Who reionized the Universe ? : dermining the photon budget of galaxies during reionization with numerical simulations, and studying the impact of dust

Joseph Lewis

To cite this version:

Joseph Lewis. Who reionized the Universe ? : dermining the photon budget of galaxies during reionization with numerical simulations, and studying the impact of dust. Astrophysics [astro-ph]. Université de Strasbourg, 2020. English. NNT : 2020STRAE041 . tel-03199136

HAL Id: tel-03199136
https://theses.hal.science/tel-03199136v1

Submitted on 15 Apr 2021

UNIVERSITÉ DE STRASBOURG

École doctorale
Physique, chimie-physique
ED 182
Université de Strasbourg

ÉCOLE DOCTORALE 182
UMR 7550, Observatoire astronomique de Strasbourg

THÈSE présentée par :

Joseph Lewis

soutenue le : **25 septembre 2020**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**
Discipline/ Spécialité : Astrophysique

Qui a réionisé l'Univers ?

Détermination par la simulation numérique du budget de photons des galaxies pendant l'époque de la Réionisation, et étude de l'impact des poussières

THÈSE dirigée par :

M. AUBERT Dominique	Professeur des universités, Université de Strasbourg

RAPPORTEURS :

M. GONZALES Mathias	Maître de conférences, Université de Paris
M. LANGER Mathieu	Professeur des universités, Université Paris-Saclay

---

AUTRES MEMBRES DU JURY :

M. SEMELIN Benoît	Professeur des universités, Sorbonne Université
Mme. LANCON Ariane	Professeur des universités, Université de Strasbourg
M. OCVIRK Pierre	Astronome adjoint, Université de Strasbourg

ACKNOWLEDGEMENT

First and foremost, I'd like to express my deepest gratitude and respect to Pierre and Dominique. Though they will definitely be humble about the quality of their supervision, I am thoroughly impressed by their availability and readiness to lend a hand, explain something or just have a chat. From the start, I found them very easy to work with, their laid back and informal but involved and supportive style is un-intimidating, pleasant, and productive. This, combined with a hands on approach provided me with a fantastic learning opportunity. I look forward to continue working and interacting with them whenever possible.

I also want to thank Timothé and Federico, who both endured my loud keystrokes, sporadic ranting, and extensive kettle usage with a smile. I couldn't have hoped for better office mates. Hopefully, the next generation can continue to annoy each other with the office nerf gun.

On the whole, I feel hugely indebted to everyone at the *Observatoire* for their warm and friendly they help maintain, it made me feel welcome, and at home in Strasbourg. In fact I hadn't anticipated making so many friends and memories here, both of which I'll be clinging to dearly. Thus, for (heavy)lifting my spirits I'd like to mention JB, Jérôme, Jonathan, Julie, Julien, Lucie, Mathieu, the Nicolas (D., G., and L.), Paolo, and Timothé; but also Ada, Katarina, Lorenzo, Mari, Oliver, and Yelena.

I'd also like to thank Aline, Elena, Ludmilla, Melchior, and Orlane for putting up with me at home (and the amount of dirty dishes I produce). In particular, thank-you Orlane and Elena for making writing in quarantine bearable.

A special mention goes out to François and Raphael for the hours of fun over discord, as well as to Benoît and to Cyril who also continue to enjoy my company, which is no mean feat.

Naturally, I want to say to my parents how much I appreciate their enduring support and love, enabling me on all possible fronts.

Finally, I'd like to renew my expression of gratitude to Roser, for turning me on to astrophysics all the way back in 2014.

TABLE OF CONTENTS

FOREWORD

Hierarchical structure formation has enjoyed much success in explaining the gradual bottom up formation of the large scale structure of the Universe, and galaxies. However, because it is extremely difficult to observationally constrain the formation of the first structures, a large gap in our knowledge persists. One possible avenue of investigation is to study the Epoch of Reionization(EoR): When the first ionising sources formed in the first galaxies a few hundred million years after the Big Bang, they ionised the intergalactic Hydrogen gas over the next several hundred million years. During this process, and as time went by, and the number of ionising sources grew, so did the ionised regions surrounding galaxies, until eventually the intergalactic matter (IGM) was fully ionised. Evidently, the EoR is directly connected to the build up of structure. Therefore, following the density and luminosity of ionising sources, as well as the growth and distribution of ionised regions in the IGM at very high redshift gives a window into the formation of the first structure, its spatial distribution, and its evolution.

In fact, the study of the EoR, the main topic of interest in this thesis, is timely. Indeed, over the next decades a new generation of instruments is set to revolutionise our observations of Reionization. First, new facilities capable of targetting reionising galaxies such as the James Webb Space Telescope (JWST)[1], or the Extremely Large Telescope (ELT)[2] will see first light. Second, new radio experiments such as the Square Kilometre Array (SKA)[3] or the Low Frequency Array (LOFAR)[4] will allow us to directly probe the density and distribution of neutral Hydrogen gas. Interpreting these new datasets will require extensive modelling and simulating, to retrieve physical constraints on Reionization.

Understanding how Reionization unfolds requires understanding the relative contribution of the various ionising sources to Reionization, and how it evolved. Therefore, one question at the heart of my work is "Which ionising sources drove Reionization?". This has proved a difficult question to answer using observations, as it requires extensive assumptions and modelling. This is all the more so the case if one is interested in a more detailed answer, such as the nature of the galaxies that contributed the most. Numerical simulations are therefore a useful tool to explore answers and their implications, and that I relied on in this thesis. For instance, in simulations it is possible to fully explore the contributions of individual galaxies or populations of galaxies, according to their properties: the ionising photon budget of galaxies. However, due to computational limitations, numerical studies have tended to lack either in resolution, scale, or the fully coupled physics that would enable a consistent investigation of how galaxies propel large scale Reionization of the IGM. As of such, renewed efforts must be made to realistically depict Reionization in order to be able to investigate its galactic drivers. One often neglected aspect of the problem is the impact of dust on Reionization. Although its effects

---

[1] Currently the JWST's launch is planned for 2021.

[2] First light with the ELT is planned for 2025.

[3] SKA's first light is planned for 2020.

[4] LOFAR has already provided upper limits on the global average 21cm signal of neutral Hydrogen during the EoR [Mertens et al., 2020].

have sometimes been explored, it has often been in the form of simple post-processing models, that cannot account for the coupling of dust with gas and radiation. It is the investigation of the above shortcomings, that led me to explore the following questions, central to this manuscript :

- What is the ionising photon budget of galaxies?
    - Which galaxies dominate the photon budget? Why?
    - Does it vary during Reionization? Why?
- What missing physics could increase the realism of the EoR in simulations? How including them in our simulations affect the photon budget?

To find an answer, I first analysed an existing large scale fully coupled radiation and hydrodynamics simulation of the EoR : Cosmic Dawn II (CoDa II), and determined the most suitable way to measure the contribution of galaxies to Reionization. By investigating the resulting photon budget of galaxies, it emerged that an intermediate mass range ($6 \times 10^{8}\,\mathrm{M}_{\odot}$ to $2 \times 10^{10}\,\mathrm{M}_{\odot}$) of galaxies drove Reionization in the simulation. To take this answer further, I sought to improve the realism of future simulations by working on the RAMSES-CUDATON code. Accordingly, I strived to implement new more realistic recipes for the ionising emissivities of stellar populations, additional cooling rates from metals and Helium, as well as dust, that could play an important role in Reionization. I then evaluated the impact of these additions to the code, and calibrated the new dust model by running a series of new calibration simulations. At the same time, I explored new parametrisations of the existing simulation physics, and a different setup of the sub-grid recipe for star formation, so as to better match observational constraints on Reionization. Finally, I performed a new production run sized simulation with the new setup. In this new simulation, I found that the new setup and new physics favoured a late Reionization scenario, in which the dominant sources of Reionization are lower mass galaxies than in CoDa II : $10^{8}\,\mathrm{M}_{\odot}$ to $10^{10}\,\mathrm{M}_{\odot}$. Moreover, I was able to directly investigate the role of dust in Reionization, finding that the ionising photon budget is measurably affected at the massive bright end, and that the ultra violet luminosity function (UVLF) is significantly impacted for $\mathrm{M}_{\mathrm{AB1600}} \lesssim$ -19 after $z \approx 8.5$.

Organisation of the manuscript:

- The Introduction, where I properly situate the EoR in cosmic history, and review the existing studies of the contribution of galaxies to Reionization. Finally, I list some constraints on the EoR that can be used to calibrate simulations, and briefly present the RAMSES-CUDATON simulation code.
- Chp. 1 presents the CoDa II simulation (including the article that presented it: Ocvirk et al. [2020]), and my analysis of the ionising photon budget of galaxies in CoDa II (which was published in my first paper as first author, also included: Lewis et al. [2020]). At the end of the chapter, I show results that were not within the scope of my publication, but that enable me to expand my analysis.

- Chp. 2 explains the new physics that I added to the RAMSES-CUDATON simulation code. In particular length, it details the implementation of dust, and compares my calibration of the implemented dust model with work from observations, semi-analytical models, and simulations. At the same time, I explore the possible impact of each new feature on the EoR.
- Chp. 3 outlines how I set out to change the simulation setup (with respect to that of CoDa II) in order to better match constraints on the ionisation of the IGM (from Lyα forest data[5]). Afterwards, I compare my new simulation to constraints on the ionisation rate of Hydrogen, the UVLF, the star formation rate density (SFRD), the stellar mass to halo mass relation, and the integrated Thomson optical depth from the scattering of CMB photons.
- Chp. 4 investigates the combined effect of the new simulation setup (explained in Chp. 3) and new physics in the code (presented in Chp. 2) on the ionising photon budget of galaxies (and the differences with the photon budget from CoDa II as presented in Chp. 1). The results from this chapter are planned to be published (along with elements from the previous chapters presenting the new code and setup).
- Conclusions and perspectives reviews the work I carried out as a whole, gives a summary of what remains to be done, and reconnects the above with the context of the field.

Publication list (during PhD):

Papers presented in the manuscript:

- Lewis Joseph S. W., Pierre Ocvirk, Dominique Aubert, Jenny G. Sorce, Paul R. Shapiro, Nicolas Deparis, Taha Dawoodbhoy, et al. « Galactic ionising photon budget during the Epoch of Reionisation in the Cosmic Dawn II simulation. » Monthly Notices of the Royal Astronomical Society, 1 juin 2020. https://doi.org/10.1093/mnras/staa1748.
- Ocvirk Pierre, Dominique Aubert, Jenny G. Sorce, Paul R. Shapiro, Nicolas Deparis, Taha Dawoodbhoy, Joseph Lewis, et al. « Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization ». Monthly Notices of the Royal Astronomical Society, 1 mai 2020. https://doi.org/10.1093/mnras/staa1266.

Other papers (Co-authored papers in which I played a limited role):

- Gronke Max, Pierre Ocvirk, Charlotte Mason, Jorryt Matthee, Sarah E. I. Bosman, Jenny G. Sorce, Joseph Lewis, et al. « Lyman-alpha transmission properties of the intergalactic medium in the CoDaII simulation ». arXiv e-prints 2004 (1 avril 2020): arXiv:2004.14496.
- Chardin Jonathan, Grégoire Uhlrich, Dominique Aubert, Nicolas Deparis, Nicolas Gillet, Pierre Ocvirk, et Joseph Lewis. « A Deep Learning Model to Emulate Simulations of Cosmic Reionization ». Monthly Notices of the Royal Astronomical Society 490, °1 (novembre 2019): 1055. https://doi.org/10.1093/mnras/stz2605.

---

[5] See Sec. Constraints on Reionization for details on Lyα constraints.

- Ocvirk P., D. Aubert, J. Chardin, N. Deparis, et J. Lewis. « Impact of the Reduced Speed of Light Approximation on the Post-Overlap Neutral Hydrogen Fraction in Numerical Simulations of the Epoch of Reionization ». Astronomy and Astrophysics 626 (juin 2019): A77. https://doi.org/10.1051/0004-6361/201832923.
- Deparis Nicolas, Dominique Aubert, Pierre Ocvirk, Jonathan Chardin, et Joseph Lewis. « Impact of the Reduced Speed of Light Approximation on Ionization Front Velocities in Cosmological Simulations of the Epoch of Reionization ». Astronomy and Astrophysics 622 (février 2019): A142. https://doi.org/10.1051/0004-6361/201832889.

BIBLIOGRAPHY

Lewis J.S.W., et al.. Galactic ionising photon budget during the Epoch of Reionisation in the Cosmic Dawn II simulation. *Monthly Notices of the Royal Astronomical Society*, 2020.

Mertens F.G., et al.. Improved upper limits on the 21 cm signal power spectrum of neutral hydrogen at z $\approx$ 9.1 from LOFAR. *Monthly Notices of the Royal Astronomical Society*, 493:1662–1685, 2020.

Ocvirk P., et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 2020.

AVANT-PROPOS

La description de la naissance des grandes structures de l'Univers via une histoire de formation hiéarchique a connu beaucoup de succès. Cependant, dû à la difficulté extrême des observations des premières structures et de leur formation, de grandes zones d'ombre persistent. Dès lors, une approche d'investigation possible est d'étudier l'époque de la Réionisation : lorsque, quelques centaines de millions d'années après le Big Bang, les premières sources ionisantes sont apparues dans les premières galaxies, la lumière ionisante ainsi produite ionisa le gaz d'Hydrogène intergalactique au cours des centaines de millions d'années qui ont suivi. Au fil du temps, le nombre de sources ionisantes augmenta, avec le nombre et la taille des régions de gaz ionisées, jusqu'à ce que le gaz intergalactique soit intégralement ionisé. Clairement, l'époque de la Réionisation est directement liée à la formation des structures. Il en découle que suivre la densité et la luminosité des sources ionisantes, ainsi que la densité et le volume de régions ionisées, constitue une sonde de la formation des premières structures, et de leur évolution.

L'étude de l'époque de la Réionisation, le sujet principal de cette thèse, est d'actualité. En effet, au cours des prochaines décénies, une nouvelle génération d'instruments révolutionnera les observations de la Réionisation. Premièrement, de nouvelles installations capables de cibler les galaxies en cours de Réionisation, telles que le télescope spatial James Webb[6], ou l'ELT[7] verront le jour. Deuxièmement, des nouvelles expériences radio telles que SKA[8] ou LOFAR[9] nous permettront de sonder directement la densité, et la distribution du gaz d'Hydrogène neutre. L'interprétation de ces nouveaux jeux de données requerra des efforts en modélisation et en simulation afin de produire des constraintes physiques sur la Réionisation.

Comprendre le déroulement de la Réionisation nécessite de comprendre la contribution relative des différentes sources ionisantes à la Réionisation, et comment cela a pu évoluer pendant l'époque de la Réionisation. De ce fait, une des questions centrales dans mon travail a été "Quelles sont les sources ionisantes responsables de la Réionisation ?". La réponse à cette question s'est montrée difficile à approcher avec les observations, car cela implique beaucoup d'hypothèses et de la modélisation. C'est d'autant plus le cas si l'on souhaite parvenir à une réponse plus détaillée, qui mettrait en évidence la nature des galaxies responsables de la Réionisation. Les simulations numériques sont donc un outil indispensable auquel j'ai eu recours pour explorer d'éventuelles réponses à ces questions. Par exemple, dans des simulations il est possible d'étudier les contributions à la Réionisation de galaxies individuelles ou de populations de galaxies selon leurs propriétés : quand on compare la contribution des galaxies selon leur masses, on étudie le budget galactique des photons ionisants. Néanmoins, cette approche numérique se heurte rapidement aux limitations technologiques de capacité de calcul. Il en résulte que

---

[6] dont le lancement est prévu pour 2021.

[7] ou *Extremely Large Telescope* en anglais, dont l'entrée en service est prévu en 2025.

[8] ou *Square Kilometer Array* en anglais, qui entrera en service après 2020.

[9] ou *LOw Frequency ARray* en anglais, a déjà fourni des contraintes sur le signal global moyenné des emissions à 21cm de gaz d'Hydrogène neutre pendant la Réionisation [Mertens et al., 2020].

les études numériques du budget galactique de photons ionisants ont souvent eu recours à des sacrifices au niveau de la résolution, de l'échelle ou des modèles physiques inclus, limitant les interprétations possibles. De ce fait, de nouveaux efforts doivent être réalisés pour obtenir une description numérique réaliste de la Réionisation afin d'étudier ses sources galactiques. Un aspect souvent négligé dans ces études est l'effet potentiel de la poussière sur la Réionisation. Bien que les effets de la poussière aient été parfois explorés, le plus souvent cela a été fait grâce à des modèle simples de post-traitement, qui ne peuvent pas rendre compte du couplage entre la poussière, le gaz et les rayonnements. C'est l'étude de ces manquements qui m'a mené à me poser les questions suivantes, d'importance cruciale à ce manuscrit :

- Quel est le budget galactique des photons ionisants ?
    - Quelles galaxies dominent ce budget des photons ? Pourquoi ?
    - Comment est-ce-que le budget des photons évolue pendant la Réionisation ? Pourquoi ?
- Quels modèles physiques absents pourraient augmenter le réalisme de nos simulations de la Réionisation ? Comment ces modèles affectent le budget des photons ionisants ?

