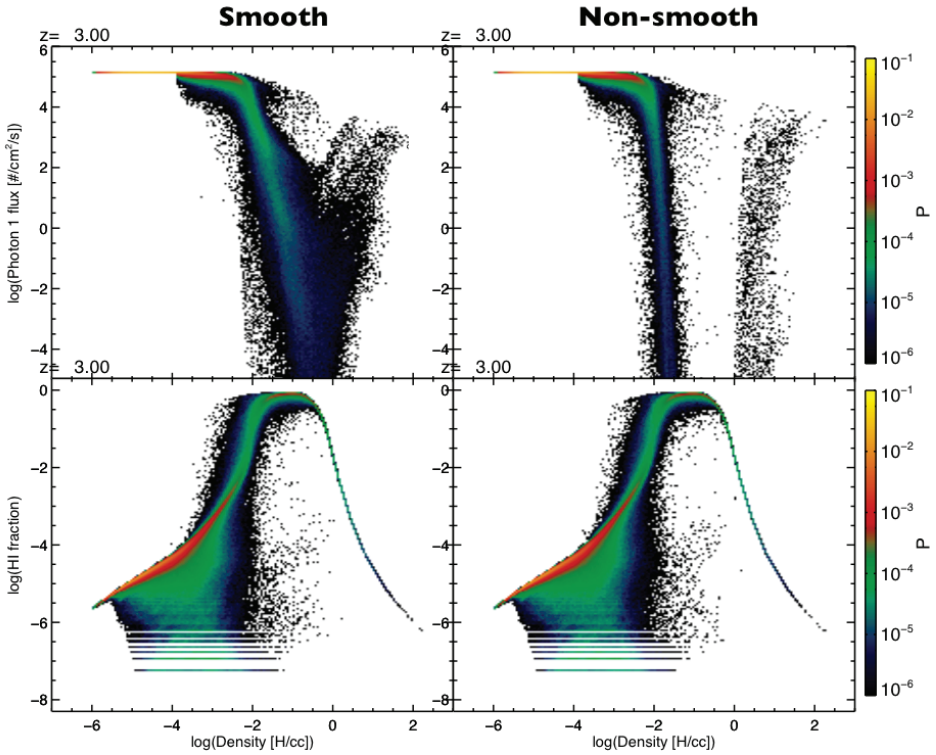


Figure 5.14: *Smooth RT: Comparison of photon fluxes* ***(top)*** *and* H I *fractions* ***(bottom)*** *at redshift 3 in low resolution equivalents of the H1 simulation that are identical save that the left one uses smooth RT and the right one does not.*

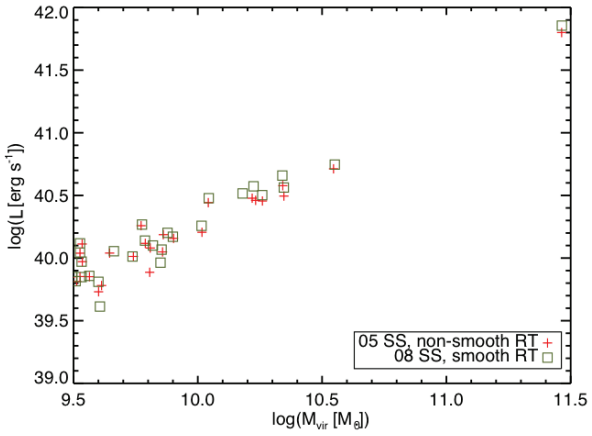


Figure 5.15: *Smooth RT: Comparison of halo luminosities at redshift 3 in low resolution equivalents of the H1 simulation that are identical save one uses smooth RT and the other does not.*

6
Conclusions and outlook

With the arrival of the James Webb space telescope and other upcoming instruments, we will soon start to glimpse the epoch of reionization, which marks the end of the dark ages of the Universe. It is high time to focus theoretical work on this epoch, to make predictions for what we will see and be prepared to interpret upcoming observations. The motivation for this thesis is to develop state-of-the art tools to simulate and study the evolution of galaxies from the birth of the Universe until reionization.

Galaxy evolution involves the complex interplay of many factors and scales that span orders of magnitude in space and time. Fortunately a handful of cosmological codes exist that include the most important factors (gravity, dark matter, hydrodynamics of gas, stars) and can deal with scale-differences. For my work I have chosen `Ramses` (Teyssier, 2002), which is among the best cosmological codes today. One factor has been missing from `Ramses` though, which bears particular relevance to the epoch of reionization – this is the radiative transfer (RT) of ionizing (UV) photons. UV feedback from stars may have an important self-regulating effect on the cooling of gas and subsequent star-formation, and is perhaps the source that powers the reionization of the Universe. UV radiation also has an impact on observable quantities like Lyα emission/absorption of gas. Due to complexity and computational cost, RT is traditionally not included in cosmological codes, though some have started to get RT implementations during the last decade.

Development of a new RHD code: `RamsesRT`

My major achievement during the PhD is the implementation of ionizing radiation hydrodynamics (RHD) inside the `Ramses` code. The implementation stands away from most others because it utilizes the so-called moment method with the M1 closure (Aubert and Teyssier, 2008), which involves working with a field of radiation rather than rays. The most common type of implementation involves ray-tracing, which can only deal with a limited number of point sources of radiation, whereas the moment method can deal with a virtually unlimited

number of sources and even spatially continuous emission from gas regions, which can be a great advantage. The M1 closure furthermore has the benefits over other closures that it is localized and still retains the directionality of radiation. The `RamsesRT` implementation is fully integrated into the AMR structure of `Ramses` and coupled with the hydrodynamics via non-equilibrium interactions with hydrogen and helium (diffuse emission, photoionization and heating). Multi-frequency is approximated with a handful of frequency bins. I have tested the code extensively – this includes all tests from 'The Cosmological RT Comparison Project' I and II (Iliev et al., 2006a, 2009), which are used for benchmarking RT and RHD codes.

The Lyα luminosity of cold streams and the nature of blobs

I have used `RamsesRT` to study the detectability of extended Lyα emission from cold accretion streams at redshift 3. These streams present an upgrade to the theory of how galaxies obtain their baryons, which traditionally involves only galaxy mergers and spherically symmetric hot gas accretion. They are routinely predicted by simulations, but have avoided the gaze of observers, either because they don't exist or are hard to detect. Simulation work has been done recently on their detectability, with contradictory results (e.g. Goerdt et al., 2010; Faucher-Giguère et al., 2010), but a presumably important factor is always missing – which is the propagation of ionizing photons. The Lyα emissivity of cold streams is very sensitive to the temperature and ionization state of the gas, which in turn is sensitive to self-shielding from UV radiation. RHD is therefore considered important to make correct predictions. I ran and analyzed cosmological `RamsesRT` simulations that include the propagation of the UV background and unprecedented resolution in accretion streams. The main conclusions of this work are that (i) cold streams are on the verge of being unambiguously observable with upcoming instruments via direct Lyα emission which is powered predominantly by gravitational heating and (ii) they are in principle capable of powering the mysterious Lyα blobs, which are the tip of extended Lyα emission focused on giant accreting groups/clusters of galaxies. An interesting result of this work is also that radiative transfer is actually not necessary for the UV background in this context – a much simpler self-shielding approximation is quite sufficient. However, when local UV radiation will later be added to the problem, RT will likely become important.

Ongoing and future work: Towards understanding early galaxy evolution

This thesis has a largely technical focus, but it serves as a springboard into future scientific studies of the interplay of photons and gas in a cosmological context. There are many interesting subjects to tackle in this field, which are made even more exciting by the advent of next generation instruments.

Lyα emission

My work on **extended Lyα emission** at redshift 3 has so far focused on the efficiency of gravitational heating in accretion streams. This is hailed as one of four possible models of what powers Lyα blobs. The other three are (i) scattering in the IGM of Lyα photons

originating from galaxies, (ii) UV photo-fluorescence powered by nearby galaxies/AGN, and (iii) cooling radiation in dense extended shells of gas formed by galactic winds. With simulations analogous to the ones I have been running, but with added components, I intend to investigate each of these models. In addition to UV radiative transfer, passive Lyα radiative transfer is also important here to model scattering effects. This can be done with a line RT code, `MCLy`α, which has been developed by my colleague Anne Verhamme (Verhamme et al., 2006).

With Anne, I also plan to work on **compact Lyα emission** from galaxies. The spectrum in the vicinity of the Lyα line – the Lyα line profile – has the potential to give us important information about the emitting galaxy. Interpreting the profile requires detailed modelling, of the complex scattering process, of the scattering medium, and the Lyα emission sources. Anne has applied her `MCLy`α code to `Ramses` simulation outputs to study and interpret the line profiles that emerge from galaxies (e.g. Verhamme et al., 2008, 2009). As a next step, I will work with her on processing Lyα transfer through `RamsesRT` simulation outputs which accurately model the interstellar gas ionization state resulting from local UV feedback and make it possible to identify the sites at which Lyα photons originate via UV photo-fluorescence. We have done preliminary checks and it indeed appears that the non-homogeneous UV component is vital in reproducing observable properties of Lyα line profiles of galaxies.

Early galaxies and cosmic reionization

The main theme I want to study here is the role of UV feedback in early galaxy evolution and the reionization of the Universe. The team of Adrianne Slyz and Julien Devriendt at Oxford University runs and analyzes a suite of pc-scale resolution `Ramses` zoom simulations they call the Nut (e.g. Powell et al., 2011; Geen et al., 2012), into which they are incorporating `RamsesRT` runs. Fig. 6.1 shows redshift 9 maps from the `RamsesRT` run Nut-RHD, where the IGM is in the process of being ionized by young stars. So far their preliminary investigation on the Nut-RHD has focused on the effect of stellar UV feedback on star formation and the inflow of cold gas at very high redshift. I hope to continue working with them and my supervisor Jeremy Blaizot on running and analyzing large cosmological RHD simulations and on my part focus on the first gigayear of galaxy evolution. These simulations should also go into the work I do with Anne Verhamme on studying Lyα scattering in galaxies.

Reionization powered by gas accretion is a novel idea recently explored by Dopita et al. (2011). They propose to solve the problem of an apparent impotency of stars and quasars to ionize the Universe at early redshift, by employing instead the UV radiation produced by accretion flows as they crash onto galactic disks. They use analytics and 1d simulations to conclude that sufficient ionizing radiation should be produced by this mechanism to ionize the Universe by $z \sim$ 7-8. They call for the idea to be further investigated with 3d simulations. `RamsesRT` gives a unique advantage to carry out this work because it includes UV emission from gas and not just point sources, thus allowing for a direct study of the abundance of UV photons produced in this way and how effective they are at ionizing the IGM.