Pour trouver une réponse, j'ai commencé par analyser une simulation de la Réioinsation à très grande échelle où l'hydrodynamique et le transfert du rayonnement sont couplés : Cosmic Dawn II (CoDa II). Après avoir déterminé la méthode la plus appropriée pour mesurer la contribution des galaxies à la Réionisation dans la simulation, j'ai étudié le budget de photons des galaxies. J'ai montré qu'une gamme de masse intermédiaire (6×10^8 à $3 \times 10^{10}\,\mathrm{M}_\odot$) de galaxies alimente la Réionisation en photons dans la simulation. Pour aller plus loin, j'ai cherché à améliorer le réalisme de simulations futures en travaillant sur le code de simulation RAMSES-CUDATON. Dans ce sens, j'ai oeuvré à l'implémentation de recettes plus réalistes pour l'émissivité ionisante des populations stellaires, de nouveaux processus de refroidissement du gaz dû à l'hélium et aux métaux, ainsi que de modèles de poussière qui pourrait avoir un effet important sur la Réionisation. Par la suite j'ai évalué les conséquences de ces ajouts au code, et ai étalonné le modèle de poussière retenu en lançant une série de simulations de calibration. En même temps, pour mieux reproduire les contraintes observationnelles sur la Réionisation, j'ai étudié de nouvelles paramétrisations des modèles physiques qui figuraient déjà dans le code, ainsi qu'une nouvelle configuration pour la recette sous-grille de formation stellaire. Finalement, j'ai réalisé une nouvelle simulation de taille adaptée à l'étude du budget des photons. Dans cette nouvelle simulation, j'ai découvert que les nouveaux modèles physiques et la nouvelle calibration favorisaient une Réionisation tardive, dans laquelle les sources dominantes de la Réionisation sont moins massives que dans CoDa II : entre10^8 et $10^{10}\,\mathrm{M}_\odot$. De plus, j'ai pu directement explorer le rôle de la poussière pendant la Réionisation, et ai trouvé que le budget de photons ionisants des galaxies massives est mesurablement modifié, et que la fonction de luminosité ultra-violette des galaxies est très affectée par l'extinction due aux poussières pour des magnitudes $\lesssim$ -19 après z $\approx$ 8.5 .

Organisation du manuscrit:

- Dans Introduction, je situe la Réionisation dans l'Histoire de l'Univers, et passe en revue les approches employées pour déterminer les sources de la Réionisation. Finalement, j'énumère les contraintes observationnelles principales de la Réionisation, et présente brièvement le code de simulation RAMSES-CUDATON.

- Le Chp. 1 présente la simulation CoDa II (et inclut l'article la présentant: Ocvirk et al. [2020]), ainsi que mon analyse du budget de photon ionisant des galaxies (qui a figuré dans mon premier article publié dans un journal scientifique, lui aussi inclus: Lewis et al. [2020]). A la fin du chapitre, je montre des résultats supplémentaires qui ne figuraient pas dans ma publication, et qui me permettent d'étendre mon analyse.

- Le Chp. 2 explique les nouvelles physiques que j'ai ajouté au code RAMSES-CUDATON. De façon détaillée, j'expose l'implémentation de la poussière que j'ai réalisée, ainsi que l'étalonnage des modèles de poussière et leur comparaison avec des études observationnelles, semi-analytiques, et numériques. En même temps, j'étudie les effets possibles de chaque nouvelle implémentation sur la Réionisation.

- Le Chp. 3 détaille mon approche pour modifier la configuration des simulations afin de mieux reproduire des contraintes observationnelles sur l'ionisation du milieu inter-galactique (tirées de données forêt Lyman α, c.f. Sec. Constraints on Reionization). Ensuite, j'ai comparé ma nouvelle simulation aux contraintes sur le taux d'ionisation de l'hydrogène neutre, la fonction de luminosité ultra-violette des galaxies, la densité de formation stellaire, la relation entre la masse stellaire et la masse de halo de matière noire des galaxies, ainsi que la profondeur optique Thomson intégrée produite par la diffusion des photons du CMB par des électrons libres.

- Le Chp. 4 étudie les effets combinés de la nouvelle configuration des simulations (expliquée dans le Chp. 3) et des nouveaux modèles physiques dans le code (présentés dans le Chp. 2) sur le budget de photons ionisants des galaxies. Les résultats figurant dans ce chapitre seront publiés (avec des éléments explicatifs des chapitres précédents).

- Dans Conclusions et perspectives, je passe en revue le travail accompli, résume le travail restant, et resitue mon travail dans le contexte actuel du domaine.

INTRODUCTION

A brief note on standard cosmology

The 2019 Nobel prize for physics was awarded to Jim Peebles for his extensive contributions to the foundation of what is now considered standard cosmology, monumentalising the revolutions that occurred in the field over the past ≈ 60 years[10] as one of the great recent advances in Physics. In this section, we shall briefly recall some elements of standard cosmology or ΛCDM[11] cosmology, that are fundamental for understanding the Epoch of Reionization, the main topic of this manuscript (for a recent full review see Zavala and Frenk [2019]).

ΛCDM cosmology

One of the early fundamental observations of cosmology is that galaxies appear to recede, with distant galaxies receding faster than nearer ones [as described in Lemaître, 1927; Hubble, 1929]. Rapidly, this was interpreted as the sign that we inhabit a dynamical, expanding Universe, in which physical distances evolve with time. Let us define $r(t)$, a physical distance as $r(t) = a(t)R(t)$, where $R(t)$ is a distance that follows expansion, or a comoving distance, and $a(t)$ is the scale factor at a time t. We take the scale factor to be normalised, so that $a(t{=}0){=}1$, and $a(t{=}t_{BB}){=}0$ (where $t{=}0$ is the present time, and t_{BB} is the time at the moment of the Big Bang). We also define $z(t)$ the cosmological redshift, related to $a(t)$ by $z(t) = \frac{1}{a(t)} - 1$. Then we can express the Hubble parameter, or the rate of cosmological expansion as $H(t) = \frac{\dot{a}(t)}{a(t)}$ (with $\dot{a}(t)$, the total time derivative of $a(t)$). The recession speed of a distant galaxy is then $H(t)r(t)$.

Today, standard cosmology (or ΛCDM cosmology) describes the evolution of a homogeneous and isotropic Universe with a set of fundamental measurable parameters, using the formalism of General Relativity. It allows for the following expression of $H(t)$ as a function of redshift and of the cosmological parameters that describe the content of the Universe:

$$H^2 = H_0^2(\Omega_\Lambda 0 + \Omega_{k0}(z+1)^2 + \Omega_{m0}(z+1)^3 + \Omega_{r0}(z+1)^4)\,, \tag{1}$$

with H_0 the present value of H (when z=0 and a=1), $\Omega_{\Lambda 0}$ the dark energy density parameter, Ω_{k0} the curvature parameter, Ω_{m0} the matter energy density parameter, Ω_{r0} the relativistic species'[12] energy density parameter.

Crucially, the predictions of the model have allowed for precise determinations of the cosmological parameters. For instance, by measuring the angular power spectrum of fluctuations in the cosmic mi-

---

[10] closer to ≈ 100 years if including the discovery of other galaxies, and cosmological expansion.

[11] i.e. : Cold Dark Matter with a cosmological constant Λ

[12] Here matter is used in the largest possible sense, and encompasses all energy densities that are not in the dark energy density term or the curvature density term. Relativistic species refers to particles that travel at relativistic speeds (e.g. photons, neutrinos).

crowave background (CMB), it is possible to constrain the relative densities of dark matter, baryons, dark energy, and the average curvature of space time with precision, as done by Planck (and WMAP and COBE before). In Tab. 1, I list some of the fundamental parameters as given by Planck Collaboration et al. [2018]. Other independent (depending on the measured parameter) determinations can be made, some of which we list here :

- Using the evolution shape of the universal of supernova luminosity with time as a standard ruler, one can determine the rate of cosmological expansion and it's evolution [as in Perlmutter et al., 1999; Scolnic et al., 2018].
- When the CMB photons decoupled from the plasma, the sound speed horizon of acoustic oscillations caused by density fluctuations was set into the matter power spectrum, creating a bump at the corresponding scale. Thus, the sound speed horizon can be used as a standard ruler to determine H_0, and in combination with CMB data, $\Omega_{\Lambda 0}$ [Eisenstein and Hu, 1998; Viel et al., 2002; Alam et al., 2017].
- Weak lensing is a small distortion of galaxy images by mass near the line of sight. Wide scale surveys of weak lensing can act as probes of the matter power spectrum by constraining $\sigma_8\Omega_{\mathrm{m}0}$ [Blandford et al., 1991; Hikage et al., 2019].
- Predictions for the abundance of massive haloes can be compared with the observations for massive galaxy clusters, which are sensitive to both the amplitude of the matter power spectrum σ_8 and the matter density $\Omega_{\mathrm{m}0}$ [Peebles et al., 1989; Bocquet et al., 2019].
- Conditions in the first minutes following the Big Bang were conducive to the nucleosynthesis of low mass nuclides. Because the relative mass fractions of the different species produced during Big Bang Nucleosynthesis (BBN) were sensitive to the conditions in the early Universe, and since they can be independently determined via observations, BBN can be used to test cosmologies [Walker et al., 1991; Cyburt et al., 2016].

Observations suggest that we inhabit a flat ($\Omega_{\mathrm{k}0}\approx 0$) expanding Universe with dark matter, with accelerating expansion ($\Omega_{\Lambda 0} > 0$). Of particular note is the high value of $\Omega_{\Lambda 0}$ compared to the other density parameters. $\Omega_{\Lambda 0}$ is far greater than the matter density parameter $\Omega_{\mathrm{m}0}$, who's dominant constituent is dark matter. The agreement between the variety of cosmological probes and complementary constraints, is illustrative of the success that ΛCDM has enjoyed. Appropriately, it is referred to as a concordance cosmology.

Despite the many successes of ΛCDM, there still remain a number of hurdles to jump before its description of the Universe can be considered complete [for a recent review of successes and issues, see Bull et al., 2016; Bullock and Boylan-Kolchin, 2017]. For example, the determinations of the Hubble expansion parameter H_0 from CMB data and BAO are in tension with the observations from supernova light curves [Knox and Millea, 2020]. Particularly topical are the so called problems that arise at small scales when comparing numerical simulations' predictions to observations of structure. For instance: simulations over-predict the number of satellite dwarf galaxies of Milky Way-like galaxies (missing satellites problem, see Klypin et al. [1999]); simulations predict the presence of massive sub-haloes

Planck 2018 estimates		
Ω_{b0}	0.04897 ± 0.00030	Baryon energy density parameter
$\Omega_{\mathrm{c}0}$	0.26067 ± 0.00199	Dark matter energy density parameter
$\Omega_{\Lambda 0}$	0.6889 ± 0.0056	Dark energy density parameter
$\Omega_{\mathrm{k}0}$	0.0007 ± 0.0019	Curvature density parameter
H_0	$67.66 \pm 0.042\ \mathrm{km} \cdot \mathrm{s}^{-1} \cdot \mathrm{Mpc}^{-1}$	Hubble constant
$\mathrm{n_s}$	0.9665 ± 0.0038	Primordial spectral index
σ_8	0.8102 ± 0.006	Normalisation of the matter power spectrum
$\mathrm{Y_p^{BBN}}$	$0.243 \pm^{0.023}_{0.024}$	Primordial mass fraction of Helium
τ	0.0561 ± 0.0071	Optical depth to CMB photons due to free electrons

Table 1: Some of the fundamental cosmological parameters from Planck Collaboration et al. [2018] that are useful in this manuscript.

.

that aren't observed, and would be too bright to be missed or suppressed (too big to fail problem, as in Boylan-Kolchin et al. [2011]); the density profiles of galaxies in observations have 'cores' when compared to denser simulated profiles that are 'cuspier' (core vs cusp problem, seen first in Flores and Primack [1994]). Better understanding the initial stages of structure formation by studying the EoR, is a promising approach to overcome these inconsistencies.

A short timeline of cosmic history

To properly situate the moment of cosmic history that interests us, we broadly review the history of the Universe according to standard cosmology, and inspired by Barkana et al. [2001]. Fig. 1 shows a representative history of the universe, starting at the Big Bang, and ending at some time close to the present. Particularly clear in this image, is the gradual change in the nature of structure at small scales, from early times when it is simple and homogeneous, to the present day when it is complex and heterogeneous.

In the moments following the Big Bang, the Universe was extremely hot and dense. It is thought that Inflation or a period of extremely rapid exponential expansion imprinted a random Gaussian field of small primordial density fluctuations onto the matter distribution. After Inflation, the cosmological expansion of the Universe continued according to the evolution of H(t) as in Eq. 1.

During the first minutes after the Big Bang, the high temperature of primordial matter allowed Big Bang Nucleosynthesis (or BBN) [for a recent review, see Cyburt et al., 2016]. Through this process, Helium, Lithium, and Beryllium formed. In fact, BBN resulted in a large mass fraction of Helium nuclides: $0.243 \pm^{0.023}_{0.024}$ [?].

Because the plasma was so hot, and because it was permeated by black body radiation, the growth of baryonic density fluctuations was opposed by radiation pressure, but by $z \approx 3800$ dark matter density fluctuations grew unopposed. Cosmological expansion gradually cooled the primordial plasma, reducing its density and temperature, and increasing the mean free path of CMB photons. When $z \approx 1100$ (The Universe was $\approx 380\,000$ yr old), the mean free path of the cosmic microwave background

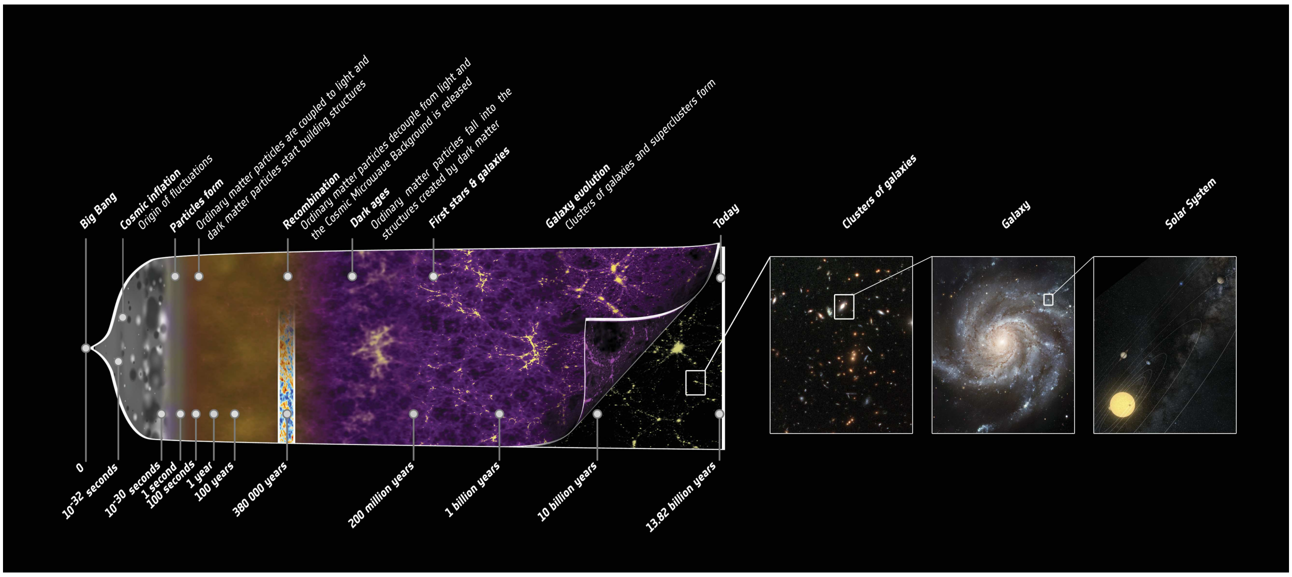


Figure 1: Illustrative timeline of the history of the Universe from the Big Bang to today. Credit:ESA – C. Carreau

became essentially infinite, allowing its free propagation and observation in the present day.

By z $\approx$ 1000, cosmological expansion had sufficiently cooled the plasma for it to recombine, ushering in what is known as the Dark Ages, so called because of the absence of stellar radiation. Around this period, the coupling between the photons and the residual free electrons in the gas became too weak to oppose the gravitational attraction of baryons to the growing dark matter density fluctuations. Resultantly, baryons can cool and start to fall onto the dark matter density perturbations, and the power spectrum of baryonic fluctuations approaches the power spectrum of the dark matter density fluctuations.

When dark matter density perturbations grew to sufficient overdensities, they collapsed into virialised dark matter haloes. Baryons could then accumulate in the haloes that were massive enough. After collapsing from perturbations, dark matter haloes grew via mergers with other haloes. In dark matter haloes that reached sufficient masses($\approx 10^6\,\mathrm{M}_\odot$), there was enough in-fall of cool gas for them to form the first stars via H_2 cooling, marking Cosmic Dawn, and ending the Dark Ages $\approx$ 300-100Myr(z$\approx$ 15-30) after the Big Bang.

Epoch of Reionization

The Epoch of Reionization[13], sometimes referred to as the last cosmological phase transition, is defined as the period separating Cosmic Dawn, brought on by the formation of the first UV bright objects, from the complete Reionization of the IGM [See Dayal and Ferrara [2018]; Barkana and Loeb [2007] for reviews on the EoR]. The corresponding time frame has evolved much as our constraints and

---

[13] Here and throughout this manuscript, we address the Epoch of Hydrogen Reionization, and not the later phenomenon known as Helium Reionization.

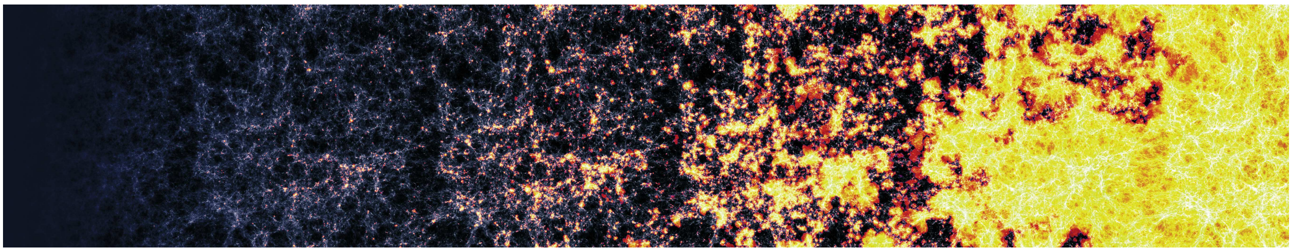


Figure 2: "Light cone" courtesy of N. Deparis. Generated from the CoDa II RHD cosmological simulation, time advances from left to right. The blue and white hues depict large scale structure and its evolution, whilst the bright yellows and reds portray the ionisation and temperature (respectively). This image, produced by assembling slices of the simulation at fixed coordinates over time, is highly illustrative of Reionization : one can see the first sources appear within the densest structures, and gradually grow ionised and hot bubbles, that finally meet which yields a fully ionised IGM.

knowledge thereof have improved. It is believed that the first and foremost (initially, at least) ionising sources were the first stars, formed at some time after z=30 (or 100 Myr after the Big Bang) [Bromm, 2017]. Analysis of the Lyα forests of distant quasars (see Sec. Constraints on Reionization) places the end of Reionization, when the network of ionised regions surrounding ionised sources completely merged, at some time between as late as z=5.3 [Becker et al., 2015a; Kulkarni et al., 2019] (roughly 1.1 Gyr after the Big Bang) and as early as z≈6 [Fan et al., 2006b] (about 0.94 Gyr after the Big Bang). The progress and morphology of this large scale transformation, is depicted in Fig. 2 from left to right.

Owing to its source driven nature, the study of the Epoch of Reionization is closely tied to many open questions in astrophysics. Indeed, to constrain the progress of Reionization, is to constrain the evolution of the density of ionising sources at cosmological scales. In turn, the study of Reionization provides a unique window into the study of the formation and evolution of the large scale structure that harbour stars (See Sec. Sources of ionisation for details) and active galactic nuclei (AGN). Finally, this leads to striving to understand the nature and birth of the first of these sources. Reionization then, provides a unique probe into cosmology and the early stages of formation of the structure we observe around us.

This unique perspective of study is ultimately rendered further interesting by the lasting impact of Reionization on the Universe. For instance, the intense ionising background radiation prevalent during the Epoch of Reionization is thought to have been able to suppress star formation in galaxies by heating and ionising the inter-stellar-medium (ISM) of galaxies. In doing so the ultra-violet background (UVBG) is thought to hinder the accretion of cold gas, as well as photo-evaporate low mass gas structures; both suppressing star formation in already formed dark matter haloes, as well as the formation of new haloes [Shapiro et al., 1994; Efstathiou, 1992]. This suppression effect has been put forward as a possible explanation for the missing satellite problem [Bullock and Boylan-Kolchin, 2017] : cosmological simulations over-predict the number of satellite galaxies in the vicinity of Milky Way-like galaxies, when compared to observations. Hence one possible explanation being the non-detection of a fraction of the local dwarf galaxies since they are either rendered fainter, or quenched by the radiative suppression arising during Reionization [Ocvirk et al., 2014; Brown et al., 2014; Dawoodbhoy et al., 2018; Katz et al., 2019a].

To complete this short review of the science of Reionization, we shall start by covering in detail how the first stars and galaxies formed, and how studies have estimated the contribution of galaxies to Reionization in Sec. Sources of ionisation. Then, we shall move on to Sec. Constraints on Reionization which deals with the observational evidence of Reionization. Finally, Sec. The simulation code RAMSES-CUDATON reviews the RAMSES-CUDATON code that we use to simulate Reionization.

Sources of ionisation

Here we cover the main sources that are thought to ionise the IGM : stars in the first galaxies. Active galactic nuclei are believed to be key in the later stages of Reionization, and in maintaining the ionisation of IGM's gas after the overlap of the ionised regions during the EoR [Kulkarni et al., 2019a]. However, we do not deal with them directly herein (they are not represented in the simulations we present further on), and therefore chose to omit them. Similarly, although x-ray binaries and gamma radio burst events are thought to happen during the time-frame of Reionization, they are too rare to contribute as significantly.

The first stars and galaxies

The first stars Understanding the physics of star formation is a difficult task in the Local Universe, and an active area of research. The theory of star formation in the high-redshift Universe is derivative of the local one. A complete review is beyond the scope of this manuscript, but exceedingly briefly :star formation is thought to occur when pressure supported giant molecular clouds collapse and fragment into clumps driven by gravity and turbulence. Further accretion from the gas in the cloud can then result in the formation of a proto-stellar core (for a full review, see McKee and Ostriker [2007]).

The first stars, therefore, form during the collapse of the first gaseous over-densities. In order to contract adiabatically, the first gaseous objects must be more massive than the Jeans mass, defined as (Eq. 2 from Barkana et al. [2001]):

$$\mathrm{M_J} = 5.73 \times \left(\frac{\Omega_{\mathrm{m0}}\mathrm{h}^2}{0.15}\right)^{-1/2} \left(\frac{\Omega_{\mathrm{b0}}\mathrm{h}^2}{0.022}\right)^{-3/5} \left(\frac{1+\mathrm{z}}{10}\right)^{3/2} \mathrm{M}_\odot\,, \tag{2}$$

where $\mathrm{M_J}$ is the Jeans mass, and h is the reduced hubble parameter or $\mathrm{H_0}/100$.

However, here we are more interested in the dark matter halo mass scale above which such baryonic objects can form in dark matter haloes. Following Barkana et al. [2001], and considering the hydrostatic equilibrium of baryons within a non-linear dark matter potential, we can find a minimum halo mass in order to achieve a given baryon over-density in its centre. This leads to an approximate minimum halo mass for the formation of the first stars of $\mathrm{M_{min}} \gtrsim 3 \times 10^4\,\mathrm{M}_\odot$ [Barkana et al., 2001]. In truth, the objects that are the closest to the limit form first, and it is they that host the first stars.

This initial phase of adiabatic collapse can only continue if the gas can cool. The first stars form

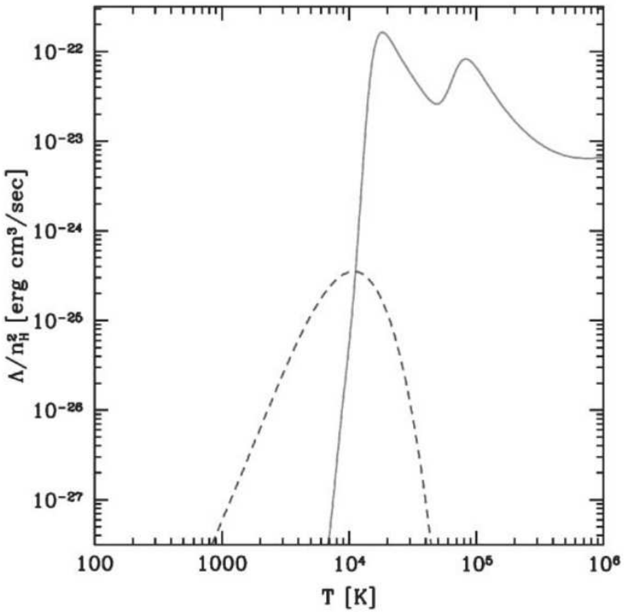


Figure 3: Atomic Hydrogen and Helium (full lines, first and second bump respectively) and molecular Hydrogen (dashed lines) cooling rates, [Λ/n_H^2 $\mathrm{ergs} \cdot \mathrm{cm}^3 \cdot \mathrm{s}^{-1}$] from Barkana and Loeb [2001].

in a pristine environment, where the dominant coolant above (below) 10^4 K[14] for the gas is atomic (molecular, H_2) hydrogen . Fig. 3 shows that under 10^3 K, there are no efficient cooling mechanisms for pristine Hydrogen gas. This sets another minimum gas temperature for the gas to collapse. In turn, and since the clouds are pressure supported, this means there is a minimum virial temperature and virial mass under which stars cannot form : around $5 \times 10^4\,M_\odot$ for molecular cooling, and $10^8\,M_\odot$ for atomic cooling (both at z=10). These cooling threshold masses scale with redshift as $\left(\frac{z+1}{10}\right)^{3/2}$.

In gasses that have been enriched in metals, other coolants allow the collapse of baryons into stars in smaller baryon clumps. This is the rationale for the well established idea that the first GMCs were hotter and more massive than in the Local Universe, and hence the first generations of stars (or pop IIIs) are thought to be more massive (anywhere in the order of 10 to $100 M_\odot$) [Bromm and Larson, 2004]. The properties of pop III stars are linked to their higher masses and their extremely low metallicities, which affect the nuclear processes that occur within the stars as they follow the main sequence of star formation. In short, and appropriately for our topic, pop III stars are considerably more emissive sources of LyC photons [Schaerer, 2002], and as very massive stars, are short lived (a few Myr) [See Bromm et al. [2009] for a succinct review].

Although the ionising radiation of the newly formed pop III stars can encourage H_2 formation by increasing the number density of available electrons by ionising HI gas, in fact, it has been shown that the soft ionising background (precisely, Lyman-Werner radiation $h\nu$ =11.2 eV to 13.6 eV and able to freely penetrate neutral atomic hydrogen, and propagate outwards from halos) rapidly dissociates H_2 before ionisation is complete [Haiman et al., 1997]. This leads to the complete suppression of H_2

---

[14] The dominant processes being the collisional excitation of Hydrogen and of Helium II, this requires a certain amount of gas ionisation, hence the high temperature cut-off.

cooling shortly after the first generation of stars appear. Therefore, in order for subsequent generations of stars to form, haloes must accrete and merge in order to reach virial masses $> 10^8\,\mathrm{M}_\odot$. This raises the virial temperature, rendering atomic cooling efficient, and allowing star formation.

Cosmological simulations have shown that the very first stars form in pristine gas between z=30 and z=20 (180 Myr after the Big Bang) in the centre of low mass dark matter haloes ($\sim 10^5\,\mathrm{M}_\odot$ to $10^6\,\mathrm{M}_\odot$), which lends itself well to the outlined theoretical predictions [Yoshida, 2019]. In addition, they find that the intense ionising radiation that pop III stars produce, could significantly delay (By approximately 10^2 Myr [Yoshida et al., 2006; Kitayama et al., 2004]) the creation of the second generation of stars. Not only does the ionising radiation rapidly photo-dissociate the H_2 molecules and thus the main coolant, but by heating and ionising the surrounding region, it can reduce gas accretion onto the halo, further delaying stellar formation. Supernova feedback from the first generation of stars also plays a similar role, but can also directly deprive haloes of their gas reservoirs by ejecting them.

The ejected gas from the first supernovas contains the metals produced by stellar and explosive nucleosynthesis in the first stars. The expelled metal rich gas gradually cools, partially falls back onto the halo from which it originates, where it mixes with the remaining pristine gas. The metallicity of the halo gas after mixing is difficult to predict, however simulations suggest that a single core collapse supernova of a pop III star can raise the halo gas metallicity to values of [Fe/H]=-5 to -3, thus allowing the formation of metal enriched stars analogous to the extremely metal-poor stars observed in the galactic halo Chiaki et al. [2018]; Chiaki and Wise [2019]. As this enrichment process continues, less and less of the total stellar mass is contained in pop III stars. This, in turn, means that the average number of LyC ionising photons per stellar baryon decreases.

As the enrichment process proceeds, and stellar generations succeed one another, the hierarchical growth of dark matter haloes and galaxies continues. From the seeds of H_2 cooling haloes of $10^5\,\mathrm{M}_\odot$ to $10^6\,\mathrm{M}_\odot$, more massive atomic cooling ($\gtrsim 10^8\,\mathrm{M}_\odot$) haloes emerge through continued accretion and mergers. Thus, allowing later more massive haloes to accumulate larger densities of gas, and ultimately form more stars, progressively populating the Universe with star forming galaxies.

Populating the Universe with galaxies Press and Schechter [1974] were the first to present a formalism for predicting the number density of collapsed virialised dark matter haloes (the halo mass function), as a function of their virial mass, and of time. Fig. 4 shows the evolution of the mass function over time for $10^2\,\mathrm{M}_\odot$ to $10^{15}\,\mathrm{M}_\odot$. The distinctive shape of the mass function is produced by the hierarchical nature of dark matter halo formation across time : more and more massive haloes form due to mergers of low mass haloes. Note that for z>5, the mass function increases between each redshift for all plotted masses. For z≤5, this is no longer the case as the low mass end of the halo mass function is gradually depleted, by fuelling the creation of more and more massive haloes. At these later times, fewer new dark matter haloes are forming, and the total number of predicted haloes falls as the more numerous less massive haloes merge.

At the same time, it is possible to estimate the large scale build up of star formation. Using stellar population synthesis models applied to high redshift galaxy surveys, the density of star formation can be estimated across time from the spectral energy density (SED) of galaxies [for instance Madau et al., 1998; Kennicutt, 1998]. Fig. 5 shows such an estimation of the universal star formation rate

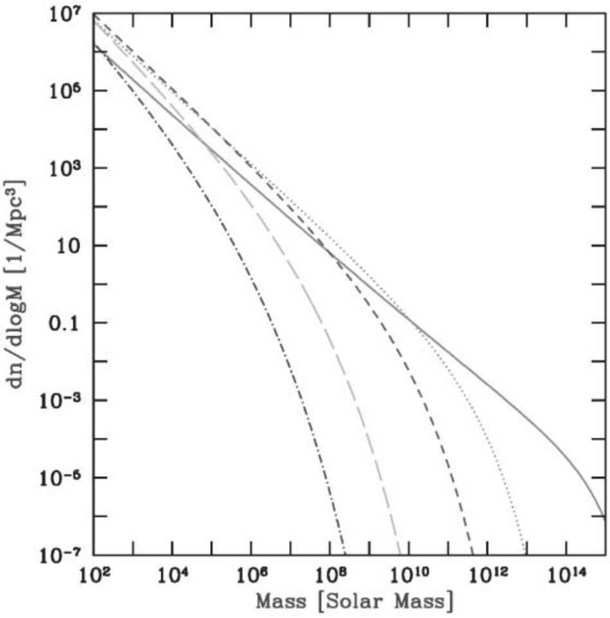


Figure 4: Halo mass functions from Barkana and Loeb [2001]. Different curve correspond to different redshifts : z = 0(solid curve), 5(dotted curve), 10(short-dashed curve), 20(long-dashed curve), and 30(dot–dashed curve).