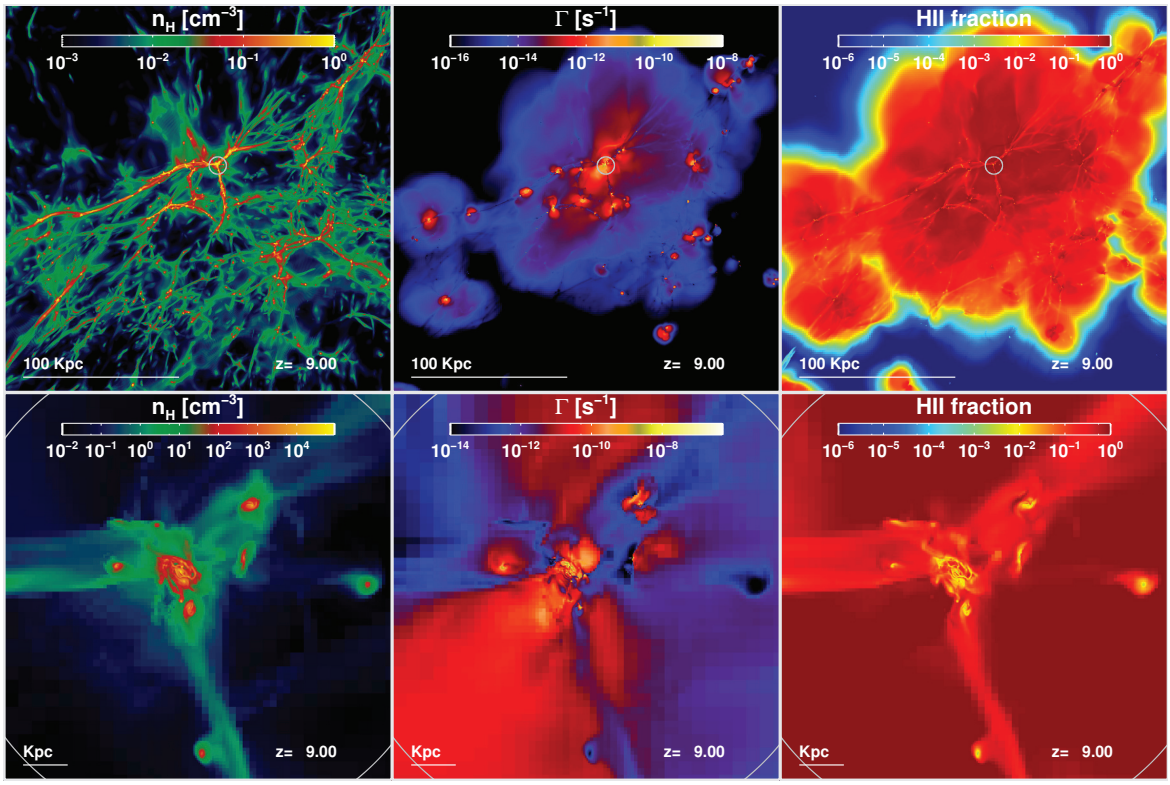


Figure 6.1: *Redshift 9 maps from the* `RamsesRT` *run Nut-RHD cosmological simulation, where stellar particles are used as radiative sources. The bottom row zooms in on the main halo ($6.5\,10^9\ M_\odot$) from the upper row maps, which is indicated by circles marking the virial radius (5.5 kpc). The length scales represent physical distance. The maximum resolution here is about 10 physical pc.* ***Left:*** *Maxima of gas density along the line-of-sight (LOS).* ***Middle:*** *LOS mass weighted average of hydrogen photoionization rate, $\Gamma = c_r \sum_i N_i \bar{\sigma}_{i\mathrm{HI}}$, showing radiation from the stellar sources.* ***Right:*** *LOS mass weighted average of the ionized hydrogen fraction.*

Improvements of physical models

Stellar feedback happens on scales well beyond the current resolution limits of cosmological simulations. To incorporate this, simulations adopt sub-resolution recipes where thermal and/or kinetic energy is injected into star-formation sites. This is somewhat ad-hoc and typically gives ineffective feedback because of insufficient resolution. I have started work with Sam Geen with the objective of improving stellar feedback recipes, including the effects of SNe, UV radiation and stellar winds. In the spirit of Ceverino and Klypin (2009), the idea is to start with semi-idealized simulations on scales that resolve relevant processes and then increase the scale in steps while keeping the effect of stellar feedback the same.

Metal cooling plays an important part in allowing population I and II stars to form, and the effect of soft X-rays has been suggested as important in slowing down star-formation and solving the so-called over-cooling problem that appears in simulations (Cantalupo, 2010). `RamsesRT` has not yet been equipped with metal cooling that takes into account the local effect of UV and X-ray radiation. Rather the metal-cooling rate assumes a homogeneous Haardt & Madau UV background model everywhere. A more consistent implementation of

(equilibrium) metal cooling has been broadly drawn up by me and Stephanie Courty, and we intend to continue with this.

Molecules are essential ingredients in the formation of the first stars (e.g. Gnedin et al., 2009), and the formation and destruction of those is sensitive to UV radiation. Inclusion of molecules could come naturally as an extension to `RamsesRT`, and their addition is made easy by the explicit solvers already in place.

Radiation pressure, where the momentum of photons is transferred onto atoms and/or dust, is another effect of UV feedback that likely has an important effect on regulating star formation, on the gas budget, on shaping the ISM, on outflows, and the initial mass function (e.g. Krumholz and Matzner, 2009; Fall et al., 2010; Hopkins et al., 2011b; Lopez et al., 2011; Murray et al., 2011; Krumholz and Thompson, 2012). It is therefore of great interest to add momentum transfer to `RamsesRT`, and it can also be done rather straightforwardly. I have already started preliminary work on the technical aspect, where I only consider the direct transfer of momentum onto H and He atoms, and neglect dust absorption for the time being.

Coupling `Aton` and `RamsesRT` for AMR reionization

`RamsesRT` is unfortunately not very well equipped to simulate large-scale reionization, due to the reduced speed of light approximation. However, it is complementary to the `Aton` GPU code (Aubert and Teyssier, 2008, 2010), which is fast enough to be able to use a full light-speed but not good at doing so in an AMR grid. An interesting idea is to mesh these implementations together to couple the large and small scales directly. Here, the codes could work in unison on the same simulation, with `RamsesRT` working on the fine (refined) scales and `Aton` on the coarse scale linking together the refined islands. The radiation field could be transferred between the codes whenever it crosses from coarse to refined, or vice versa. Doing this would likely give us a unique advantage over other reionization simulations in terms of resolution, and a correct light speed, which our RT tests hint at being vital to produce correct results.

A

The RT equation in co-moving coordinates

The radiative transfer equation describes the transport of photons through space. A nice derivation of it can be found in Gnedin and Ostriker (1997) (and later Petkova and Springel, 2009): We let $f_\gamma(\mathbf{x}, \mathbf{p}, t)$ be the distribution function for photons in comoving coordinates $\mathbf{x}$ and comoving momentum

$$\mathbf{p} = a\frac{h\nu}{c}\mathbf{n}, \tag{A.1}$$

where a is the cosmological scale factor, h is the Planck constant, ν the photon frequency, c the speed of light, and $\mathbf{n}$ a unit vector in the direction of photon propagation. The comoving coordinate system expands along with the Universe and a proper (or physical) distance can be found by $dr = a\,dx$. The total number of photons in the Universe is

$$N_\gamma = \int f_\gamma \, d^3x \, d^3p, \tag{A.2}$$

which changes in time only due to processes of photon emission and absorption. The continuity equation for f_γ is thus

$$\frac{df_\gamma}{dt} = \frac{\partial f_\gamma}{\partial t} + \dot{\mathbf{x}} \cdot \frac{\partial f_\gamma}{\partial \mathbf{x}} + \dot{\mathbf{p}} \cdot \frac{\partial f_\gamma}{\partial \mathbf{p}} \quad = \; - \left.\frac{\partial f_\gamma}{\partial t}\right|_{\text{absorption}} + \left.\frac{\partial f_\gamma}{\partial t}\right|_{\text{emission}}, \tag{A.3}$$

where $\dot{\mathbf{x}}$ and $\dot{\mathbf{p}}$ are time derivatives. Rather than the f_γ, it is customary to use the radiation specific intensity, $I_\nu(\mathbf{x}, \mathbf{n}, t)$, where

$$I_\nu \, d\nu \, d\Omega \, dA \, dt$$

is the energy of photons with frequency over the range $d\nu$ around ν propagating through the area dA in a solid angle $d\Omega$ around the direction $\mathbf{n}$, i.e.

$$I_\nu(\mathbf{x}, \mathbf{n}, t) = h\nu \; f_\gamma(\mathbf{x}, \mathbf{n}, t) \; \frac{d^3x \, d^3p}{d\nu \, d\Omega \, dA \, dt}. \tag{A.4}$$

Using the identities $a^3\ d^3x = c\ dt\ dA$ and $d^3p = p^2\ dp\ d\Omega$, we can convert from f_γ to I_ν with

$$I_\nu = \frac{h^4\ \nu^3}{c^2} f_\gamma. \tag{A.5}$$

By substituting (A.5) into (A.3), the radiative transfer equation is obtained:

$$\boxed{\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\mathbf{n}}{a}\cdot\nabla I_\nu - \frac{H}{c}\left(\nu\frac{\partial I_\nu}{\partial \nu} - 3I_\nu\right) = -\kappa_\nu I_\nu + \eta_\nu,} \tag{A.6}$$

where $H \equiv \dot{a}/a$ is the Hubble expansion rate, $\kappa_\nu(\mathbf{x}, \mathbf{n}, t)$ is the absorption coefficient and $\eta_\nu(\mathbf{x}, \mathbf{n}, t)$ is the source function. In the last step we used the following identities, which can be straightforwardly verified:

$$\begin{aligned}
\dot{\nu} &= -\nu H, \\
\dot{\mathbf{p}} &= 0, \\
\dot{\mathbf{x}} &= \frac{c}{a}\mathbf{n}, \\
\frac{\partial f_\gamma}{\partial x} &= \frac{c^2}{h^4\nu^3}\frac{\partial I_\nu}{\partial x}, \\
\frac{\partial f_\gamma}{\partial t} &= \frac{c^2}{h^4\nu^3}\left(\frac{\partial I_\nu}{\partial t} - \nu H\frac{\partial I_\nu}{\partial t} + 3I_\nu H\right).
\end{aligned}$$

According to (A.6), the change in beam intensity, in a fixed point, is described by the propagation of the beam (second LHS term), the cosmological effects of wavelength stretching and volume dilution (LHS term in parentheses), which together make the radiative energy density dilute as a^{-4}, and the baryonic interactions of photon absorption and emission (RHS). As we discussed is §2.1.2 and §3.3, the cosmological terms in parentheses are usually ignored in RT and RHD implementations. As discussed in §3.3, we do keep it *partially* in our implementation, as we allow the number density of photons to dilute with the co-moving coordinates, as a^{-3}.

B

On the propagation speed of ionization fronts

The main justification for using the reduced speed of light approximation (RSLA) is that the limiting speed in RT and RHD problems is not the speed of light but rather the speed of ionization fronts, or I-fronts (see §3.2.4). The light speed can thus in theory be reduced to any value c_r which is still faster than the fastest I-front in a simulation, without affecting the results. In the end, our choice of c_r is motivated by practical constraints (to get a manageable computing time), but it is nontheless good to have an idea of how it compares against I-front speeds.