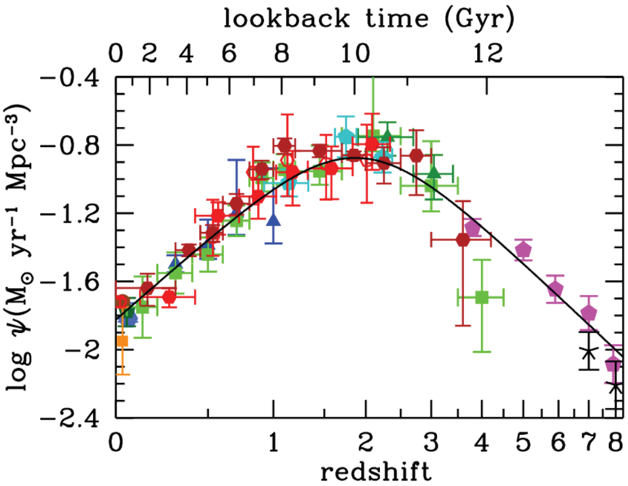


Figure 5: Joint far UV and far IR estimations of the cosmic star formation rate density over time from Madau and Dickinson [2014].

density from Madau and Dickinson [2014]. The total star formation rate increases with decreasing redshift until a maximum near z ≈ 2 (often referred to as "Cosmic noon"), and decreasing until z=0. The increase of the SFRD during the EoR (z$\gtrsim$ 6 in the figure), suggests that as time goes on more and more photons could be available to reionise the IGM.

By observing the cosmic evolution of the number of galaxies, and the amount of star formation, its possible to estimate their contribution to Reionization.

Galaxies propelling Reionization

Observations In order to estimate the participation of galaxies to reionising the Universe, one must estimate the number of ionising photons produced in galaxies that ionise the IGM. A first approach then, is to count observed galaxies as a function of their UV luminosity during the EoR, in a large volume. The left panel of Fig. 6 shows the spatial density of UV galactic sources per magnitude (at λ=1600 Å) for several redshifts (the ultra-violet luminosity function or UVLF), observed with the Hubble Space Telescope. The shape and redshift evolution of the UVLF is reminiscent of the hierarchical formation of galaxies and mirrors the shape of the mass function of Fig. 4 : the brightest galaxies form in the most massive dark matter haloes where star formation is the most efficient, and are therefore rarer and appear later (at lower redshifts) than more abundant lower magnitude galaxies.

Since we have determined that there is a minimum mass under which dark matter haloes do not host star formation, and since stars have a minimum mass and UV luminosity set by their formation conditions, we can infer that there is some minimum UV luminosity at a given epoch for galaxies. Furthermore, star formation in dwarf galaxies could be suppressed by the feedback from supernovas and ionising radiation during Reionization, further decreasing the UVLF at the faint end [as seen in Ocvirk et al., 2016, 2020]. Consequently, the rise of the UVLF in the left panel of Fig. 6 should stall for the faintest magnitudes. This trend change in the UVLF, commonly referred to as the 'turnover', is a crucial ingredient in determining the relative contribution of faint galaxies to Reionization, as well as the total contribution of galaxies. Indeed the fainter the magnitude where the turnover occurs, the greater the total amount of star forming galaxies producing photons, and the greater the relative contribution of faint galaxies to the total number of produced ionising photons is. The turnover could occur at magnitudes fainter than -14 [compatible with the observations from Atek et al., 2018]. To observe such faint objects during the EoR requires deep surveys, or fields lensed by galaxy clusters. The right panel of Fig. 6 shows the impact of lensing uncertainties on the shape of the UVLF at low magnitudes. It is thought that the planned JWST mission will extend the flux limit of high redshift galaxies by around two magnitudes, which may prove sufficient to constrain the turnover of the UVLF, thereby placing better constraints on the role of galaxies in reionising the Universe.

As we have already seen, we encounter difficulties when counting galaxies, due to the brightness and completeness limitations of observations. However, ideally one would need the density of ionising photons produced by galaxies to estimate their role in Reionization. As it turns out, actually estimating the number of ionising photons that galaxies produce adds many layers of complexity.

In Fig. 6, we showed the abundance of high redshift galaxies as a function of their magnitude at λ=1600 Å. In fact, the ionising luminosities of reionising galaxies are extremely difficult to measure

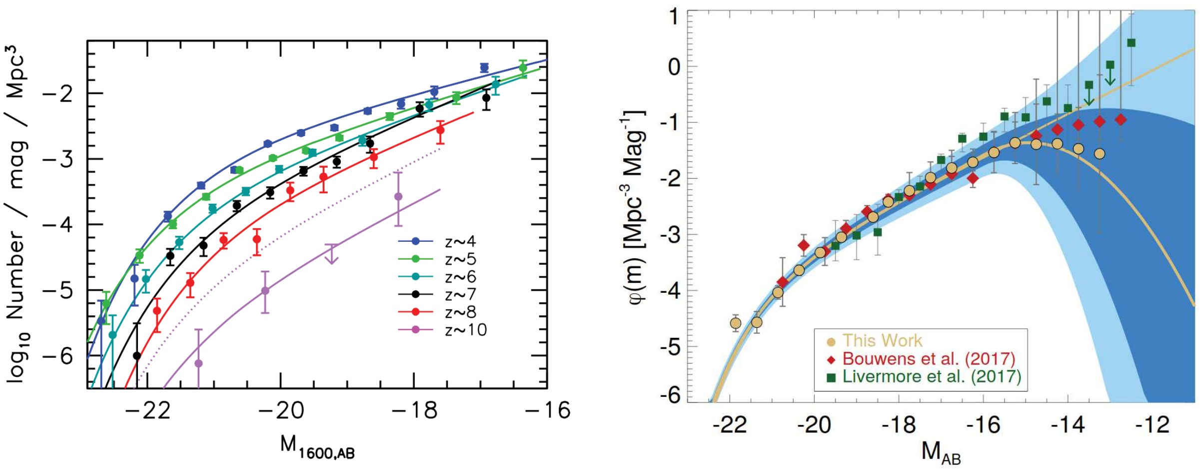


Figure 6: *Left*: UVLF at z=4, 5, 6, 7, 8, 10 from Bouwens et al. [2015]. *Right*: UVLFs from various groups at z=6 from Atek et al. [2018]. Note the blue regions, which correspond to confidence regions derived using lensing uncertainties relevant to the computation of the UVLF.

directly due to the large column densities of neutral H gas in the IGM before the end of Reionization ($z \gtrsim 6$, and depending on the sight line, before z $\gtrsim$ 5.5). Though this gives an indication of the density of galactic UV sources at a given time, measuring the luminosity of sources in wavelength bands $\lambda > 912$Å[15], gives no direct constraint on the density of ionising sources, or on the density of ionising photons. Therefore converting the observed UV luminosities to ionising UV luminosities requires extensive assumptions and modelling.

Modelling Generally [for instance in Robertson et al., 2013; Finkelstein et al., 2019], studies take that :

$$\dot{\mathrm{n}}_{\mathrm{ion}} = \rho_{\mathrm{UV}}\xi_{\mathrm{ion}}\mathrm{f}_{\mathrm{esc}}\,, \tag{3}$$

Where $\dot{\mathrm{n}}_{\mathrm{ion}}$ is the total galaxy ionising emissivity, ρ_{UV} is the UV photon density, and ξion is the Lyman continuum photon emission rate per unit UV. Finally, $\mathrm{f}_{\mathrm{esc}}$ is the fraction of produced ionising photons that traverse the ISM and circum-galactic matter (CGM) of galaxies without being absorbed, leading to possible ionisation of the IGM.

Simply, the density of UV photons can be determined by integrating the UVLF weighted luminosities over the observed magnitude range. To determine ξ_{ion}, used to convert from UV luminosities to ionising UV luminosities, Robertson et al. [2013] compare the observed average slope of the UV continuum as a function of redshift, with predictions from stellar populations synthesis models. Making assumptions about the stellar ages, metallicities and dust content of the stellar populations, then allows one to estimate ξ_{ion}. Another approach is to estimate ξ_{ion} using H_α observations, and an empirically calibrated relation between H_α and LyC in galactic spectral energy distributions [as done in Bouwens et al., 2016]. This allows for a more direct probe of ξ_{ion}.

---

[15] Or h$\nu < 13.6$eV, i.e.: Photons without enough energy to ionise a Hydrogen atom from its ground state.

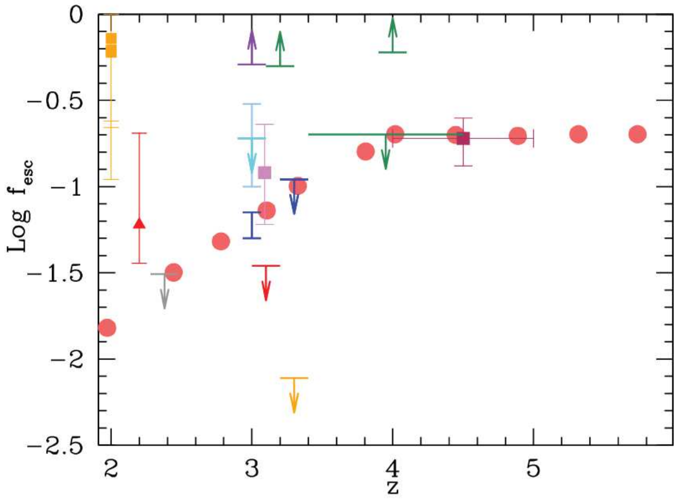


Figure 7: Figure taken from Dayal and Ferrara [2018], showing a compilation of f_{esc} measurements using various techniques. The red filled dots represent an analytical fit to observation data performed by Inoue et al. [2006].

Finally, one must assume what fraction of the produced LyC photons reaches the IGM from the neutral and optically thick[16] ISM and IGM of galaxies, or f_{esc}. Indeed, it has been shown that large fractions of the LyC light produced within galaxies can be absorbed by galactic HI gas [for instance Matthee et al., 2017]. Though a key ingredient for understanding Reionization by galaxies, it is a difficult quantity to measure, as it also relies on multi-wavelength observations and assumptions about the stellar populations of galaxies. Fig. 7 displays a compilation of f_{esc} values of various groups by Dayal and Ferrara [2018]. The few measurements that exist are mostly at low redshift and vary from almost 1.0 to $\lesssim$ 1%. Observations of f_{esc} in the high-redshift Universe prove rare and more difficult. During Reionization, it is very challenging to measure directly, due to the large column densities of neutral Hydrogen in the Universe before Reionization end.

To constrain $\dot{n}_{ion}$, and the parameters from Eq. 3 requires further assumptions and modelling. Madau et al. [1999] show that the global evolution of the filling factor of ionised Hydrogen can be described by :

$$\dot{Q}_{H_{II}} = \frac{\dot{n}_{ion}}{\bar{n_H}} - \frac{Q_{H_{II}}}{t_{rec}} \,, \tag{4}$$

where $\dot{Q}_{H_{II}}$ is the rate of change of the ionised Hydrogen filling factor $Q_{H_{II}}$, $\dot{n}_{ion}$ is the total ionising emissivity, $\bar{n_H}$ is the average Hydrogen density, and t_{rec} is the recombination timescale.

By solving the differential equation in Eq. 4, it's possible to place constraints on the minimum value of $\dot{n}_{ion}$ necessary to ionise the Universe, and to maintain ionisation[17]. These in turn can be used to place indirect constraints on $\xi_{ion} f_{esc}$ [as in Robertson et al., 2013, 2015; Bouwens et al., 2016;

---

[16] Both due to absorption by neutral Hydrogen and to dust grains.

[17] Because ionised atoms can recombine, and assuming each ionisation requires one photon (i.e. :there are no multiple ionisations from high energy photons), then Reionization required several photons per Hydrogen atom

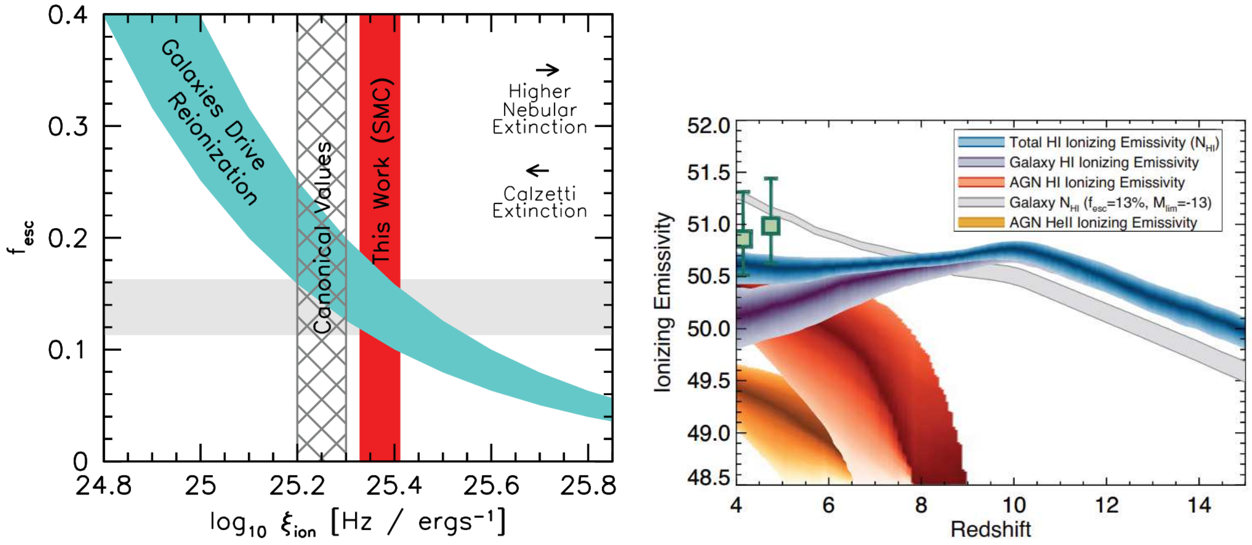


Figure 8: *Left*: Constraints on f_{esc} from Bouwens et al. [2016], using measurements of ξ_{ion} inferred from UV and H_α observations. *Right*: Total ionising emissivity in $ph \cdot s^{-1} \cdot cMpc^{-3} \cdot h^3$ as predicted by Finkelstein et al. [2019]. Becker and Bolton [2013] compare data from observations of the Lyα forest to numerical simulations to constrain the total ionisation rate of Hydrogen. In combination with estimations of the mean free path of ionising photons, this allows them to constrain the total ionising emissivity from all sources. The result is shown by the two green square markers.

Finkelstein et al., 2019].

The value assumed for f_{esc} in Eq. 3 is often arbitrary and represents the average f_{esc} of the population of galaxies. Sometimes recipes or models inspired by numerical simulations are used. The left panel of Fig. 8 shows the constraints on the galactic average f_{esc} depending on the value of ξ_{ion}, and the assumed UV extinction by dust in Bouwens et al. [2016]. In fact, dust in high redshift galaxies is a relatively unknown but plausibly decisive parameter. Observations have shown that there could be large dust masses, (and large dust mass to stellar mass ratios) in massive star forming galaxies during Reionization [e.g. Burgarella et al., 2020; Laporte et al., 2017; Béthermin et al., 2015]. Underestimating the dust masses of such galaxies could dramatically alter the observed UVLF due to extinction, and thus the predictions made for the ionising emissivities of galaxies. At the same time, high column densities of dust in galaxies could dramatically alter their f_{esc} by reprocessing ionising photons.

Recent modelling efforts seem to suggest that a varying $\xi_{ion} f_{esc}$ is required so that the total ionising emissivity better matches constraints after z=6 [e.g. Puchwein et al., 2019; Finkelstein et al., 2019]. In fact, they report that this leads to the total ionising emissivity reaching a peak during Reionization, after which it decreases. The right panel of Fig. 8 shows the total ionising emissivity predictions in Finkelstein et al. [2019]. To determine the galaxy ionising emissivity, they used the f_{esc} from high resolution numerical simulations. Note the very different trends of the total ionising emissivity from galaxies when they assume a constant global value for f_{esc} [as in Robertson et al., 2013, 2015, for instance]. In the latter case the total $\dot{n}_{ion}$ from galaxies increases for all redshifts, whereas in the former it decreases by a factor of ≈ 3 from z=10 to z=4. This illustrates the strong impact of one's assumptions about the $\mathrm{f_{esc}}$ of galaxies on the ionising emissivities of galaxies.

Simulations In this context of difficult observations, and many free model parameters, simulations are a useful tool to explore the physical processes and galactic properties that govern the ionising emissivities and the ionising photon budget, as well as the escape of ionising photons to the IGM during Reionization. Although dark matter only simulations have long been a popular and successful tool in cosmology [for instance Springel et al., 2005], Reionization is driven and experienced by baryons : stellar baryons produce the ionising radiation that reionises the baryons in the IGM. Ideally, simulations should use fully coupled radiative-hydrodynamic simulations. Further, in order to be able to follow individual stars, properly capturing their formation, and the effects of stellar and radiative feedback, simulations should feature sub-parsec resolution. At the same time, the simulated boxes should span hundreds of Mpc so as to realistically reproduce the inhomogeneous nature of Reionization [Iliev et al., 2014], and represent the cosmic variance inherent in Reionization. However, such simulations are computationally very far from our grasp. As of such, numerical studies of Reionization have varied from very detailed parsec level simulations of molecular clouds [Kim et al., 2019], to detailed galaxy simulations (e.g. : Yoo et al. [2020]; Trebitsch et al. [2017]), through simulations of galaxy clusters[Trebitsch et al., 2020], to Mpc scale Universe simulations (as in Katz et al. [2018, 2019]; Rosdahl et al. [2018]; Ocvirk et al. [2016, 2020]).

Escape fraction Of the work that studied $\mathrm{f_{esc}}$ using simulations, one could distinguish higher resolution simulations (those that resolve the ISM), from coarser ones that do not, the larger ($\gtrsim 10\,\mathrm{cMpc \cdot h^{-1}}$) boxes from the smaller ones, and those that perform fully coupled radiation-hydrodynamics (RHD) simulations from those that post-process radiative transfer. However, beyond this more superficial categorization, there are substantial differences that reside in the implementation of sub-grid physics such as star formation, and stellar feedback mechanisms, that affect the $\mathrm{f_{esc}}$. The smallest, most detailed simulations that resolve the ISM [with parsec scale resolution, i.e. Kimm et al., 2017; Wise et al., 2014; Paardekooper et al., 2015] seem to show $\mathrm{f_{esc}}$ values that decrease on average with halo mass from $\approx 10^6\,\mathrm{M_\odot}$ to $10^9\,\mathrm{M_\odot}$. Coarser simulations that do not resolve the ISM [with kpc resolution or more, as in Kimm and Cen, 2014; Katz et al., 2018; Yajima et al., 2011] show similar trends with mass, but for more massive objects of $\mathrm{M_{halo}} \approx 10^8\,\mathrm{M_\odot}$ to $10^{11}\,\mathrm{M_\odot}$. Using a fixed dust to metal ratio, Yajima et al. [2011] found that dust could affect $\mathrm{f_{esc}}$ by $\approx 10\%$ at z=6. Interestingly, the simulations of Razoumov and Sommer-Larsen [2010], that employ zoom-in simulations that can resolve the ISM in galaxies of similar masses ($10^8\,\mathrm{M_\odot}$ to $10^{12}\,\mathrm{M_\odot}$), also find decreasing $\mathrm{f_{esc}}$ with halo mass. Using a fixed dust to metal ratio, they found that dust could reduce $\mathrm{f_{esc}}$ by an increasing amount as time goes on, giving up to 10% lower $\mathrm{f_{esc}}$ values in some galaxies. Ma et al. [2020] use a similar technique but with higher resolution, and found that for the zoom-in simulations with the highest resolution, $\mathrm{f_{esc}}$ increases between $10^8\,\mathrm{M_\odot}$ to $10^{11}\,\mathrm{M_\odot}$ before stagnating. In the simulations of their highest mass haloes, that have poorer resolution, they report a downward trend of $\mathrm{f_{esc}}$ with halo mass for masses $\gtrsim 5 \times 10^{10}\,\mathrm{M_\odot}$. In fact, they show that this trend is driven by high dust densities in high mass galaxies (they use a similar dust prescription to Yajima et al. [2011]; Razoumov and Sommer-Larsen [2010]). Overall, although this sheds light on disagreements on the values of $\mathrm{f_{esc}}$, it also shows that there is some form of agreement, as most studies report some downward trend of $\mathrm{f_{esc}}$ with increasing halo mass. Interestingly, it seems that dust could have a significant effect on $\mathrm{f_{esc}}$ at high redshift. In turn, probably effecting the photon budget. However, it remains to be seen if this is also true in simulations that directly model dust, avoiding simple prescriptions that convert gas metallicity to dust.

The global evolution with redshift of the SFR weighted $\mathrm{f_{esc}}$ is studied in Kimm and Cen [2014]; Katz et al. [2018]; Rosdahl et al. [2018]. In both cases, relatively flat trends are reported, with values fluctuating around a mean value of nearly $\mathrm{f_{esc}}= 0.1$. Studies of the time variability of $\mathrm{f_{esc}}$ in idividual galaxies show that there is considerable fluctuation in the values of $\mathrm{f_{esc}}$[such as in Trebitsch et al., 2017; Rosdahl et al., 2018]. For instance, Trebitsch et al. [2017]; Ma et al. [2020] reported a correlation between bursts of supernova explosions and increases in $\mathrm{f_{esc}}$, separated by a time delay.

Photon budget Studying the photon budget in simulations requires large enough, well resolved enough, boxes to capture the full range of atomic cooling haloes that hosted star forming galaxies, and participated in Reionization. As of such, few studies can directly compute the photon budget. For instance, Sharma et al. [2016] use a recipe based on the star formation density in the EAGLE simulations [Crain et al., 2015] to compute $\mathrm{f_{esc}}$, and find that the brightest galaxies drove Reionization. This conclusion puts their results at odds with most of the literature, as the consensus that seems to emerge is that high mass, bright galaxies had the least significant contribution to Reionization. For instance, Anderson et al. [2017] post-processed radiative transfer around sources in a smaller coupled N-body and hydrodynamics simulation, and found that faint galaxies ($\mathrm{M_{AB1600}}$>-17) drove Reionization, contributing more than 85% of photons at any single time. Following a comparable process, but including a prescription for dust, Yajima et al. [2011] reach a similar conclusion. They find that the haloes $\lesssim 3 \times 10^9\,\mathrm{M_\odot}$ contribute around 45% of the ionising photons at z=6. Using fully coupled RHD simulations of Reionization, Katz et al. [2018, 2019] found that the photon budget was dominated by intermediate mass haloes ($10^9\,\mathrm{M_\odot}$ to $10^{10}\,\mathrm{M_\odot}$). It could be that when not using fully coupled RHD simulations, the effect of feedback on star formation is not properly modelled, yielding higher contributions to the photon budget by low mass and faint haloes. Katz et al. [2019] also show the dominant contributors to Reionization evolve over time. Using a different approach, based on very high resolution, fully coupled RHD zoom-in simulations, Ma et al. [2020] compute the $\mathrm{f_{esc}}$ of their highly resolved galaxies. By convolving their reported relation between $\mathrm{f_{esc}}$ and the halo mass function, they also find that Reionization was dominated by halos of intermediate masses at z=6. Just as in Katz et al. [2018], they report that earlier on, near z=10, the dominant contributors were lower mass haloes.

Constraints on Reionization

In this section we briefly list some of the main constraints on the evolution of the ionised fraction of Hydrogen in the IGM, used extensively to calibrate simulations. For complete reviews, see Becker et al. [2015b]; McQuinn [2016].

The Lyα line is a resonant emission line of Hydrogen. One of the main astrophysical sources of Lyα photons is the recombining gas in the $\mathrm{H_{II}}$ regions generated by ionising sources such as galaxies, or quasars. Due to its resonant nature, it appears as a very strong spectral line. Accordingly, it is used to detect high redshift ionising sources. Because of the line's associated high cross section [see Dijkstra, 2017, for a full review of Lyα physics], the line centre and blue wing are often dampened by absorption from surrounding residual neutral Hydrogen gas. In fact, very low residual fractions of neutral Hydrogen ($\gtrsim 10^{-4}$) can absorb most flux blue-wards of the line. It is this property that makes the Lyα line a successful probe of the neutral fraction of the IGM. Let us consider photons blue-ward of the Lyα line emitted before the end of Reionization. As the ionising photons travel, they are redshifted, so that when they encounter neutral gas at a redshift of z, some of the photons that were emitted at $\lambda < \lambda_{\mathrm{Ly}\alpha}$[18] are redshifted into the blue wing of the Lyα line and absorbed. The photons that are redder than the Lyα line at z can pass freely and reach the observer. When the photons encounter a sufficiently ionised pocket in the IGM, they pass freely, creating a bump in the final observed spectrum around $\lambda = (1+\mathrm{z})\lambda_{\mathrm{Ly}\alpha}$. As Reionization progresses and the ionised pockets begin to grow in size and spatial frequency, wider and wider portions of the observed spectrum become highly transmissive. The residual spikes in transmission that are observed in the spectra of high redshift quasars are referred to as the Lyα forest. The area of low transmission blue-wards of the Lyα line is the Gunn-Peterson (GP) trough. A collection of high redshift quasar spectra exhibiting these features is shown in Fig. 9. The Lyα line is apparent, and shifts in the spectra according to the source's redshift. At z=6.42, only a portion of the spectrum blue-wards of the Lyα line has non-zero transmission. Whereas at z=5.74, many more wavelengths on the blue side of the Lyα line have higher transmissions. This highlights the ongoing changes in the IGM as Reionization progresses. The figure also illustrates the heterogeneous nature of Reionization, as the Lyα forest transmission varies from source to source at similar redshifts, when the average ionisation was similar.

Gunn and Peterson [1965] first predicted an absence of signal blue-wards of the Lyα line of distant quasars. Because the transmission in the Lyα forest is linked to the abundance and ionisation of ionised bubbles along the line of sight (LoS) of the observed source, the Lyα forest can be used to constrain the volume filling factor of ionised Hydrogen, and the progression of Reionization. However, since the Lyα forest can be entirely absorbed along lines of sight with very low (10^{-4}) fractions of neutral Hydrogen, this is only possible in the later stages of Reionization, when a considerable volume is already ionised. Briefly, we summarize some of the constraints that can be obtained on Reionization by analysing the spectra of distant quasars.

- The average transmission in the Lyα forest gives the effective optical depth over a certain portion of the LoS. Using some model of the density of Hydrogen gas, it is then possible to estimate the average neutral fraction along the line of sight [as in Fan et al., 2006].
- For the highest redshift spectra, there may be few spikes in transmission in the Lyα forest. In those cases, the remainder of the spectrum blue-wards of Lyα is 'dark', or completely absorbed. By counting the number and length of the dark portions of such spectra, and making assumptions about the growth and size of ionised bubbles around sources, and about the abundance of sources, estimates of the neutral fraction of Hydrogen can be made [see Fan et al., 2006].
- The dampening of the blue wing of the Lyα emission of quasars depends on the size and ionisation

---

[18] 1215.6 Å

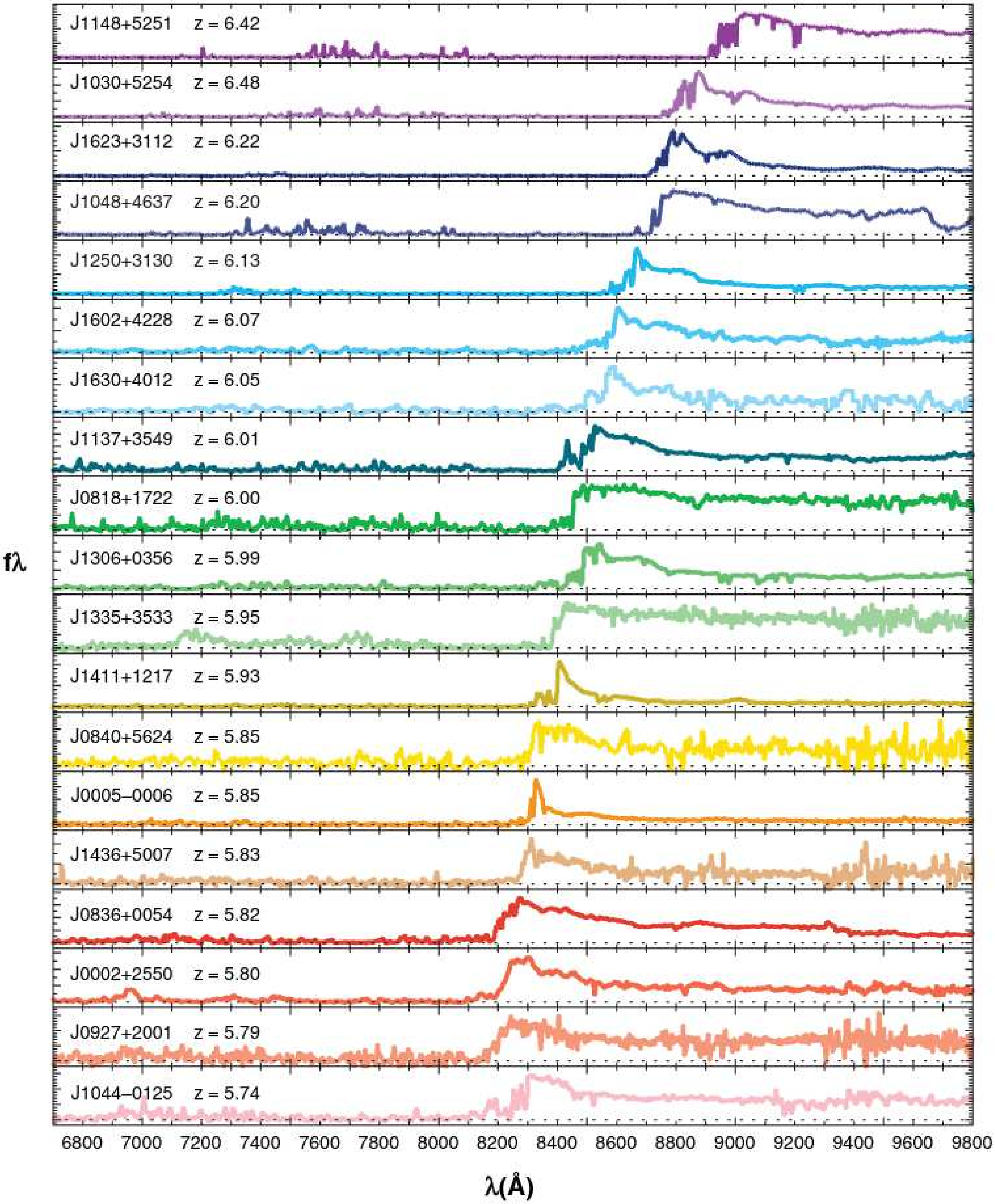


Figure 9: The spectra from 19 high redshift quasars from Fan et al. [2006], organised by descending redshift.

of the ionised bubble in which they reside. It is possible then, to produce measurements of the ionization in the environments of quasars [see Ďurovčíková et al., 2019] (Similar measurements are made using the dampening of Lyα emission from region ionised by gamma ray bursts [for instance Totani et al., 2016]). Ionising sources in large enough ionised regions may show little to no dampening in the blue wing of their Lyα emission. As Reionization progresses, the number of regions large enough for this to occur increases. Thus, the evolution of the equivalent widths and clustering[19] of detected Lyα lines, can be linked to the evolution of the ionisation of the IGM.

Free electrons in the IGM can scatter and polarise CMB photons, dampening the small scale CMB anisotropies [as described in Zaldarriaga, 1997]. The total optical depth due to scattering by free electrons is τ. Because it is an integrated optical depth, it constrains the average ionisation history of the IGM, and thus the timing and length of Reionization is linked to the value of τ. Some of the most recent constraints are from **?**.

The simulation code RAMSES-CUDATON

In this section, we outline the physical and numerical choices in the RAMSES-CUDATON code for radiative hydrodynamical simulations of the Universe. First, we examine RAMSES, then we shall focus on ATON, finally we explain the coupling of the two codes.

RAMSES

RAMSES [Teyssier, 2002] is a coupled N-body and hydrodynamics simulation code. It features an Eulerian scheme with adaptive mesh refinement (AMR), allowing for higher resolution where it is most impactful and at a lesser computational cost than a full high resolution simulation.

Adaptive mesh refinement

In RAMSES, data is organised into a hierarchical system of groups of 8 cells[20] or octs, at several levels of refinement. Adaptive mesh refinement allows local increases of resolution. Refinement is triggered by a user defined criterion; for cosmological uses, dark matter density, gas density or both are used. Following this system, octs are grouped into trees starting at the base level of no refinement. Every time a cell meets the conditions for refinement, it is represented at a higher level of refinement by a new oct. Care is taken to ensure a smooth transition between refinement levels. Note that this feature is disabled in RAMSES-CUDATON.