An analytic expression for the I-front speed can easily be derived for the Strömgren sphere setup (Strömgren, 1939; Spitzer, 1978). Although the setup is idealized, we have seen in some of the tests described in Chapter 4 that the Strömgren description is not far from (simulated) reality even if cooling/heating and gas dynamics are turned on.

We shall first derive the relevant quantities; the Strömgren radius and the speed and position of the I-front as a function of time. Then we will use these expressions to study the I-front speed and dimensions for a few choices of density and photon emission rate. This should give a general idea of what values of c_r can be used in a given simulation.

B.1 The Strömgren sphere model

The Strömgren radius

We picture a source isotropically emitting ionizing photons at a rate $\dot{N}$ into a homogeneous hydrogen-only medium of number density n_{H}. We fix the temperature at a constant value: A bit above 10^4 Kelvin is typical in photoionized hydrogen gas. We assume equilibrium has been reached, i.e. that the I-front has reached a stable Strömgren radius r_{S}. Then the number of recombinations within the ionized sphere equals the central photon emission rate, i.e.

$$\dot{N} = \frac{4}{3}\pi\, r_{\mathrm{S}}^3\, \alpha^{\mathrm{B}}(T)\, n_{\mathrm{H}}^2,$$

where α^{B} is the recombination rate, and we assumed that all gas within r_{S} is 100% ionized, i.e. $n_{\mathrm{e}} = n_{\mathrm{HII}} = n_{\mathrm{H}}$. We solve for the Strömgren radius:

$$r_{\mathrm{S}} = \left(\frac{3\dot{N}}{4\pi\alpha^{\mathrm{B}} n_{\mathrm{H}}^2} \right)^{1/3}. \tag{B.1}$$

Time development of the Strömgren I-front

We establish a frame which moves with the expanding I-front, which has radius r_{I} and speed v_{I} from the emitting source. The gas is completely ionized for $r < r_{\mathrm{I}}$ and completely neutral for $r > r_{\mathrm{I}}$. In this frame, the flux $F(r_{\mathrm{I}})$ of ionizing photons from the ionized side must equal the flux of neutral atoms from the other, i.e.

$$F(r_{\mathrm{I}}) = n_{\mathrm{H}}\, v_{\mathrm{I}}.$$

The photon flux at r_{I} is reduced from the emission rate $\dot{N}$ due to geometric dilution and the absorption of the in-between gas:

$$F(r_{\mathrm{I}}) = \frac{\dot{N} - 4/3\pi r_{\mathrm{I}}^3 n_{\mathrm{H}}^2 \alpha^{\mathrm{B}}}{4\pi r_{\mathrm{I}}^2}.$$

Equating the fluxes of photons and atoms from each side of the I-front and switching to the rest-frame then gives the front speed as a function of radius:

$$v_{\mathrm{I}} = \frac{dr_{\mathrm{I}}}{dt} = \frac{1}{3t_{\mathrm{rec}}} \frac{r_{\mathrm{S}}^3 - r_{\mathrm{I}}^3}{r_{\mathrm{I}}^2}, \tag{B.2}$$

where we have defined $t_{\mathrm{rec}} = (n_{\mathrm{H}}\alpha^{\mathrm{B}})^{-1}$. Moving the dt across, integrating on both sides and solving for r_{I} gives

$$r_{\mathrm{I}}(t) = r_{\mathrm{S}} \left[1 - e^{-t/t_{\mathrm{rec}}} \right]^{1/3}. \tag{B.3}$$

Inserting this expression into (B.2) then gives

$$v_{\mathrm{I}}(t) = \frac{r_{\mathrm{S}}}{3t_{\mathrm{rec}}} \frac{e^{-t/t_{\mathrm{rec}}}}{\left(1 - e^{-t/t_{\mathrm{rec}}}\right)^{2/3}}. \tag{B.4}$$

Note that in these derivations we have implicitly assumed an infinite speed of light.

B.2 The speed of I-fronts for a range of luminosities and gas densities.

In Fig. B.1 we present the idealized I-front speed v_{I} as a function of distance r_{I} from the radiative source, according to Eq. B.2, for a range of gas densities and source luminosities. A horizontal line in each plot marks the speed of light. The Strömgren model does not take the speed of light into account at all, and we initially have $v_{\mathrm{I}} > c$, which is clearly unphysical. It is possible though that some ray-tracing RT implementations similarly manifest $v_{\mathrm{I}} > c$ I-fronts, since they effectively use an infinite light speed. At some point in the Fig. B.1 plots,

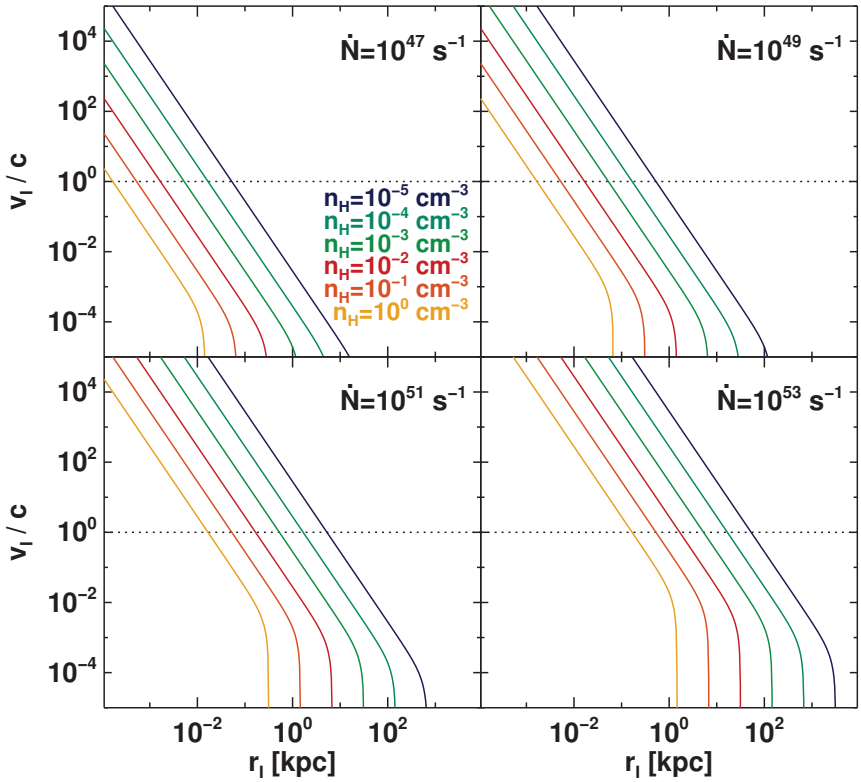


Figure B.1: *Idealized I-front speeds, in units of the light speed, for a range of densities and luminosities. The speed of the I-front, $v_{\rm I}$, is calculated from the Strömgren sphere model. Each of the four plots represents a source luminosity. Each of the solid colored lines represents a gas density. Horizontal dotted lines represent the light speed.*

there is a transition to $v_{\rm I} < c$, corresponding to the intersection of the solid and horizontal dotted lines. After this point the real speed of I-fronts should be well presented by the plots, as well as by infinite light speed RT implementations. However, if an RT implementation uses a reduced speed of light c_r, like we do, it starts out with a too slow I-front, $v_{\rm I} \sim c_r$, and only after the intersection of $v_{\rm I} \sim c_r$ will the simulated I-front speed be realistic. An implementation that uses the correct speed of light should get the I-front speed right at all points. Presumably this would be close to c in the initial stage, where the plots show $v_{\rm I} > c$ fronts and then it would follow the plot curves (roughly, since this is an idealized model).

To summarize:

- An infinite speed of light approximation presumably simulates a premature (too fast) I-front where $v_{\rm I} > c$ in the plots of Fig. B.1, but gets the speed right after that.

- A reduced speed of light approximation (RSLA) has the I-front speed initially lagging (too slow), but has it roughly correct at radii where $v_{\rm I} \lesssim c_r$ in the plots. The approximation thus roughly becomes OK at radii larger than where the $v_{\rm I}$ profile intersects with c_r.

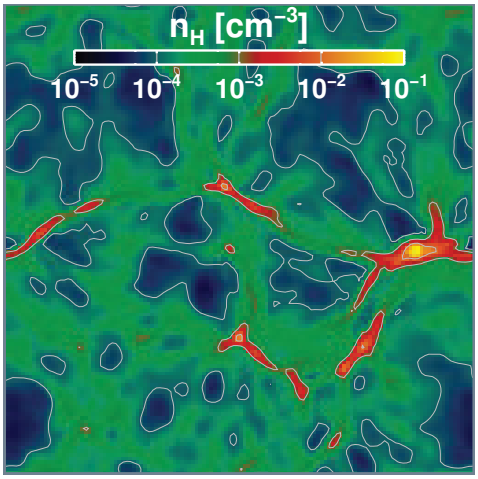


Figure B.2: *Gas density slice at i_z=64/128 from the cosmological field test from Iliev et al. (2006a) (test 4, see §4.3.5). The width of the (redshift 9) box is 71.5 physical kpc. Contours represent the densities 10^{-2}, 10^{-3}, and 10^{-4} cm^{-3}.*

- Simulating an I-front in `RamsesRT` with the correct light speed should presumably give us the speed profile of a 'real' I-front. This should be initially close to c, but roughly follow the plot profiles in Fig. B.1 where $v_{\rm I} < c$.

Let us now consider the propagation speed for three 'regimes': A reionization epoch cosmological scale, a galactic environment akin to those we have studied in our work on Lyα blobs, and an ISM environment.

A reionization scale

Here we can take for a finger-in-the-air reference the cosmological box we used in the 4th benchmark test from Iliev et al. (2006a) (Il06) and described in §4.3.5. We map a slice of the density field in Fig. B.2. From the map, we can determine the relevant density range, which is $n_{\rm H} \gtrsim 10^{-1}$ cm^{-3} in the dense regions immediately surrounding the radiative sources and $n_{\rm H} \gtrsim 10^{-5}$ cm^{-3} in the void regions. From the Il06 test setup we also determine the luminosities for the radiative sources, which in the test are in the range $\dot{N} \sim 10^{52} - 10^{53}$ photons s^{-1} for the 16 radiative sources in the box. We are thus concerned with the lower right plot in Fig. B.1.

If a RSLA is used with $f_c \sim 10^{-2}$, the intersection of light speed and $v_{\rm I}$ is at $\sim 100 - 500$ kpc at the void densities ($\sim 10^{-5} - 10^{-4}$ cm^{-3}). As this is larger than the cosmological box size presented in Fig. B.2, it is clear that such a low light speed will poorly represent reionization of these void regions, and likely this applies to reionization simulations in general. Even with densities as high as 10^{-2} cm^{-3} is the intersection radius ~ 10 kpc, which is still a considerable fraction of the box width and well above the resolution.

However, using an infinite light speed may also be a bad approximation: The front is likely too fast (faster than light), up to the point where the $v_{\rm I}$ profile intersects with c. At the void densities the I-front speed will be too fast for the first $\sim 20 - 50$ kpc, which is

comparable to the box width.