N-body dynamics

Though RAMSES is a eulerian code, that uses cells and grids to compute dynamics (as opposed to a smooth particle hydrodynamics, lagrangian approach), it tracks dark matter and stars in the form of

---

[19] At high enough redshifts there are few regions large enough for this to happen, so fraction of spectra detection with some transmission blue-wards of Lyα must cohabit in a smaller number of regions.

[20] if 3 dimensions are used

particles. To solve the Vlasov-Poisson equations (Eq. 5) that describe collision-less N-body systems, it is therefore necessary to interpolate the positions of the particles onto densities on a eulerian grid. This is done using a cloud in cell interpolation scheme.

$$\begin{cases} \frac{\mathrm{d}\mathbf{x}_\mathrm{p}}{\mathrm{dt}} & = & \mathbf{v}_p \,, \\ \frac{\mathrm{d}\mathbf{v}_\mathrm{p}}{\mathrm{dt}} & = & -\nabla_x \phi \,, \\ \text{With } \Delta_\mathrm{x}\phi & = & 4\pi G\rho \,, \end{cases} \tag{5}$$

where particles with positions and velocities : $\mathbf{x}_\mathrm{p}$ and $\mathbf{v}_\mathrm{p}$, form a density field ρ, and with t the time in the eulerian referential, and G the gravitational constant.

The potential can then be computed using the interpolated densities. At the coarsest level (no refinement) it is computed using a Fast Fourier Transform (FFT) technique, whereas at refined levels the potential is computed using a 7-points finite difference approximation of the Laplacian, with Dirichlet boundary conditions. Then, acceleration can be computed using a 5-points finite difference approximation of the gradient. Afterwards, an inverse cloud in cell interpolation scheme is used to assign an acceleration to each particle. Finally, particle velocities and positions are updated using a second order midpoint scheme with variable time steps.

Hydrodynamics

In RAMSES, gas dynamics are tracked by solving the conservative Euler equations :

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) & = & 0 \,, \\ \frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p & = & -\rho \nabla \phi \,, \\ \frac{\partial(\rho e)}{\partial t} + \nabla \cdot (\rho \mathbf{u}[e + p/\rho]) & = & -\rho \mathbf{u} \cdot \nabla \phi \,, \end{cases} \tag{6}$$

where ρ is gas density, t is the time in the Eulerian reference frame, $\mathbf{u}$ is the velocity vector, e is the specific total energy, p the gas thermal pressure, ϕ the gravitational potential. We also have that :

$$p = (\gamma - 1)\rho(e - \frac{1}{2}u^2) \,, \tag{7}$$

Where γ=5/3 is the adiabatic index (assuming a mono-atomic perfect gas).

This set of equations can be generalised as follows :

$$\frac{\partial \mathrm{U}}{\partial \mathrm{t}} + \nabla \cdot \mathbf{F}(\mathrm{U}) = \mathrm{S}(\mathrm{U}) \,, \tag{8}$$

where U is the conserved quantity, $\mathbf{F}$(U) is its associated flux, and S(U) its associated source term. In this form, the equations are then solved using a second-order unsplit Godunov scheme, that uses a Riemann solver to compute the fluxes at cell boundaries.

Star formation and Supernova feedback

Star formation In RAMSES, star formation occurs according to a phenomenological approach based on a Schmidt law [Schmidt, 1959]. Stars are tracked by massive stellar particles that represent entire stellar populations, who's masses are based on the gas masses of cells. In practice, the stellar particle masses are multiples of a minimum stellar mass based on a user defined gas density threshold. Star formation is allowed in cells that exceed a gas density criterion of $\rho_{\rm SF}$. When this is the case, the mass of each new stellar particle is determined by sampling a discrete Poisson distribution of expected value $\frac{\rm M_{gas}}{\rm M_\star}$, where $\rm M_{gas}$ is the gas mass available for star formation (set by Eq. 9), and $\rm M_\star$ is the minimum stellar mass.

$$\mathrm{SFR} = \epsilon_\star \frac{\rho_{\rm gas}}{\rm t_{ff}}, \text{only if } \rho_{\rm gas} > \rho_{\rm SF}\,, \tag{9}$$

where $\epsilon_\star$ is the star formation efficiency, $\rho_{\rm gas}$ is the gas density, and $\rm t_{ff} = \sqrt{\frac{3\pi}{32 G \rho_{gas}}}$. This setup ensures that $\mathrm{SFR} \propto \rho_{\rm gas}^{1.5}$ [hence the reference to Schmidt, 1959]. Upon its creation, every particle is assigned a metallicity value corresponding to the gas metallicity in the cell where it is created.

Supernova feedback When stellar particles are older than 10 Myr, supernovas are triggered for a mass fraction of the stellar population of the particle (corresponding to the mass fraction of massive stars) : $\eta_{\rm SN}$. This number also gives the fraction of the stellar particle's mass that is ejected by supernovas following the kinetic feedback described in Dubois and Teyssier [2008]. At the same time, 10^{51} ergs of thermal energy are injected into the gas per $10\,\rm M_\odot$ of ejected stellar mass. The ejected gas is enriched in metals according to the yield parameter : a yield fraction of the ejected gas is ejected in the form of gaseous metals, the rest is ejected as a gas with the same metallicity as the gas that formed the stellar particle.

ATON

ATON [Aubert and Teyssier, 2008] is a code for computing radiative transfer and the non-equilibrium ionisation chemistry of Hydrogen. It solves the integrated moment equations of radiative transfer, using the M1 closure model [Levermore, 1984].

Radiative transfer

Radiative transfer equations The radiative transfer equation is given by :

$$\frac{1}{\rm c}\frac{\partial \rm I_\nu}{\partial \rm t} + \mathbf{n}\cdot\nabla \rm I_\nu = -\kappa_\nu \rm I_\nu + \eta_\nu\,, \tag{10}$$

where c is the speed of light, $\rm I_\nu(\mathbf{x}, \mathbf{n}, t)$ is the radiation specific intensity, $\kappa_\nu(\mathbf{x}, \mathbf{n}, \rm t)$ is the absorption coefficient, and $\eta_\nu(\mathbf{x}, \mathbf{n}, \rm t)$ is the source term. $\mathbf{x}$ is the position vector, $\mathbf{n}$ is a normalised direction vector, and t is time.

The absorption term couples the photons to the gas as $\kappa_\nu = \rm n_{HI}\sigma_\nu$. σ_ν is the cross section for the frequency ν, and $\rm n_{HI}$ is the neutral density of Hydrogen. The source term can be split into an ionising stellar source field $\dot{\rm N}_\nu^\star$, and a diffuse recombination radiation term $\dot{\rm N}_\nu^{\rm rec} = \rm n_e n_{HII}(\alpha_A - \alpha_B)$. Where $\rm n_e$ is the density of free electrons, $\rm n_{HII}$ is the density of free protons, $\alpha_{\rm A}$ and $\alpha_{\rm B}$ are the case A and the case B recombination rates.

By taking the first two momenta of Eq. 10, dividing them by a single photon's energy (hν), and integrating over frequency (accordingly, we substitute the $_\nu$ subscripts with $_\gamma$ to represent the photon group and not a single frequency), we obtain :

$$\begin{cases} \frac{\partial \mathrm{N}_\gamma}{\partial \mathrm{t}} + \nabla \cdot \mathbf{F}_\gamma & = & -\mathrm{n}_{\mathrm{HI}} \mathrm{c} \sigma_\gamma \mathrm{N}_\gamma + \mathrm{n}_{\mathrm{e}} \mathrm{n}_{\mathrm{HII}} (\alpha_{\mathrm{A}} - \alpha_{\mathrm{B}}) + \dot{\mathrm{N}}^\star_\gamma \,, \\ \frac{\partial \mathbf{F}_\gamma}{\partial \mathrm{t}} + \mathrm{c}^2 \nabla \cdot \mathbf{P}_\gamma & = & -\mathrm{n}_{\mathrm{HI}} \mathrm{c} \sigma_\gamma \mathbf{F}_\gamma \,, \end{cases} \tag{11}$$

where N_γ is the number of group photons, $\dot{\mathrm{N}}^\star_\gamma$ is the number of ionising photons from sources, $\mathbf{F}_\gamma$ is the flux of group photons, and $\mathbf{P}_\gamma$ is the pressure tensor.

Thermochemistry

The full coupling of gas and photons requires a set of equations that track Hydrogen gas ionisation and temperature. The state of the Hydrogen gas is governed by a balance between collisional ionisation, photo-ionisation, and collisional recombination. The evolution of the neutral density of Hydrogen (n_{HI}) can be expressed as follows :

$$\frac{\mathrm{D}\mathrm{n}_{\mathrm{HI}}}{\mathrm{Dt}} = \alpha_{\mathrm{A}} \mathrm{n}_{\mathrm{e}} \mathrm{n}_{\mathrm{HII}} - \beta \mathrm{n}_{\mathrm{e}} \mathrm{n}_{\mathrm{HI}} - \Gamma_{\gamma,\mathrm{HI}} \mathrm{n}_{\mathrm{HI}} \,, \tag{12}$$

where α_{A} is the case a recombination rate, β is the collisional ionisation rate, and $\Gamma_{\gamma,\mathrm{HI}} = \mathrm{c}\sigma_\gamma \mathrm{N}_\gamma$.

At the same time, the thermal evolution of the gas is set by :

$$\rho_{\mathrm{gas}} \frac{\mathrm{D}}{\mathrm{Dt}} \left(\frac{\mathrm{e}}{\rho_{\mathrm{gas}}} \right) = \mathcal{H} - \mathcal{L} \,, \tag{13}$$

where ρ_{gas} is the gas density, $\mathcal{H}$ is the total heating rate, and $\mathcal{L}$ is the total cooling rate. e is the specific gas kinetic energy, e= $\frac{3}{2} \mathrm{n}_{\mathrm{tot}} \mathrm{k}_{\mathrm{b}} \mathrm{T}$.

The total cooling rate $\mathcal{L}$ is given by cell gas density and temperature using fits of the cooling rates due to collisional ionisation and recombination, given by Hui and Gnedin [1997]; Maselli et al. [2003]. The total heating rate due to photo-ionisation is given by $\mathrm{c}\epsilon_\gamma \sigma_\gamma \mathrm{N}_\gamma$, where ϵ_γ is the average photon group energy.

M1 closure of the moment equations To close the set of equations in Eq. 11, we need to express $\mathbf{P}_\gamma$ as a function of N_γ and $\mathbf{F}_\gamma$. In ATON, the M1 approximation is used as a closure relation [Levermore, 1984]. It is assumed that the angular distribution of radiation is axisymmetric around the flux vector, giving $\mathbf{P} = \mathbf{D}\mathrm{N}$ and allowing the following expression of the Eddington tensor $\mathbf{D}$:

$$\mathbf{D} = \frac{1-\chi}{2} \mathbf{I} + \frac{3\chi - 1}{2} \mathbf{u} \otimes \mathbf{u} \,, \tag{14}$$

where $\mathbf{D}$ is the Eddington tensor, $\mathbf{u}$ is a unit vector aligned with the flux vector $\mathbf{F}$, and χ is the Eddington factor. The expression of $\mathbf{D}$ describes the geometry of radiation with a linear combination of two geometries : the first term is an isotropic diffusive term, and the second is a transport term and aligned with the local flux.

χ is defined using the reduced local ionising flux $\mathbf{f}$. If we define $\mathbf{f} = \frac{\mathbf{F}}{\mathrm{cN}} = \mathrm{f}\mathbf{u}$, then Levermore [1984] give that :

$$\chi = \frac{3+4|\mathbf{f}|^2}{5+2\sqrt{4-3|\mathbf{f}|^2}}\,, \tag{15}$$

With this definition of *chi*, the value of the reduced local ionising flux dictates the geometry of radiation :when the local reduced ionising flux is 1, the diffusive term of **D** is 0; when the local reduced ionising flux is 0, the flux aligned transport term of **D** is null.

The M1 closure is easy to adapt to Eulerian codes such as RAMSES. Solving the integrated moment equations circumvents the numerical cost of ray tracing methods when there are many sources separated by large distances[21]. However, M1 remains an approximation. For instance, one caveat of M1 is that photons have a fluid like behaviour, meaning opposing fluxes of photons collide instead of passing through one another.

Solving the resulting coupled equations An operator splitting scheme is used in ATON to solve the above sets of coupled equations. In practice, the solving of the various equations is split into steps.

- Source step : The previous source field (photon densities) are updated. In RAMSES-CUDATON, the new source field is obtained from RAMSES, that sums the ionising luminosities of every stellar particle in each cell before calling ATON.
- Transport step : Here the transport (without a source term) is solved. The system of equations obtained in Eq. 11 can be re-expressed as a system of hyperbolic conservation laws, just like in Eq. 8 in RAMSES, and are solved using an explicit second order Godunov scheme. Because the speed of light is so great, the Courant stability condition[22] imposes very short time steps compared to the time steps of the RAMSES code.
- Thermochemical step : Here the full set of coupled gas/photon equations are solved. Because of the smaller time scales associated with the atomic processes involved, an implicit scheme is chosen to guarantee the stability of the solution. Since the temperature update depends only on the updated local values of N_γ, F_γ, T, and x (the ionised fraction) , it can be solved separately in an explicit manner. Due to the extremely small time scales involved, and to guarantee stability, the temperature equations and cooling rates are sub-cycled for each cell until temperature converges.

Coupling of RAMSES and ATON

The small time steps imposed by the explicit transport step in ATON, are much smaller than the time steps required for stability in solving the N-body dynamics and hydrodynamics in RAMSES. In order to speed up the coupled code, one approach has been to adopt a reduced or adaptive speed of light, making the Courant condition less strict. However, this can result in sizeable differences in the Reionization history and ionising front speeds in simulation [as show in Ocvirk et al., 2019; Deparis

---

[21] In ray tracing schemes, the numerical cost scales as $\mathrm{N}_{\mathrm{rays}}\mathrm{N}_{\mathrm{sources}}$. Furthermore, since resolution decreases as the distance from a source increases, many rays are required to properly resolve radiative transfer in cosmological simulations. In the case of ATON, the numerical cost increases as the number of cells, independently of the number of sources.

[22] $3\mathrm{c}\Delta\mathrm{t}/\Delta\mathrm{x} < 1$, where Δt is the time step, Δx is the spatial resolution, and c is the speed of light.

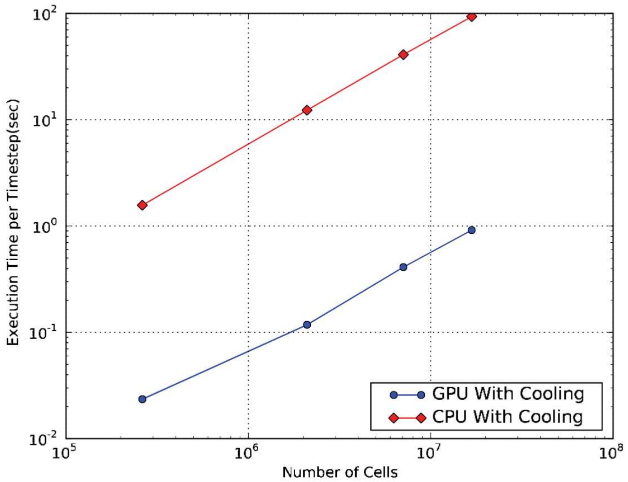


Figure 10: ATON (blue) and CUDATON (red) performance as a function of the size of the Eurlerian grid, when post-processing RAMSES cosmological simulations using CPUs (ATON) or GPUs (CUDATON). From Aubert and Teyssier [2010].

et al., 2019]. This is the technique used in RAMSES-RT Rosdahl et al. [2013]. In RAMSES-CUDATON, the computational cost of ATON's short time steps is offset by the increases in performance obtained when using the GPU version of ATON : CUDATON [Aubert and Teyssier, 2010].

At the end of every RAMSES time step (on CPUs), the source field is computed and sent to CUDATON (on GPUs) along with the Hydrogen gas state (ionisation fraction, density, temperature) and the fluxes and photon densities from the last CUDATON step. Once the new CUDATON step is finished, it sends back new ionised fractions, temperature, fluxes, and photon densities to RAMSES, to be used in the next RAMSES step. Fig. 10 shows the considerable speed gain (≈ 100 times faster) of CUDATON (on GPUs) when compared to ATON (on CPUs). It is this performance gain, that significantly reduces the cost of radiative transfer, and allows RAMSES-CUDATON to run large scale cosmological simulations with fully coupled RHD and a full speed of light. However, this impressive performance gain is in part due to aggressive optimisation of the CUDATON code on GPUs, that precludes running the code with AMR. RAMSES-CUDATON must therefore be run on a uniform single grid.

BIBLIOGRAPHY

Alam S., et al.. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 470:2617–2652, 2017.

Anderson L., et al.. The little Galaxies that could (reionize the universe): predicting faint end slopes & escape fractions at z>4. *Mon Not R Astron Soc*, 468(4):4077–4092, 2017.

Atek H., et al.. The extreme faint end of the UV luminosity function at z $\sim$ 6 through gravitational telescopes: a comprehensive assessment of strong lensing uncertainties. *Mon Not R Astron Soc*, 479(4):5184–5195, 2018.

Aubert D. and Teyssier R. A radiative transfer scheme for cosmological reionization based on a local Eddington tensor. *Mon Not R Astron Soc*, 387(1):295–307, 2008.

Aubert D. and Teyssier R. Reionization Simulations Powered by Graphics Processing Units. I. On the Structure of the Ultraviolet Radiation Field. *The Astrophysical Journal*, 724:244–266, 2010.

Barkana R. and Loeb A. In the beginning: the first sources of light and the reionization of the universe. *Physics Reports*, 349(2):125–238, 2001.

Barkana R. and Loeb A. The physics and early history of the intergalactic medium. *Reports on Progress in Physics*, 70:627–657, 2007.

Barkana et al. The Reionization of the Universe by the First Stars and Quasars. *Annual Review of Astronomy and Astrophysics*, 39(1):19–66, 2001.

Becker G.D. and Bolton J.S. New measurements of the ionizing ultraviolet background over $2 < z < 5$ and implications for hydrogen reionization. *Monthly Notices of the Royal Astronomical Society*, 436:1023–1039, 2013.

Becker G.D., et al.. Evidence of patchy hydrogen reionization from an extreme Lyα trough below redshift six. *Monthly Notices of the Royal Astronomical Society*, 447:3402–3419, 2015a.

Becker G.D., et al.. Reionisation and High-Redshift Galaxies: The View from Quasar Absorption Lines. *Publications of the Astronomical Society of Australia*, 32:e045, 2015b.

Blandford R.D., et al.. The distortion of distant galaxy images by large-scale structure. *Monthly Notices of the Royal Astronomical Society*, 251:600–627, 1991.

Bocquet S., et al.. Cluster Cosmology Constraints from the 2500 deg2 SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope. *The Astrophysical Journal*, 878:55, 2019.

Bouwens R.J., et al.. UV Luminosity Functions at Redshifts z ∼ 4 to z ∼ 10: 10,000 Galaxies from HST Legacy Fields. *ApJ*, 803(1):34, 2015.

Bouwens R.J., et al.. The Lyman-Continuum Photon Production Efficiency ξ ion of z ~ 4-5 Galaxies from IRAC-based Hα Measurements: Implications for the Escape Fraction and Cosmic Reionization. *The Astrophysical Journal*, 831:176, 2016.

Boylan-Kolchin M., et al.. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Monthly Notices of the Royal Astronomical Society*, 415:L40–L44, 2011.

Bromm V. The first stars: our evolving theoretical picture. *Memorie della Societa Astronomica Italiana*, 88:671, 2017.

Bromm V. and Larson R.B. The First Stars. *Annual Review of Astronomy and Astrophysics*, 42:79–118, 2004.

Bromm V., et al.. The formation of the first stars and galaxies. *Nature*, 459(7243):49–54, 2009.

Brown T.M., et al.. The Quenching of the Ultra-faint Dwarf Galaxies in the Reionization Era. *The Astrophysical Journal*, 796:91, 2014.

Bull P., et al.. Beyond Λ CDM: Problems, solutions, and the road ahead. *Physics of the Dark Universe*, 12:56–99, 2016.

Bullock J.S. and Boylan-Kolchin M. Small-Scale Challenges to the ΛCDM Paradigm. *Annual Review of Astronomy and Astrophysics*, 55:343–387, 2017.

Burgarella D., et al.. Observational and theoretical constraints on the formation and early evolution of the first dust grains in galaxies at $5 < z < 10$. *arXiv*, page arXiv:2002.01858, 2020.

Béthermin M., et al.. Evolution of the dust emission of massive galaxies up to $z = 4$ and constraints on their dominant mode of star formation. *A&A*, 573:A113, 2015.

Chiaki G. and Wise J.H. Seeding the second star: enrichment from population III, dust evolution, and cloud collapse. *Monthly Notices of the Royal Astronomical Society*, 482:3933–3949, 2019.

Chiaki G., et al.. Metal-poor star formation triggered by the feedback effects from Pop III stars. *Monthly Notices of the Royal Astronomical Society*, 475:4378–4395, 2018.

Crain R.A., et al.. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. *Monthly Notices of the Royal Astronomical Society*, 450:1937–1961, 2015.

Cyburt R.H., et al.. Big bang nucleosynthesis: Present status. *Reviews of Modern Physics*, 88:015004, 2016.

Dawoodbhoy T., et al.. Suppression of star formation in low-mass galaxies caused by the reionization of their local neighbourhood. *Monthly Notices of the Royal Astronomical Society*, 480(2):1740, 2018.

Dayal P. and Ferrara A. Early galaxy formation and its large-scale effects. *PhR*, 780:1–64, 2018.

Deparis N., et al.. Impact of the reduced speed of light approximation on ionization front velocities in cosmological simulations of the epoch of reionization. *Astronomy and Astrophysics*, 622:A142, 2019.

Dijkstra M. Saas-Fee Lecture Notes: Physics of Lyman Alpha Radiative Transfer. *arXiv e-prints*, 1704:arXiv:1704.03416, 2017.

Dubois Y. and Teyssier R. On the onset of galactic winds in quiescent star forming galaxies. *Astronomy and Astrophysics*, 477:79–94, 2008.

Efstathiou G. Suppressing the formation of dwarf galaxies via photoionization. *Monthly Notices of the Royal Astronomical Society*, 256:43P–47P, 1992.

Eisenstein D.J. and Hu W. Baryonic Features in the Matter Transfer Function. *The Astrophysical Journal*, 496:605–614, 1998.

Fan X., et al.. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132:117–136, 2006a.

Fan X., et al.. Observational Constraints on Cosmic Reionization. *Annual Review of Astronomy and Astrophysics*, 44:415–462, 2006b.

Finkelstein S.L., et al.. Conditions for Reionizing the Universe with a Low Galaxy Ionizing Photon Escape Fraction. *The Astrophysical Journal*, 879:36, 2019.

Flores R.A. and Primack J.R. Observational and theoretical constraints on singular dark matter halos. *The Astrophysical Journal Letters*, 427:L1–L4, 1994.

Gunn J.E. and Peterson B.A. On the Density of Neutral Hydrogen in Intergalactic Space. *The Astrophysical Journal*, 142:1633–1636, 1965.

Haiman Z., et al.. Destruction of Molecular Hydrogen during Cosmological Reionization. *The Astrophysical Journal*, 476:458–463, 1997.

Hikage C., et al.. Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data. *Publications of the Astronomical Society of Japan*, 71:43, 2019.

Hubble E. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, 1929.

Hui L. and Gnedin N.Y. Equation of state of the photoionized intergalactic medium. *Mon Not R Astron Soc*, 292(1):27–42, 1997.

Iliev I.T., et al.. Simulating cosmic reionization: how large a volume is large enough? *Mon Not R Astron Soc*, 439(1):725–743, 2014.

Inoue A.K., et al.. The escape fraction of ionizing photons from galaxies at z = 0-6. *Monthly Notices of the Royal Astronomical Society*, 371:L1–L5, 2006.

Katz H., et al.. A Census of the LyC Photons that Form the UV Background During Reionization. *arXiv:1802.01586 [astro-ph]*, 2018. ArXiv: 1802.01586.

Katz H., et al.. How to Quench a Dwarf Galaxy: The Impact of Inhomogeneous Reionization on Dwarf Galaxies and Cosmic Filaments. *arXiv e-prints*, page arXiv:1905.11414, 2019a.

Katz H., et al.. Tracing the sources of reionization in cosmological radiation hydrodynamics simulations. *Monthly Notices of the Royal Astronomical Society*, 483:1029–1041, 2019b.

Kennicutt Jr. R.C. Star Formation in Galaxies Along the Hubble Sequence. *Annual Review of Astronomy and Astrophysics*, 36:189–232, 1998.

Kim J.G., et al.. Modeling UV Radiation Feedback from Massive Stars. III. Escape of Radiation from Star-forming Giant Molecular Clouds. *The Astrophysical Journal*, 883:102, 2019.

Kimm T. and Cen R. Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars. *The Astrophysical Journal*, 788:121, 2014.

Kimm T., et al.. Feedback-regulated star formation and escape of LyC photons from mini-haloes during reionization. *Monthly Notices of the Royal Astronomical Society*, 466:4826–4846, 2017.

Kitayama T., et al.. The Structure and Evolution of Early Cosmological H II Regions. *The Astrophysical Journal*, 613:631–645, 2004.

Klypin A., et al.. Where Are the Missing Galactic Satellites? *The Astrophysical Journal*, 522:82–92, 1999.

Knox L. and Millea M. Hubble constant hunter's guide. *Physical Review D*, 101:043533, 2020.

Kulkarni G., et al.. Evolution of the AGN UV luminosity function from redshift 7.5. *Monthly Notices of the Royal Astronomical Society*, 488:1035–1065, 2019a.

Kulkarni G., et al.. Large Ly α opacity fluctuations and low CMB τ in models of late reionization with large islands of neutral hydrogen extending to $z < 5.5$. *Monthly Notices of the Royal Astronomical Society: Letters*, 485(1):L24–L28, 2019b.

Laporte N., et al.. Dust in the Reionization Era: ALMA Observations of a z = 8.38 Gravitationally Lensed Galaxy. *The Astrophysical Journal Letters*, 837:L21, 2017.

Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques - NASA/ADS. 1927.

Levermore C.D. Relating Eddington factors to flux limiters. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 31(2):149–160, 1984.

Ma X., et al.. No missing photons for reionization: moderate ionizing photon escape fractions from the FIRE-2 simulations. *arXiv:2003.05945 [astro-ph]*, 2020. ArXiv: 2003.05945 version: 1.

Madau P. and Dickinson M. Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics*, 52:415–486, 2014.

Madau P., et al.. The Star Formation History of Field Galaxies. *ApJ*, 498(1):106, 1998.

Madau P., et al.. Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources. *The Astrophysical Journal*, 514:648–659, 1999.

Maselli A., et al.. CRASH: a radiative transfer scheme. *Monthly Notices of the Royal Astronomical Society*, 345:379–394, 2003.

Matthee J., et al.. The origin of scatter in the stellar mass-halo mass relation of central galaxies in the EAGLE simulation. *Monthly Notices of the Royal Astronomical Society*, 465:2381–2396, 2017.

McKee C.F. and Ostriker E.C. Theory of Star Formation. *Annual Review of Astronomy and Astrophysics*, 45(1):565–687, 2007.

McQuinn M. The Evolution of the Intergalactic Medium. *Annual Review of Astronomy and Astrophysics*, 54:313–362, 2016.

Ocvirk P., et al.. The Reionization of Galactic Satellite Populations. *The Astrophysical Journal*, 794:20, 2014.

Ocvirk P., et al.. Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe. *Monthly Notices of the Royal Astronomical Society*, 463(2):1462–1485, 2016.

Ocvirk P., et al.. Impact of the reduced speed of light approximation on the post-overlap neutral hydrogen fraction in numerical simulations of the epoch of reionization. *Astronomy and Astrophysics*, 626:A77, 2019.

Ocvirk P., et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 2020.

Paardekooper J.P., et al.. The First Billion Years project: the escape fraction of ionizing photons in the epoch of reionization. *Mon Not R Astron Soc*, 451(3):2544–2563, 2015.

Peebles P.J.E., et al.. Masses of rich clusters of galaxies as a test of the biased cold dark matter theory. *The Astrophysical Journal*, 347:563–574, 1989.

Perlmutter S., et al.. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *The Astrophysical Journal*, 517:565–586, 1999.

Planck Collaboration, et al.. Planck 2018 results. X. Constraints on inflation. *arXiv e-prints*, 1807:arXiv:1807.06211, 2018.

Press W.H. and Schechter P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *The Astrophysical Journal*, 187:425–438, 1974.

Puchwein E., et al.. Consistent modelling of the meta-galactic UV background and the thermal/ionization history of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 2019.

Razoumov A.O. and Sommer-Larsen J. Ionizing Radiation from z = 4–10 Galaxies. *ApJ*, 710(2):1239, 2010.

Robertson B.E., et al.. New Constraints on Cosmic Reionization from the 2012 Hubble Ultra Deep Field Campaign. *The Astrophysical Journal*, 768:71, 2013.

Robertson B.E., et al.. Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope. *The Astrophysical Journal*, 802:L19, 2015.

Rosdahl J., et al.. RAMSES-RT: radiation hydrodynamics in the cosmological context. *Monthly Notices of the Royal Astronomical Society*, 436:2188–2231, 2013.

Rosdahl J., et al.. The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Monthly Notices of the Royal Astronomical Society*, 479:994–1016, 2018.

Schaerer D. On the properties of massive Population III stars and metal-free stellar populations. *Astronomy and Astrophysics*, 382:28–42, 2002.

Schmidt M. The Rate of Star Formation. *The Astrophysical Journal*, 129:243, 1959.

Scolnic D.M., et al.. The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample. *The Astrophysical Journal*, 859:101, 2018.

Shapiro P.R., et al.. Reionization in a cold dark matter universe: The feedback of galaxy formation on the intergalactic medium. *The Astrophysical Journal*, 427:25–50, 1994.

Sharma M., et al.. The brighter galaxies reionized the Universe. *Mon Not R Astron Soc Lett*, 458(1):L94–L98, 2016.

Springel V., et al.. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, 2005.

Teyssier R. Cosmological hydrodynamics with adaptive mesh refinement - A new high resolution code called RAMSES. *A&A*, 385(1):337–364, 2002.

Totani T., et al.. High-precision analyses of Lyα damping wing of gamma-ray bursts in the reionization era: On the controversial results from GRB 130606A at z = 5.91. *Publications of the Astronomical Society of Japan*, 68:15, 2016.

Trebitsch M., et al.. Fluctuating feedback-regulated escape fraction of ionizing radiation in low-mass, high-redshift galaxies. *Mon Not R Astron Soc*, 470(1):224–239, 2017.

Trebitsch M., et al.. The Obelisk simulation: galaxies contribute more than AGN to HI reionization of protoclusters. *arXiv e-prints*, 2002:arXiv:2002.04045, 2020.

Viel M., et al.. Probing the intergalactic medium with the Lyα forest along multiple lines of sight to distant QSOs. *Monthly Notices of the Royal Astronomical Society*, 329:848–862, 2002.

Walker T.P., et al.. Primordial nucleosynthesis redux. *The Astrophysical Journal*, 376:51–69, 1991.

Wise J.H., et al.. The Birth of a Galaxy - III. Propelling reionisation with the faintest galaxies. *Monthly Notices of the Royal Astronomical Society*, 442(3):2560–2579, 2014.

Yajima H., et al.. Escape fraction of ionizing photons from high-redshift galaxies in cosmological SPH simulations. *Mon Not R Astron Soc*, 412(1):411–422, 2011.

Yoo T., et al.. On the origin of low escape fractions of ionizing radiation from massive star-forming galaxies at high redshift. *arXiv e-prints*, 2001:arXiv:2001.05508, 2020.

Yoshida N. Formation of the first generation of stars and blackholes in the Universe. *Proceeding of the Japan Academy, Series B*, 95:17–28, 2019.

Yoshida N., et al.. Formation of Primordial Stars in a ΛCDM Universe. *The Astrophysical Journal*, 652:6–25, 2006.

Zaldarriaga M. Polarization of the microwave background in reionized models. *Physical Review D*, 55:1822–1829, 1997.

Zavala J. and Frenk C.S. Dark Matter Haloes and Subhaloes. *Galaxies*, 7:81, 2019.

Ďurovčíková D., et al.. Reionization history constraints from neural network based predictions of high-redshift quasar continua. *arXiv e-prints*, 1912:arXiv:1912.01050, 2019.