A galactic environment

We are concerned here with the inter-galactic environment surrounding galaxies. We take typical densities from our simulations of extended Lyα emission, where they are in the range $n_{\mathrm{H}} \sim 10^{0} - 10^{-3}$ cm^{-3}. Presumably the luminosities are somewhat less extreme than on the cosmological scale, so we take $\dot{N} = 10^{51}$ photons s^{-1}, just to pick a value, though the luminosity can of course vary dramatically both in space and time. Thus we are concerned with the lower left plot.

We look again at the intersections of light speed and v_{I} in the plots for an RSLA simulation with $f_c = 10^{-2}$: For $n_{\mathrm{H}} = 10^{-3}$ cm^{-3} the I-front speed is badly simulated up to $r_{\mathrm{I}} \sim 5$ kpc, which corresponds to roughly 20 optimal resolution widths in our Lyα emissivity simulations. For $n_{\mathrm{H}} = 10^{0}$ cm^{-3} this goes down to $r_{\mathrm{I}} \sim 0.1$ kpc, which is well within any realistic resolution limits. The 1% RSLA thus appears to be a fair to good approximation in this regime.

An infinite light speed also appears to be a good approximation here: For $n_{\mathrm{H}} = 10^{-3}$ the light speed intersects with v_{I} at $r_{\mathrm{I}} \lesssim 1$ kpc and for $n_{\mathrm{H}} = 10^{0}$ cm^{-3} it is at ~ 0.01 kpc.

An inter-stellar medium

The ISM is of course a multi-phase medium, so let's look at the two extremes of $n_{\mathrm{H}} \sim 10^{-2}$ cm^{-3} and $n_{\mathrm{H}} \sim 10^{2}$ cm^{-3}. The source luminosities should correspond to on the order of a hundred solar masses (again, just to pick a value), or $\dot{N} = 10^{49}$ photons s^{-1}.

We consider once again the RSLA with $f_c = 1\%$. In the diffuse ISM the intersection is at $\sim 10^{-1}$ kpc and for the dense ISM it is at sub-parsec scales. Again the approximation appears to be fair to good: Although the I-front speed is well modelled at smaller scales, the resolution is also better in ISM simulations than IGM ones.

C

Interaction rate coefficients adopted in RamsesRT

Here we collect the rate coefficients used in RamsesRT for hydrogen and helium interaction, which are fitted functions taken from various sources. These are, in order of appearance, collisional ionization rates, recombination rates, cooling rates (collisional ionization, recombination, collisional excitation, bremsstrahlung, Compton and dielectric recombination), and photoionization cross sections.

Collisional ionization rate coefficients

Those are in units of [$\mathrm{cm}^3\,\mathrm{s}^{-1}$] and are taken from Cen (1992), with temperature everywhere assumed in Kelvin (see Fig. C.1):

$$\beta_{\mathrm{HI}}(T) = 5.85\ 10^{-11}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-157\ 809.1/T}$$

$$\beta_{\mathrm{HeI}}(T) = 2.38\ 10^{-11}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-285\ 335.4/T}$$

$$\beta_{\mathrm{HeII}}(T) = 5.68\ 10^{-12}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-631\ 515/T}$$

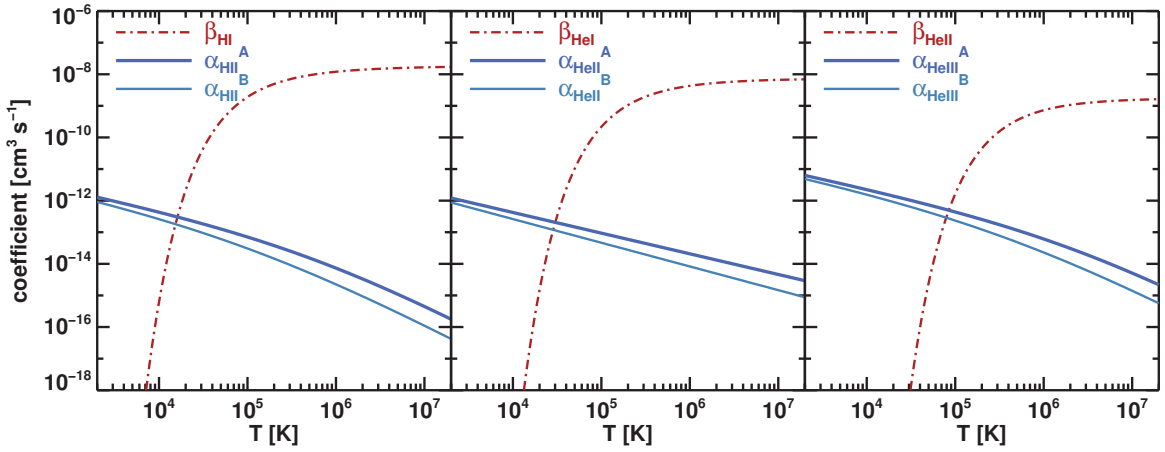


Figure C.1: *Collisional ionization and recombination rate coefficients used in* RamsesRT.

Recombination rate coefficients

These are all taken from Hui and Gnedin (1997). For readability, we use the following unitless functions:

$$\lambda_{\rm HI}(T) = \frac{315\ 614\ K}{T}$$
$$\lambda_{\rm HeI}(T) = \frac{570\ 670\ K}{T}$$
$$\lambda_{\rm HeII}(T) = \frac{1\ 263\ 030\ K}{T}$$

The coefficients are as follows, all in units of [cm^3 s^{-1}] (see Fig. C.1):

$$\alpha^{\rm A}_{\rm HII}(T) = 1.269\ 10^{-13}\ \frac{\lambda_{\rm HI}^{1.503}}{[1+(\lambda_{\rm HI}/0.522)^{0.47}]^{1.923}}$$
$$\alpha^{\rm A}_{\rm HeII}(T) = 3\ 10^{-14}\ \lambda_{\rm HeI}^{0.654}$$
$$\alpha^{\rm A}_{\rm HeIII}(T) = 2.538\ 10^{-13}\ \frac{\lambda_{\rm HeII}^{1.503}}{[1+(\lambda_{\rm HeII}/0.522)^{0.47}]^{1.923}}$$
$$\alpha^{\rm B}_{\rm HII}(T) = 2.753\ 10^{-14}\ \frac{\lambda_{\rm HI}^{1.5}}{[1+(\lambda_{\rm HI}/2.74)^{0.407}]^{2.242}}$$
$$\alpha^{\rm B}_{\rm HeII}(T) = 1.26\ 10^{-14}\ \lambda_{\rm HeI}^{0.75}$$
$$\alpha^{\rm B}_{\rm HeIII}(T) = 5.506\ 10^{-14}\ \frac{\lambda_{\rm HeII}^{1.5}}{[1+(\lambda_{\rm HeII}/2.74)^{0.407}]^{2.242}}$$

Cooling rate coefficients

The temperature used in these coefficients is assumed everywhere in Kelvin. The coefficients are plotted in Fig. C.2, and Fig. C.3 shows the resulting primordial cooling rate contributions (Eq. 3.60), normalized to $n_{\rm H}^2$, assuming collisional ionization equilibrium (CIE). Fig. C.4 then

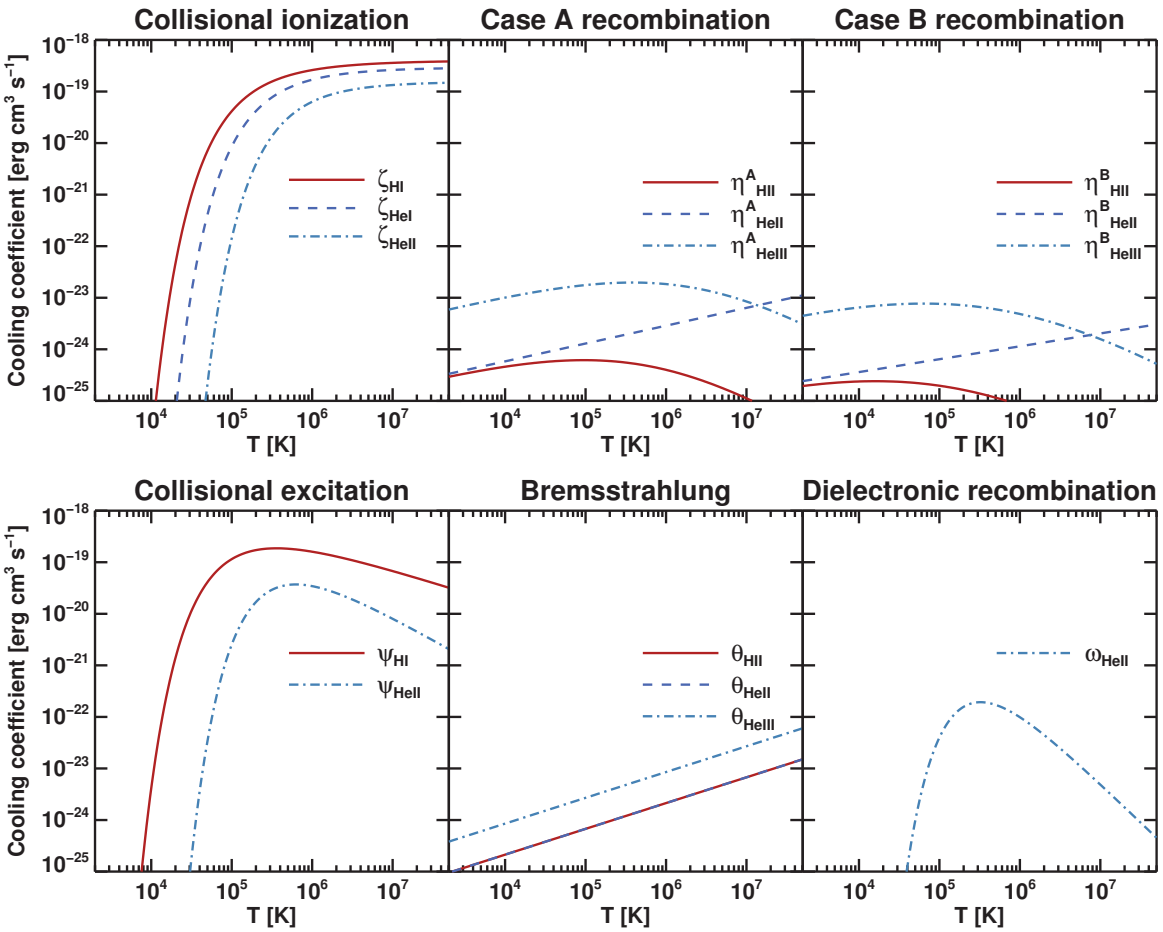


Figure C.2: *Primordial cooling rate coefficients used in* `RamsesRT`*, all in units of* $\mathrm{erg\,cm^3\,s^{-1}}$.

shows the total (normalized) primordial cooling rate assuming CIE, which is just a sum of the rates from Fig. C.3.

Collisional ionization cooling rate coefficients [$\mathrm{erg\,cm^3\,s^{-1}}$] (Cen, 1992):

$$\zeta_{\mathrm{HI}}(T) = 1.27\ 10^{-21}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-157\ 809.1/T}$$

$$\zeta_{\mathrm{HeI}}(T) = 9.38\ 10^{-22}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-285\ 335.4/T}$$

$$\zeta_{\mathrm{HeII}}(T) = 4.95\ 10^{-22}\ \sqrt{T}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-631\ 515/T}$$

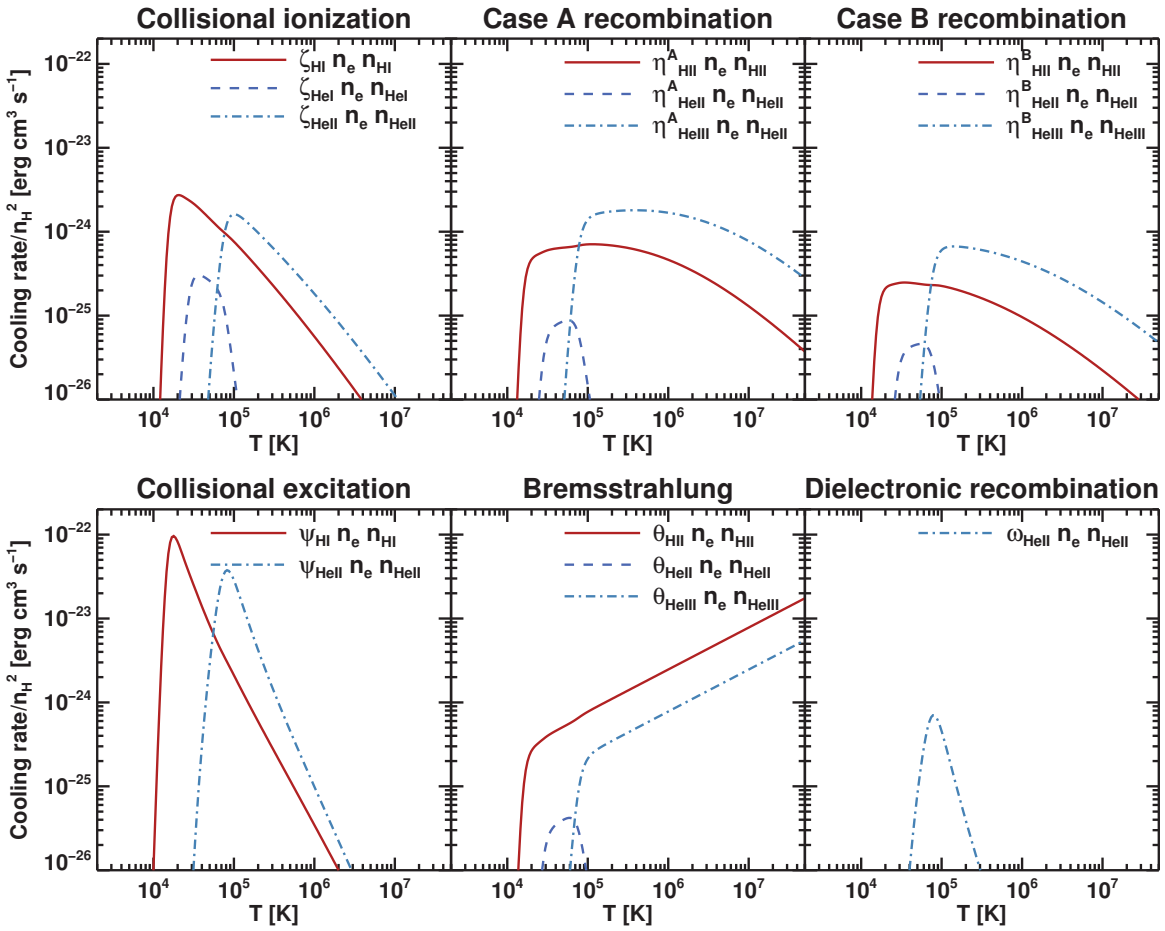


Figure C.3: *Primordial cooling rate contributions, assuming CIE abundances of the interacting species.*

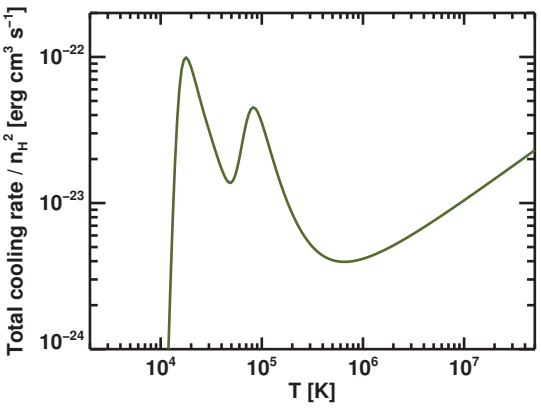


Figure C.4: *Total primordial cooling rate (sum of the rates from Fig. C.3, assuming CIE abundances, and assuming the OTSA, i.e. case B recombinations.*

Case A and B recombination cooling rate coefficients [erg cm^3 s^{-1}] (Hui and Gnedin, 1997):

$$\eta^A_{\rm HII}(T) = 1.778\ 10^{-29}\ T\ \frac{\lambda^{1.965}_{\rm HI}}{[1+(\lambda_{\rm HI}/0.541)^{0.502}]^{2.697}}$$
$$\eta^A_{\rm HeII}(T) = k_{\rm B}\ T\alpha^{\rm A}_{\rm HeII}\ =\ k_{\rm B}\ T\ 3\ 10^{-14}\ \lambda^{0.654}_{\rm HeI}$$
$$\eta^A_{\rm HeIII}(T) = 8\times 1.778\ 10^{-29}\ T\ \frac{\lambda^{1.965}_{\rm HeII}}{[1+(\lambda_{\rm HeII}/0.541)^{0.502}]^{2.697}}$$
$$\eta^B_{\rm HII}(T) = 3.435\ 10^{-30}\ T\ \frac{\lambda^{1.97}_{\rm HI}}{[1+(\lambda_{\rm HI}/2.25)^{0.376}]^{3.72}}$$
$$\eta^B_{\rm HeII}(T) = k_{\rm B}\ T\alpha^{\rm B}_{\rm HeII}\ =\ k_{\rm B}\ T\ 1.26\ 10^{-14}\ \lambda^{0.75}_{\rm HeI}$$
$$\eta^B_{\rm HeIII}(T) = 8\times 3.435\ 10^{-30}\ T\ \frac{\lambda^{1.97}_{\rm HeII}}{[1+(\lambda_{\rm HeII}/2.25)^{0.376}]^{3.72}}$$

Collisional excitation cooling rate coefficients [erg cm^3 s^{-1}] (Cen, 1992):

$$\psi_{\rm HI}(T) = 7.5\ 10^{-19}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-118\ 348/T}$$
$$\psi_{\rm HeII}(T) = 5.54\ 10^{-17}\ T^{-0.397}\ \left(1+\sqrt{\frac{T}{10^5}}\right)^{-1}\ e^{-473\ 638/T}$$

Bremsstrahlung cooling rate coefficients [erg cm^3 s^{-1}] (Osterbrock and Ferland, 2006):

$$\theta_{\rm HII}(T) = 1.42\ 10^{-27}\ \sqrt{T}$$
$$\theta_{\rm HeII}(T) = 1.42\ 10^{-27}\ \sqrt{T}$$
$$\theta_{\rm HeIII}(T) = 4\times 1.42\ 10^{-27}\ \sqrt{T}$$

Compton cooling/heating rate coefficient [erg s^{-1}] (Haiman et al., 1996):

$$\varpi(T,a) = 1.017\ 10^{-37}\ \left(\frac{2.727}{a}\right)^4\ \left(T-\frac{2.727}{a}\right)\ddagger$$

Dielectronic recombination cooling rate coefficient [erg cm^3 s^{-1}] (Black, 1981):

$$\omega_{\rm HeII}(T) = 1.24\ 10^{-13}\ T^{-1.5}\ e^{-470\ 000/T}\ \left(1+0.3\ e^{-94\ 000/T}\right)$$

Cross sections

Photoionization HI−, HeII− and HeIII− cross sections are taken from Verner et al. (1996) (via Hui and Gnedin, 1997) and are given in [cm^2] as a function of photon energy ϵ by

$$\sigma(\epsilon) = \sigma_0\ \left[(x-1)^2+y_w^2\right]\ \frac{y^{0.5P-5.5}}{(1+\sqrt{y/y_a})^P}\ ,\qquad \text{if } \epsilon \geq \epsilon_{ion}, \tag{C.1}$$

(and 0 cm^2 otherwise), where

$$x \equiv \frac{\epsilon}{\epsilon_0} - y_0,$$

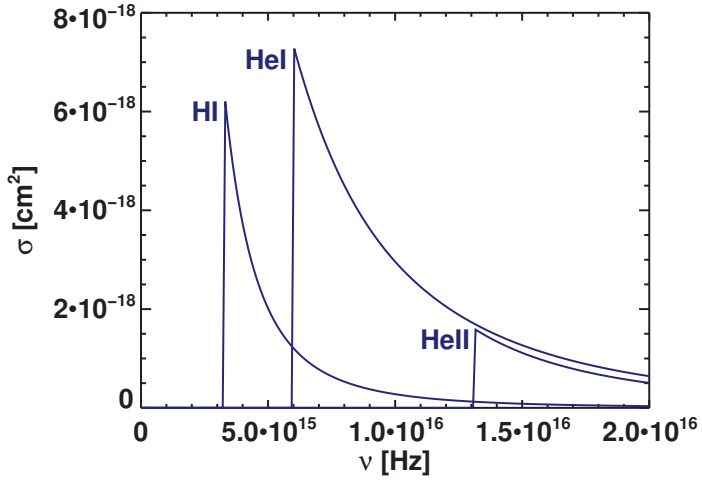


Figure C.5: *Cross sections used in* RamsesRT*, taken from Verner et al. (1996).*

and

$$y \equiv \sqrt{x^2 + y_1^2},$$

and the fitting parameters σ_0, ϵ_0, y_w, P, y_a, y_0, and y_1 are given in Table C.1. The ionization energies ϵ_{ion} and corresponding frequencies ν_{ion} are given in Table C.2. Fig. C.5 shows a plot of the cross sections.

Table C.1: *Photoionization cross section parameters – see Eq. C.1*

Ion species	ϵ_0 [eV]	σ_0 [cm^2]	P	y_a	y_w	y_0	y_1
HI	0.4298	$5.475\ 10^{-14}$	2.963	32.88	0	0	0
HeI	0.1361	$9.492\ 10^{-16}$	3.188	1.469	2.039	0.4434	2.136
HeII	1.720	$1.369\ 10^{-14}$	2.963	32.88	0	0	0

Table C.2: *Photoionization energies and corresponding frequencies*

Ion species	ϵ_{ion}	ν_{ion}
HI	$\epsilon_{\rm HI} = 13.60$ eV	$\nu_{\rm HI} = 3.288\ 10^{15}$ s^{-1}
HeI	$\epsilon_{\rm HeI} = 24.59$ eV	$\nu_{\rm HeI} = 5.946\ 10^{15}$ s^{-1}
HeII	$\epsilon_{\rm HeII} = 54.42$ eV	$\nu_{\rm HeII} = 1.316\ 10^{16}$ s^{-1}

D

RamsesRT user guide

Ramses simulation parameters are set with a number of Fortran namelists, that are described in the Ramses user guide. For RamsesRT, all the same namelists still apply, with the addition of two new ones, RT_PARAMS and RT_PACS, for setting up the radiative transfer and photon package properties. All parameters relevant to running RamsesRT simulations are in the namelists, except for the number of photon packages, which is a compilation parameter: To change the number one must recompile. The compilation is pretty straightforward: RamsesRT is a patch directory (rt) which should be included in the Makefile for compilation, but there is also a number of new source files specific to RamsesRT which must be explicitly named. An example of a Makefile is attached to the RamsesRT bundle for guidance.

The following tables describe the two namelists specific to RamsesRT.

Table D.1: RamsesRT *namelist* RT_PARAMS.

Variable name and default value	Description
X=0.76	Hydrogen mass fraction.
Y=0.24	Helium mass fraction.
rt_flux_scheme='glf'	Intercell flux function (§3.2.1). Use 'glf' or 'hll'.
hll_evals_file="	Eigenvalues file, necessary for the HLL flux (§3.2.1). Can also be set by environment variable RAMSES_HLLFILE.
rt_Tconst=-1	Constant temperature, in Kelvin, to assume for all temperature-dependent interaction rate (to run the first Iliev test). The default negative value means the actual cell temperature is used.
rt_freeflow=.false.	Free flow of photons (no interaction with gas).
rt_otsa=.true.	Assume the on-the-spot approximation (§3.2.3).
rt_c_fraction=1.	Reduced light speed fraction (§3.2.4).
rt_courant_factor=0.8	Courant factor for RT (§3.2.4).
rt_smooth=.true.	Use smooth RT (§3.2.5).
rt_star=.false.	Turn on photon emission from star particles. If rt_star=1 and rt=0 (from RUN_PARAMS), RT turns on only when the first stars are created.
sed_dir="	Directory for SED model (§3.6). This can also be set by the environment variable RAMSES_SED_DIR.
sedprops_update=-1	Frequency (per coarse timestep) of photon package updates according to SED model (§3.6). Default is never.
rt_esc_frac=1.	Escape fraction of photons from star particles (§3.6).
rt_is_init_xion=.false.	Initialize H and He ionization fractions from CIE, using temperature and density in each cell.
rt_is_outflow_bound = .false.	Force outflow boundary for RT on all box sides, regardless of how boundaries are defined for hydrodynamics. By default, boundaries are the same for RT and hydrodynamics.
upload_equilibrium_x =.false.	Set PIE ionization fractions when merging cells, instead of taking children averages (§5.1).

Table D.2: `RamsesRT` *namelist* `RT_PARAMS` *cont.: UV background parameters and RT refinement.*

Variable name and default value	Description
UV background	
`uv_file=”`	File containing UV model (§5.1). This can also be set by the environment variable `RAMSES_UV_FILE`.
`rt_uv_nhss=-1.`	Self-shielding density threshold for homogeneous UV background, set by `haardt_madau=.true.` (Chapter 5).
`rt_UVsrc_nHmax=-1.`	Density threshold for non-homogeneous UV emission (Chapter 5). Default value corresponds to no UV propagation.
`nUVpacs=npacs`	Number of photon packages dedicated to the propagated UV background. These are at the back of the photon packages: The first `npacs-nUVpacs do not carry the UV background (i.e. are for stellar feedback if turned on). Set to 0 if no UV propagation.`
RT refinement	

`rt_floor_n=1d-10`
`rt_floor_xHI=1d-10`
`rt_floor_xHII=1d-10`
	Discontinuity-based strategy: photon density and ionization fraction floor below which gradients are ignored (§5.1, §4.3.2).

`rt_err_grad_n=-1.0`
`rt_err_grad_xHI=-1.0`
`rt_err_grad_xHII=-1.0`
	Discontinuity-based strategy: photon density and ionization fraction gradients above which a cell is refined (§5.1, §4.3.2).

`rt_refine_aexp=-1.0`	Cosmological expansion at which to turn on RT refinement strategies (§5.1).

Table D.3: RamsesRT *namelist* RT_PARAMS *cont.: RT source regions.*

Variable name and default value	Description
RT source regions	
rt_nsource=0	Number of independent source (photon emission) regions in the computational box (see §4.3).
rt_source_type='square'	Geometry defining each source region. 'square' defines a generalized ellipsoidal shape with photons injected everywhere inside, 'shell' defines a finite width spherical shell into which photons are injected, and 'point' defines a point source.
rt_src_x_center=0.0 rt_src_y_center=0.0 rt_src_z_center=0.0	Coordinates of the center of each source region.
rt_src_length_x=0.0 rt_src_length_y=0.0 rt_src_length_z=0.0	Sizes in all directions of each source region. If a spherical shell is used, rt_src_length_x and rt_src_length_y represent outer and inner radius.
rt_exp_source=2.0	Exponents defining the norm used to compute distances for the generalized ellipsoid. rt_exp_source=2 corresponds to a spheroid, rt_exp_source=1 to a diamond shape, rt_exp_source$\geq$10 to a perfect square.
rt_src_pac=1	Photon packages into which photons are emitted in each source region (1 $\leq$ rt_src_pac $\leq$ M, where M is the number of photon packages).
rt_n_source=0.0 rt_u_source=0.0 rt_v_source=0.0 rt_w_source=0.0	Injection rates, in cgs units into N_i and $\mathbf{F}_i = (F_x, F_y, F_z)_i$.

Table D.4: `RamsesRT` *namelist* `RT_PARAMS` *cont.: Initialization regions.*

Variable name and default value	Description
RT initialization regions	
`rt_nregion=0`	Number of independent initialization regions in the computational box (see §4.3).
`rt_region_type='square'`	Geometry defining each initialization region. `'square'` defines a generalized ellipsoidal shape, while `'point'` defines a delta function.

`rt_src_x_center=0.0`
`rt_src_y_center=0.0`
`rt_src_z_center=0.0`
	Coordinates of the center of each initialization region.

`rt_src_length_x=0.0`
`rt_src_length_y=0.0`
`rt_src_length_z=0.0`
	Sizes in all directions of each initialization region.

`rt_exp_source=2.0`	Exponents defining the norm used to compute distances for the generalized ellipsoid. `rt_exp_source=2` corresponds to a spheroid, `rt_exp_source=1` to a diamond shape, `rt_exp_source`$\geq$`10` to a perfect square.

`rt_n_region=0.0`
`rt_u_region=0.0`
`rt_v_region=0.0`
`rt_w_region=0.0`
`rt_xion_region=0.0`
	Values in each region, in cgs units, for N_1, $\mathbf{F}_1 = (F_x, F_y, F_z)_1$ and x_{HII}.

Table D.5: RamsesRT *namelist* RT_PACS. *The default settings are* HI, HeI *and* HeII *ionizing photon packages (See §3.1) with cross section and energy properties derived from a blackbody spectrum with an effective temperature of* 10^5 *Kelvin.*

Variable name and default value	Description
`pac_L0 = 13.60, 24.59, 54.42`	Lower energy boundaries, in eV, of each photon package (see §3.1).
`pac_L1 = 24.59, 54.42, 0.0`	Upper energy boundaries, in eV, of each photon package (see §3.1). A value of `0.0` is used to represent ∞.

`pac_csn(1,:)=3.0d-18, 0.0, 0.0`
`pac_csn(2,:)=5.7d-19, 4.5d-18, 0.0`
`pac_csn(3,:)=7.9d-20, 1.2d-18, 1.1d-18`
	2d matrix representing average cross sections (cm^2) between each package (first index) and species (second index). These can either be set manually or left to RamsesRT to derive from SED models.

`pac_egy(1,:) = 17.4, 0.0, 0.0`
`pac_egy(2,:) = 31.1, 32.4, 0.0`
`pac_egy(3,:) = 62.0, 62.7, 62.3`
	2d matrix representing average photoionization energies (eV) between each package (first index) and species (second index). These can either be set manually or left to RamsesRT to derive from SED models.

`spec2pac = 1,2,3`	Determines, for each recombining species (HII, HeII, HeIII) which photon package the recombination photons are injected into (see §3.2.2 and §3.5.2).

Table D.6: *RT parameters in other namelists.*

Variable name and default value	Description
RUN_PARAMS	
`rt=.false.`	Set to activate radiative transfer
`static=.false.`	Set to run RT post-processing.
PHYSICS_PARAMS	
`neq_chem=.false.`	Set to activate non-equilibrium thermochemistry (automatically activated if `rt=.true.`).
`haardt_madau=.false.`	Set to activate homogeneous UV background.

Bibliography

Abel T, Norman ML, and Madau P. Photon-conserving Radiative Transfer around Point Sources in Multidimensional Numerical Cosmology. *ApJ* 523:66, Sep 1999.

Abel T and Wandelt BD. Adaptive ray tracing for radiative transfer around point sources. *MNRAS* 330:L53, Mar 2002.

Altay G, Croft RAC, and Pelupessy I. SPHRAY: a smoothed particle hydrodynamics ray tracer for radiative transfer. *MNRAS* 386:1931, Jun 2008.

Altay G, Theuns T, Schaye J, Crighton NHM, and Vecchia CD. Through Thick and Thin—H I Absorption in Cosmological Simulations. *AJ* 737:L37, Aug 2011.

Alvarez MA, Bromm V, and Shapiro PR. The H II Region of the First Star. *ApJ* 639:621, Mar 2006.

Anninos P, Zhang Y, Abel T, and Norman ML. Cosmological hydrodynamics with multi-species chemistry and nonequilibrium ionization and cooling. *New Astronomy* 2:209, Aug 1997.

Aubert D and Teyssier R. A radiative transfer scheme for cosmological reionization based on a local Eddington tensor. *MNRAS* 387:295, Jun 2008.

Aubert D and Teyssier R. Reionization Simulations Powered by Graphics Processing Units. I. On the Structure of the Ultraviolet Radiation Field. *ApJ* 724:244, Nov 2010.

Baek S, Matteo PD, Semelin B, Combes F, and Revaz Y. The simulated 21 cm signal during the epoch of reionization: full modeling of the Ly-α pumping. *A&A* 495:389, Feb 2009.

Baek S, Semelin B, Matteo PD, Revaz Y, and Combes F. Reionization by UV or X-ray sources. *A&A* 523:4, Nov 2010.

Belikov AV and Hooper D. How dark matter reionized the Universe. *Physical Review D* 80:35007, Aug 2009.

Black JH. The physical state of primordial intergalactic clouds. *MNRAS* 197:553, Nov 1981. A&AA ID. AAA030.161.007.

Bruzual G and Charlot S. Stellar population synthesis at the resolution of 2003. *MNRAS* 344:1000, Oct 2003.

Bunker AJ, Wilkins S, Ellis RS, Stark DP, Lorenzoni S, Chiu K, Lacy M, Jarvis MJ, and Hickey S. The contribution of high-redshift galaxies to cosmic reionization: new results from deep WFC3 imaging of the Hubble Ultra Deep Field. *MNRAS* 409:855, Dec 2010. (c) Journal compilation © 2010 RAS.

Cantalupo S. Stars quenching stars: how photoionization by local sources regulates gas cooling and galaxy formation. *MNRAS* 403:L16, Mar 2010.

Cantalupo S and Porciani C. RADAMESH: Cosmological Radiative Transfer for Adaptive Mesh Refinement Simulations. *eprint arXiv* 1009:1625, Sep 2010.

Cantalupo S, Porciani C, Lilly SJ, and Miniati F. Fluorescent Lyα Emission from the High-Redshift Intergalactic Medium. *ApJ* 628:61, Jul 2005.

Cen R. A hydrodynamic approach to cosmology - Methodology. *ApJS* 78:341, Feb 1992.

Cen R. A Fast, Accurate, and Robust Algorithm for Transferring Radiation in Three-dimensional Space. *AJS* 141:211, Jul 2002.

Cen R and Fang T. Where Are the Baryons? III. Nonequilibrium Effects and Observables. *ApJ* 650:573, Oct 2006.

Ceverino D and Klypin A. The Role of Stellar Feedback in the Formation of Galaxies. *ApJ* 695:292, Apr 2009.

Ciardi B, Ferrara A, Marri S, and Raimondo G. Cosmological reionization around the first stars: Monte Carlo radiative transfer. *MNRAS* 324:381, Jun 2001.

Ciardi B, Stoehr F, and White SDM. Simulating intergalactic medium reionization. *Monthly Notice of the Royal Astronomical Society* 343:1101, Aug 2003.

Commerçon B, Teyssier R, Audit E, Hennebelle P, and Chabrier G. Radiation hydrodynamics with adaptive mesh refinement and application to prestellar core collapse. I. Methods. *Astronomy & Astrophysics* 529:35, May 2011.

Croft RAC and Altay G. Radiation-induced large-scale structure during the reionization epoch: the autocorrelation function. *MNRAS* 388:1501, Aug 2008.

Dijkstra M and Loeb A. Acceleration of galactic supershells by Lyα radiation. *MNRAS* 396:377, Jun 2009.

Dopita MA, Krauss LM, Sutherland RS, Kobayashi C, and Lineweaver CH. Re-ionizing the universe without stars. *Ap&SS* p. 358, Jul 2011. (c) 2011: The Author(s).

Dubroca B and Feugeas J. Etude théorique et numérique d'une hiérarchie de modèles aux moments pour le transfert radiatif. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics* 329:915, Nov 1999.

Fall SM, Krumholz MR, and Matzner CD. Stellar Feedback in Molecular Clouds and its Influence on the Mass Function of Young Star Clusters. *AJ* 710:L142, Feb 2010.

Fan X, Strauss MA, Becker RH, White RL, Gunn JE, Knapp GR, Richards GT, Schneider D, Brinkmann J, and Fukugita M. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z 6 Quasars. II. A Sample of 19 Quasars. *AJ* 132:117, Jul 2006.

Faucher-Giguère CA, Kereš D, Dijkstra M, Hernquist L, and Zaldarriaga M. Lyα Cooling Emission from Galaxy Formation. *ApJ* 725:633, Dec 2010.

Faucher-Giguère CA, Lidz A, Zaldarriaga M, and Hernquist L. A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization. *ApJ* 703:1416, Oct 2009.

Ferland GJ, Korista KT, Verner DA, Ferguson JW, Kingdon JB, and Verner EM. CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. *The Publications of the Astronomical Society of the Pacific* 110:761, Jul 1998.

Finlator K, Özel F, and Davé R. A new moment method for continuum radiative transfer in cosmological re-ionization. *MNRAS* 393:1090, Mar 2009.

Fryxell B, Olson K, Ricker P, Timmes FX, Zingale M, Lamb DQ, MacNeice P, Rosner R, Truran JW, and Tufo H. FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. *AJS* 131:273, Nov 2000.

Fumagalli M, Prochaska JX, Kasen D, Dekel A, Ceverino D, and Primack JR. Absorption-line systems in simulated galaxies fed by cold streams. *MNRAS* 418:1796, Dec 2011.

Furlanetto SR and Briggs FH. 21 cm tomography of the high-redshift universe with the Square Kilometer Array. *New Astronomy Reviews* 48:1039, Dec 2004.

Gardner JP, Mather JC, Clampin M, Doyon R, Greenhouse MA, Hammel HB, Hutchings JB, Jakobsen P, Lilly SJ, Long KS, Lunine JI, McCaughrean MJ, Mountain M, Nella J, Rieke GH, Rieke MJ, Rix HW, Smith EP, Sonneborn G, Stiavelli M, Stockman HS, Windhorst RA, and Wright GS. The James Webb Space Telescope. *Space Sci. Rev.* 123:485, Apr 2006.

Garrett MA, Cordes JM, Deboer DR, Jonas JL, Rawlings S, and Schilizzi RT. Square Kilometre Array: a concept design for Phase 1. *eprint arXiv* 1008:2871, Aug 2010.

Geen S, Slyz A, and Devriendt J. Satellite Survival in Highly Resolved Milky Way Class Halos. *eprint arXiv* 1204:3327, Apr 2012.

Gibilisco M. Reionization of the Universe Induced by Primordial Black Holes. *International Journal of Modern Physics A* 11:5541, Jan 1996.

Gnedin NY. Cosmological Reionization by Stellar Sources. *ApJ* 535:530, Jun 2000.

Gnedin NY and Abel T. Multi-dimensional cosmological radiative transfer with a Variable Eddington Tensor formalism. *New Astronomy* 6:437, Oct 2001.

Gnedin NY, Kravtsov AV, and Chen HW. Escape of Ionizing Radiation from High-Redshift Galaxies. *ApJ* 672:765, Jan 2008.

Gnedin NY and Ostriker JP. Reionization of the Universe and the Early Production of Metals. *Astrophysical Journal v.486* 486:581, Sep 1997.

Gnedin NY, Tassis K, and Kravtsov AV. Modeling Molecular Hydrogen and Star Formation in Cosmological Simulations. *ApJ* 697:55, May 2009.

Goerdt T, Dekel A, Sternberg A, Ceverino D, Teyssier R, and Primack JR. Gravity-driven Lyα blobs from cold streams into galaxies. *MNRAS* 407:613, Sep 2010.

González M, Audit E, and Huynh P. HERACLES: a three-dimensional radiation hydrodynamics code. *A&A* 464:429, Mar 2007.

Haardt F and Madau P. Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background. *Astrophysical Journal v.461* 461:20, Apr 1996.

Haiman Z, Thoul AA, and Loeb A. Cosmological Formation of Low-Mass Objects. *Astrophysical Journal v.464* 464:523, Jun 1996.

Harten A, Lax PD, and van Leer B. On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws. *SIAM review* 25:35, 1983.

Hopkins PF, Keres D, Murray N, Quataert E, and Hernquist L. Realistic Stellar Feedback & Bulge Formation in Clumpy Disks. *eprint arXiv* 1111:6591, Nov 2011.

Hopkins PF, Quataert E, and Murray N. Self-regulated star formation in galaxies via momentum input from massive stars. *MNRAS* 417:950, Oct 2011.

Hubble E. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Sciences of the United States of America* 15:168, Mar 1929.

Hui L and Gnedin NY. Equation of state of the photoionized intergalactic medium. *MNRAS* 292:27, Nov 1997.

Iliev IT, Ciardi B, Alvarez MA, Maselli A, Ferrara A, Gnedin NY, Mellema G, Nakamoto T, Norman ML, Razoumov AO, Rijkhorst EJ, Ritzerveld J, Shapiro PR, Susa H, Umemura M, and Whalen DJ. Cosmological radiative transfer codes comparison project - I. The static density field tests. *MNRAS* 371:1057, Sep 2006.

Iliev IT, Mellema G, Pen UL, Merz H, Shapiro PR, and Alvarez MA. Simulating cosmic reionization at large scales - I. The geometry of reionization. *MNRAS* 369:1625, Jul 2006.

Iliev IT, Whalen D, Mellema G, Ahn K, Baek S, Gnedin NY, Kravtsov AV, Norman M, Raicevic M, Reynolds DR, Sato D, Shapiro PR, Semelin B, Smidt J, Susa H, Theuns T, and Umemura M. Cosmological radiative transfer comparison project - II. The radiation-hydrodynamic tests. *MNRAS* 400:1283, Dec 2009.

Klein RI. Star formation with 3-D adaptive mesh refinement: the collapse and fragmentation of molecular clouds. *Journal of Computational and Applied Mathematics* 109:123–152, 1999.

Komatsu E, Smith KM, Dunkley J, Bennett CL, Gold B, Hinshaw G, Jarosik N, Larson D, Nolta MR, Page L, Spergel DN, Halpern M, Hill RS, Kogut A, Limon M, Meyer SS, Odegard N, Tucker GS, Weiland JL, Wollack E, and Wright EL. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *The Astrophysical Journal Supplement* 192:18, Feb 2011.

Kravtsov AV, Klypin AA, and Khokhlov AM. Adaptive Refinement Tree: A New High-Resolution N-Body Code for Cosmological Simulations. *Astrophysical Journal Supplement v.111* 111:73, Jul 1997.

Krumholz MR, Klein RI, and McKee CF. Radiation-Hydrodynamic Simulations of the Formation of Orion-Like Star Clusters II. The Initial Mass Function from Winds, Turbulence, and Radiation. *eprint arXiv* 1203:2620, Mar 2012.

Krumholz MR, Klein RI, McKee CF, and Bolstad J. Equations and Algorithms for Mixed-frame Flux-limited Diffusion Radiation Hydrodynamics. *ApJ* 667:626, Sep 2007.

Krumholz MR and Matzner CD. The Dynamics of Radiation-pressure-dominated H II Regions. *ApJ* 703:1352, Oct 2009.

Krumholz MR and Thompson TA. Direct Numerical Simulation of Radiation Pressure-Driven Turbulence and Winds in Star Clusters and Galactic Disks. *eprint arXiv* 1203:2926, Mar 2012.

Kuhlen M and Faucher-Giguere CA. Concordance models of reionization: implications for faint galaxies and escape fraction evolution. *eprint arXiv* 1201:757, Jan 2012.

Larson D, Dunkley J, Hinshaw G, Komatsu E, Nolta MR, Bennett CL, Gold B, Halpern M, Hill RS, Jarosik N, Kogut A, Limon M, Meyer SS, Odegard N, Page L, Smith KM, Spergel DN, Tucker GS, Weiland JL, Wollack E, and Wright EL. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters. *The Astrophysical Journal Supplement* 192:16, Feb 2011.

Laursen P and Sommer-Larsen J. Lyα Resonant Scattering in Young Galaxies: Predictions from Cosmological Simulations. *ApJ* 657:L69, Mar 2007.

Leitherer C, Schaerer D, Goldader JD, Delgado RMG, Robert C, Kune DF, de Mello DF, Devost D, and Heckman TM. Starburst99: Synthesis Models for Galaxies with Active Star Formation. *AJS* 123:3, Jul 1999.

Lemaître G. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Societe Scientifique de Bruxelles* 47:49, Jan 1927.

Levermore CD. Relating Eddington factors to flux limiters. *Journal of Quantitative Spectroscopy and Radiative Transfer* 31:149, Feb 1984.

Lopez LA, Krumholz MR, Bolatto AD, Prochaska JX, and Ramirez-Ruiz E. What Drives the Expansion of Giant H II Regions?: A Study of Stellar Feedback in 30 Doradus. *ApJ* 731:91, Apr 2011.

Maiolino R, Gallerani S, Neri R, Cicone C, Ferrara A, Genzel R, Lutz D, Sturm E, Tacconi LJ, Walter F, Feruglio C, Fiore F, and Piconcelli E. Evidence of strong quasar feedback in the early Universe. *eprint arXiv* 1204:2904, Apr 2012. 5 pages, 3 figures, submitted to MNRAS Letters.

Martel H and Shapiro PR. A convenient set of comoving cosmological variables and their application. *MNRAS* 297:467, Jun 1998.

Maselli A, Ferrara A, and Ciardi B. CRASH: a radiative transfer scheme. *MNRAS* 345:379, Oct 2003.

Mellema G, Iliev IT, Alvarez MA, and Shapiro PR. C2-ray: A new method for photon-conserving transport of ionizing radiation. *New Astronomy* 11:374, Mar 2006.

Mihalas D and Mihalas BW. Foundations of radiation hydrodynamics. *New York* , Jan 1984.

Miralda-Escudé J, Haehnelt M, and Rees MJ. Reionization of the Inhomogeneous Universe. *ApJ* 530:1, Feb 2000.

Murray N, Ménard B, and Thompson TA. Radiation Pressure from Massive Star Clusters as a Launching Mechanism for Super-galactic Winds. *ApJ* 735:66, Jul 2011.

Nakamoto T, Umemura M, and Susa H. The effects of radiative transfer on the reionization of an inhomogeneous universe. *MNRAS* 321:593, Mar 2001.

Ocvirk P and Aubert D. A signature of the internal reionisation of the Milky Way? *eprint arXiv* 1108:1193, Aug 2011.

O'Shea BW, Bryan G, Bordner J, Norman ML, Abel T, Harkness R, and Kritsuk A. Introducing Enzo, an AMR Cosmology Application. *ArXiv Astrophysics e-prints* , 2004.

Osterbrock DE and Ferland GJ. Astrophysics of gaseous nebulae and active galactic nuclei. 2006.

Pawlik AH and Schaye J. TRAPHIC - radiative transfer for smoothed particle hydrodynamics simulations. *MNRAS* 389:651, Sep 2008.

Pawlik AH and Schaye J. Photoheating and supernova feedback amplify each other's effect on the cosmic star formation rate. *MNRAS* 396:L46, Jun 2009.

Pawlik AH and Schaye J. Multi-frequency, thermally coupled radiative transfer with TRAPHIC: Method and tests. *eprint arXiv* 1008:1071, Aug 2010.

Penzias AA and Wilson RW. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ* 142:419, Jul 1965.

Perlmutter S, Aldering G, Goldhaber G, Knop RA, Nugent P, Castro PG, Deustua S, Fabbro S, Goobar A, Groom DE, Hook IM, Kim AG, Kim MY, Lee JC, Nunes NJ, Pain R, Pennypacker CR, Quimby R, Lidman C, Ellis RS, Irwin M, McMahon RG, Ruiz-Lapuente P, Walton N, Schaefer B, Boyle BJ, Filippenko AV, Matheson T, Fruchter AS, Panagia N, Newberg HJM, Couch WJ, and Project SC. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *ApJ* 517:565, Jun 1999.

Petkova M and Springel V. An implementation of radiative transfer in the cosmological simulation code GADGET. *MNRAS* 396:1383, Jul 2009.

Petkova M and Springel V. A novel approach for accurate radiative transfer in cosmological hydrodynamic simulations. *eprint arXiv* 1012:1017, Dec 2010. 20 pages, 23 figures, submitted to MNRAS.

Petkova M and Springel V. Simulations of galaxy formation with radiative transfer: Hydrogen reionisation and radiative feedback. *eprint arXiv* 1008:4459, Aug 2010. 13 pages, 15 figures, submitted to MNRAS.

Pierleoni M, Maselli A, and Ciardi B. CRASHα : coupling continuum and line radiative transfer. *MNRAS* 393:872, Mar 2009.

Powell LC, Slyz A, and Devriendt J. The impact of supernova-driven winds on stream-fed protogalaxies. *MNRAS* 414:3671, Jul 2011.

Press WH, Teukolsky SA, Vetterling WT, and Flannery BP. Numerical recipes in FORTRAN. The art of scientific computing. 1992.

Rasera Y and Teyssier R. The history of the baryon budget. Cosmic logistics in a hierarchical universe. *A&A* 445:1, Jan 2006.

Raskutti S, Bolton JS, Wyithe JSB, and Becker GD. Thermal constraints on the reionization of hydrogen by Population II stellar sources. *MNRAS* 421:1969, Apr 2012.

Razoumov AO and Cardall CY. Fully threaded transport engine: new method for multi-scale radiative transfer. *MNRAS* 362:1413, Oct 2005.

Razoumov AO and Sommer-Larsen J. Ionizing Radiation from z = 4-10 Galaxies. *ApJ* 710:1239, Feb 2010.

Reynolds DR, Hayes JC, Paschos P, and Norman ML. Self-consistent solution of cosmological radiation-hydrodynamics and chemical ionization. *Journal of Computational Physics* 228:6833, Oct 2009.

Richards GT, Strauss MA, Fan X, Hall PB, Jester S, Schneider DP, Berk DEV, Stoughton C, Anderson SF, Brunner RJ, Gray J, Gunn JE, Ivezić Ž, Kirkland MK, Knapp GR, Loveday J, Meiksin A, Pope A, Szalay AS, Thakar AR, Yanny B, York DG, Barentine JC, Brewington HJ, Brinkmann J, Fukugita M, Harvanek M, Kent SM, Kleinman SJ, Krzesiński J, Long DC, Lupton RH, Nash T, Neilsen EH, Nitta A, Schlegel DJ, and Snedden SA. The Sloan Digital Sky Survey Quasar Survey: Quasar Luminosity Function from Data Release 3. *AJ* 131:2766, Jun 2006.

Riess AG, Filippenko AV, Challis P, Clocchiatti A, Diercks A, Garnavich PM, Gilliland RL, Hogan CJ, Jha S, Kirshner RP, Leibundgut B, Phillips MM, Reiss D, Schmidt BP, Schommer RA, Smith RC, Spyromilio J, Stubbs C, Suntzeff NB, and Tonry J. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ* 116:1009, Sep 1998.

Rijkhorst EJ, Plewa T, Dubey A, and Mellema G. Hybrid characteristics: 3D radiative transfer for parallel adaptive mesh refinement hydrodynamics. *A&A* 452:907, Jun 2006.

Rosdahl J and Blaizot J. Extended Lyα emission from cold accretion streams. *MNRAS* p. 2837, Apr 2012.

Sokasian A, Yoshida N, Abel T, Hernquist L, and Springel V. Cosmic reionization by stellar sources: population III stars. *MNRAS* 350:47, May 2004.

Spitzer L. Physical processes in the interstellar medium. *New York Wiley-Interscience* , Jan 1978.

Springel V. The cosmological simulation code GADGET-2. *MNRAS* 364:1105, Dec 2005.

Springel V, Frenk CS, and White SDM. The large-scale structure of the Universe. *Nature* 440:1137, Apr 2006.

Steinacker J, Bacmann A, and Henning T. Application of adaptive multi-frequency grids to three-dimensional astrophysical radiative transfer. *Journal of Quantitative Spectroscopy & Radiative Transfer* 75:765, Dec 2002.

Strömgren B. The Physical State of Interstellar Hydrogen. *ApJ* 89:526, May 1939.

Sun P. A Pseudo-Non-Time-Splitting Method in Air Quality Modeling. *Journal of Computational Physics* 127:152, Aug 1996.

Susa H. Smoothed Particle Hydrodynamics Coupled with Radiation Transfer. *Publications of the Astronomical Society of Japan* 58:445, Apr 2006.

Teyssier R. Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. *A&A* 385:337, Apr 2002.

Theuns T, Schaye J, Zaroubi S, Kim TS, Tzanavaris P, and Carswell B. Constraints on Reionization from the Thermal History of the Intergalactic Medium. *ApJ* 567:L103, Mar 2002.

Toro EF. Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction. 1999.

Trac H and Pen UL. A moving frame algorithm for high Mach number hydrodynamics. *New Astronomy* 9:443, Jul 2004.

Vaytet N, Audit E, and Dubroca B. A Numerical Model for Multigroup Radiation Hydrodynamics. *Numerical Modeling of Space Plasma Flows* 429:160, Sep 2010.

Verhamme A, Dubois Y, Slyz A, and Devriendt J. Lyα radiation transfer in simulated galaxies : variation of Lyα escape fraction, spectra, and images with viewing angle, for different galaxy formation scenarios. *SF2A-2009: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics* p. 225, Nov 2009.

Verhamme A, Schaerer D, Atek H, and Tapken C. 3D Lyα radiation transfer. III. Constraints on gas and stellar properties of z 3 Lyman break galaxies (LBG) and implications for high-z LBGs and Lyαemitters. *A&A* 491:89, Nov 2008.

Verhamme A, Schaerer D, and Maselli A. 3D Lyα radiation transfer. I. Understanding Lyα line profile morphologies. *A&A* 460:397, Dec 2006.

Verner DA, Ferland GJ, Korista KT, and Yakovlev DG. Atomic Data for Astrophysics. II. New Analytic FITS for Photoionization Cross Sections of Atoms and Ions. *Astrophysical Journal v.465* 465:487, Jul 1996.

Wadsley JW, Stadel J, and Quinn T. Gasoline: a flexible, parallel implementation of TreeSPH. *New Astronomy* 9:137, Feb 2004.

Whalen D and Norman ML. A Multistep Algorithm for the Radiation Hydrodynamical Transport of Cosmological Ionization Fronts and Ionized Flows. *AJS* 162:281, Feb 2006.

Wise JH and Abel T. Resolving the Formation of Protogalaxies. III. Feedback from the First Stars. *ApJ* 685:40, Sep 2008.

Wise JH and Abel T. ENZO+MORAY: radiation hydrodynamics adaptive mesh refinement simulations with adaptive ray tracing. *MNRAS* 414:3458, Jul 2011.

Wise JH and Cen R. Ionizing Photon Escape Fractions From High-Redshift Dwarf Galaxies. *ApJ* 693:984, Mar 2009.

Yajima H, Choi JH, and Nagamine K. Escape fraction of ionizing photons from high-redshift galaxies in cosmological SPH simulations. *MNRAS* 412:411, Mar 2011.

Yajima H, Li Y, Zhu Q, and Abel T. ART2 : Coupling Lα Line and Multi-wavelength Continuum Radiative Transfer. *eprint arXiv* 1109:4891, Sep 2011. 19 pages, 16 figures, submitted to MNRAS.

Zahn O, Lidz A, McQuinn M, Dutta S, Hernquist L, Zaldarriaga M, and Furlanetto SR. Simulations and Analytic Calculations of Bubble Growth during Hydrogen Reionization. *ApJ* 654:12, Jan 2007.