Galactic ionising photon budget in Cosmic Dawn II

1

1.1 Presenting Cosmic Dawn II

Before motivating Cosmic Dawn II's usage for determining the ionising photon budget of galaxies, we must first remind ourselves of its characteristics. In this section I give a brief overview of the simulation, and its published results. The published article presenting the simulation is featured at the end of the section, and can furnish the reader with a more complete view of the simulation, and important definitions for the rest of the manuscript. At the same time, the article allows the reader to better grasp the context of this work.

1.1.1 RAMSES-CUDATON

Cosmic Dawn II (CoDa II) is a simulation performed in the context of the CoDa project, that was described in Ocvirk et al. [2020] (included further on in the chapter). The CoDa II simulation employs the RAMSES-CUDATON code [Ocvirk et al., 2016], utilising a unique approach that couples the computation of the dynamics of dark matter, hydrodynamics and star formation on CPUs using RAMSES [Teyssier, 2002], with the numerically expensive computation of radiative transfer and Hydrogen photo-chemistry and cooling on GPUs using ATON [Aubert and Teyssier, 2008]. This hybrid CPU/GPU approach gives performance increases that allow the circumvention of a commonly used performance approximation : reducing the speed of light. Typically, this could result in unrealistic propagation speeds of ionising fronts, as well as lower ionisation rates [see : Ocvirk et al., 2019; Deparis et al., 2019; Gnedin, 2016]. The approach herein suffers from neither of these setbacks. However, the aggressive optimisation of ATON on GPU in RAMSES-CUDATON requires the use of a fixed grid, precluding the use of AMR. As of such, to produce all the simulations discussed in this manuscript, RAMSES-CUDATON was run using a uni-grid with single global simulation grids.

1.1.2 Simulation set-up

Aiming at producing large scale RHD simulations of the Reionization of the Universe, the CoDa II simulation's most important specifications are given in table 1.1. The large size of the CoDa II simulation (94.44 cMpc), allows for a massive sample of galaxies[1] during Reionization, assuring a proper sampling of the dark matter halo mass function. Such large scales are essential in order to realistically portray the EoR, as too small boxes under represent the brightest galaxies, fundamentally affecting the inhomogeneous geometry of Reionization. Iliev et al. [2014] show that volumes greater than $\approx 100\,\mathrm{cMpc} \cdot \mathrm{h}^{-1}$ are required in order for realisations to properly depict the full range of bright galaxies, and to not be susceptible to cosmic variance.

---

[1] over 1.31×10^7 detected dark matter haloes ranging in mass from 2×10^7 to 1.5×10^{12} $\mathrm{M}_\odot$ at z=6

CoDa II's resolution was set so as to resolve the least massive atomic cooling star forming dark matter haloes during Reionization ($M_{\text{halo}} \gtrsim 10^8\,M_\odot$). With a physical force resolution of 3.3 kpc at z=6, $10^8\,M_\odot$ CoDa II haloes are described by a handful of resolution elements[2]. Massive ($10^{11}\,M_\odot$ and above) haloes are resolved by several thousands of cells, which is sufficient to describe detailed and intricate structure within the virial radius[3], but not the structures of the ISM.

CoDa II was run from z=150 onwards using constrained initial conditions from the CLUES (Constrained Local UniversE Simulations) project Sorce et al. [2016]. The initial conditions were constrained using the velocity fields in the Local Universe. When used to perform simulations down to z=0, they result in structure that echoes the local structure of the Universe (in mass, velocity, and position). However, due to the numerical cost of running the full simulation, it was stopped at z=5.8. In order to use the CoDa II simulation to study the progenitors of Local Group structure during Reionization, a companion dark matter simulation of resolution 2048^3 and 4096^3 was ran down to z=0.

The ambitious scope of the CoDa II simulation, that combines large scales, high resolution (for a Reionization simulation), and fully coupled RHD, required the use of considerable computing resources. CoDa II was run on the Oak Ridge National Laboratory's Titan supercomputer, on 65 536 CPU cores and 16 384 GPUs. The shear size of the 4096^3 grid and the number of physical fields, that span over 2 PB of data, means that storing and analysing the CoDa II simulation is a challenge in of itself. So as to lower the memory imprint of data analysis, but also to ease storage and transfer of the data, the various physical fields (gas density, temperature, photon density, photon flux, ionisation fraction, 3d gas velocity) were cut into 512 512^3 cell sub cubes of data. Dark matter particle data (3d position and 3d velocity, ids) is stored in the original output formats of RAMSES, in one file per MPI process. Stellar particle data was extracted into ascii files (containing id, 3d position, mass and age).

Dark matter halo detection in CoDa II was performed using the Friends of Friends algorithm [Roy et al., 2014] with a linking length of 0.2. A spherical definition of DM haloes was adopted, that uses the widespread convention of assimilating the r_{200} of DM haloes to their virial radius to define their limits. We explore some of the drawbacks of this definition in Sec. 1.2.4.

1.1.3 Results

Realistic depiction of Reionization

Fig. 1.1.3 shows a combination image of CoDa II I produced at z=7, when Reioinization is well underway. It effectively illustrates the realistic heterogeneous, "patchy" nature of Reionization in CoDa II. Indeed, at this redshift there are both large ionised regions formed from the merger of many individual regions, and small bubbles around galaxies that are just beginning to ionise the surrounding IGM. In between these extrema, a wide variety of ionised areas of many shapes and sizes exist. The regions with the greatest densities of gas (attracted by the DM over-densities) form the brightest objects with the most stars, and are ionised first. These objects sit in the central areas of the largest ionised patches. The bright pink/red hues that lie at the centres of these most massive galaxies, show

---

[2] The r_{200} of such haloes, representative of the virial radius, is ≈ 0.5 cells, and increases by a factor of 3 for every dex in halo mass

[3] That we approximate as r_{200} as described further on in Sec. 1.2.4

Cosmology (Planck 2014)	
Ω_{Λ}	0.693
Ω_{m}	0.307
Ω_{b}	0.048
H_0	$67.7\,\mathrm{km}\cdot\mathrm{s}^{-1}\cdot\mathrm{Mpc}^{-1}$
σ_{s}	0.8288
n	0.963
Resolution	
Grid size	4096^3
Comoving box size	$94.44\,\mathrm{cMpc}$ $(64\,\mathrm{cMpc}\cdot\mathrm{h}^{-1})$
Comoving force resolution	$23.06\,\mathrm{ckpc}$
Physical force resolution at z=6	$3.3\,\mathrm{kpc}$
Dark matter particle number	4096^3
Dark matter particle mass	$4.07\times10^{5}\,\mathrm{M}_{\odot}$
Stellar particle mass	$11\,732\,\mathrm{M}_{\odot}$
Star formation and feedback	
Density threshold for star formation	$50{<}\rho_{\mathrm{gas}}{>}$
Star formation efficiency	0.02
Massive star lifetime	$10\,\mathrm{Myr}$
Supernova energy	10^{51} ergs
Radiation	
Stellar ionising emissivity	$4.32\times10^{46}\,\mathrm{ph}\cdot\mathrm{s}^{-1}\cdot\mathrm{M}_{\odot}^{-1}$
Stellar particle escape fraction	0.42
Effective photon energy	$20.28\,\mathrm{eV}$
Effective HI cross-section	$2.493\times10^{-22}\,\mathrm{m}^2$

Table 1.1: Some of the essential parameters of the CoDa II simulation. See Ocvirk et al. [2020] for the full list.

regions of very high temperatures ($>10^5$ K) induced by supernova explosions. Within the confines of the large blue ionised regions, there are slightly red hued waves that show the rise in temperature in recently ionised gas.

In CoDa II, Reionization ends near z=6. Ocvirk et al. [2020] show that CoDa II agrees well with several constraints on the timing of Reionization : The evolution of the average neutral fraction and the average ionised fraction agree well with constraints during Reionization, from z=10 to z≈6. Though, after Reionization, the average neutral fraction is too low when compared to constraints from the Lyα forest [Fan et al., 2006]. At the same time, the average J_{21} (integrated flux) in the simulation is too high after z=6, suggesting that there are too many photons after the overlap of the ionised regions around z=6. The simulation's average Thomson optical depth is a good match with observations from Planck [Planck Collaboration et al., 2018]. CoDa II's UVLF agrees very well with the observed UVLF of Bouwens et al. [2015]; Finkelstein et al. [2015]; Atek et al. [2018], showing that the recipes for star formation and feedback used in CoDa II, in combination with the computation of dark matter and gas dynamics, yield highly realistic source densities. This is fundamental in order to properly describe Reionization in a simulation, and further hints at CoDa II's massive population of galaxies being realistic.

Feedback and star formation suppression

CoDa II showcases well the impact of radiative feedback on star formation, which occurs predominantly in the haloes of mass $\lesssim 10^9\,M_\odot$. Suppression from both radiative and stellar feedback mainly affects the least massive haloes ($\lesssim 10^9\,M_\odot$), but also has an appreciable impact on the star formation in $10^9\,M_\odot$ to $10^{10}\,M_\odot$ haloes. Suppression is already apparent by z=10, and increases rapidly between z=7 and z=6. In particular, Ocvirk et al. [2020] show the considerable impact of radiative feedback on star formation in the lower mass star formation haloes of CoDa II at z=5. Examination of the relationship between local over-density and SFR in low mass galaxies, showed that suppression happens earlier in the highest over-density regions where star formation is stronger earlier on, due to the proximity with the most massive, brightest galaxies. Star formation, and its suppression, arise later and more intensely in the under-dense regions of CoDa II. This only happens after z=6, when the ionised regions of the simulation have overlapped, and the under-dense regions are flooded by the UV background. This is in line with the results from Dawoodbhoy et al. [2018], based on CoDa I.

Reionization of progenitors of Local Universe objects

Reionization time maps suggest that the Milky Way and M31 reionised separately as islands. As did the Virgo cluster, and the groups of M81 and CenA. This further supports what was hinted at by previous work [Ocvirk et al., 2013], and CoDa I AMR [Aubert et al., 2018].

1.1.4 Galaxy sample

Fig. 1.2, 1.3, 1.4, 1.5, 1.6, 1.7 show maps from a small sample of the galaxies simulated in CoDa II at z=6. From top to bottom, and from left to right each 3x3 panel shows : ionising photon flux,

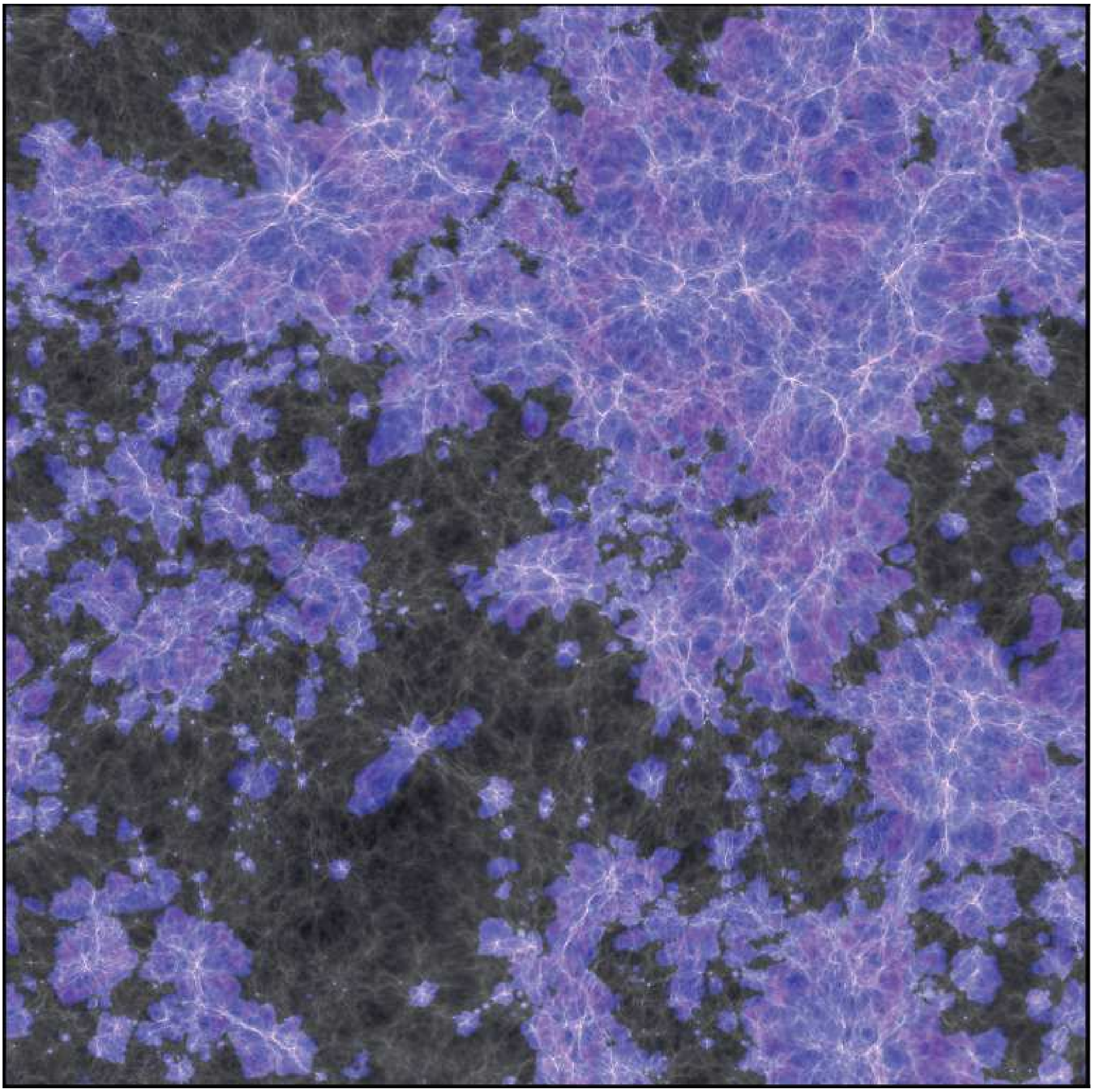


Figure 1.1: A combined map of ionisation (blue), temperature (red), gas density (luminosity), photon density (alpha channel) at z=7.0 in CoDa II. The map spans the entire simulation box in a plane of x,y : so a side is $64\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$ or $94.44\,\mathrm{cMpc}$ in length.

$\log_{10}$(ionising photon density), $\log_{10}$(gas neutral fraction), $\log_{10}$(gas over-density[4]), $\log_{10}$(gas temperature), $\log_{10}$(stellar mass), $\log_{10}$(young (<10 Myr, that emit LyC photons in CoDa II) stellar mass), M_{AB1600}. Also included is a table halo properties that were computed for Lewis et al. [2020]. The central circles represent r_{200}[5]. The central cross represents the halo centre as detected by pFOF [Roy et al., 2014].

One of the striking features of these maps is how the densest gas regions at the centres of massive star forming galaxies are both neutral (x_{HI}>0.1) and dense in young stars that have high ionising emissivities (high young stellar mass). The fact that these regions are able to recombine makes them more opaque to LyC photons (at constant gas density, UV transmission in cells decreases exponentially with increasing neutral fraction), which can be directly seen in the photon density maps.

Recombination is not as effective in the less dense regions that surround these neutral peaks of star formation. In the most massive haloes, this gas is heated past 10^5 K by accretion shocks on the edges of accreting filamentary structures of gas, and to 10^6 K and beyond by supernova shock waves.

These few maps illustrate how well CoDa II resolves detailed features and sub-structure in the most massive haloes. At the same time, they illustrate how very different the environments of massive haloes can be : in some cases, the central halo is surrounded by a swarm of less massive stellar densities, whereas in others it is isolated. However, also apparent in some maps is that a certain amount of star formation occurs far from the dark matter haloes in the maps' centres. Ocvirk et al. [2020] show this to be due to our spherical definition of dark matter haloes, despite their often triaxial nature, and the tendency of FoF to "over-link" haloes, i.e. in some cases, if a dense enough "bridge" of DM exists between two haloes, they are detected as one structure. When that happens, the centre of mass computed by FoF does not correspond to a physical over-density. We will discuss this issue in further detail in Sec. Ionising photon budget of galaxies.

---

[4] Here we define gas over-density as the cell density in units of the average baryon density.

[5] Defined in Sec. Ionising photon budget of galaxies.

1.1.5 Ocvirk et al. 2020

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **496,** 4087–4107 (2020)
Advance Access publication 2020 May 13

doi:10.1093/mnras/staa1266

Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization

Pierre Ocvirk,[1]★ Dominique Aubert,[1] Jenny G. Sorce,[1,2,3] Paul R. Shapiro,[4] Nicolas Deparis,[1] Taha Dawoodbhoy,[4] Joseph Lewis,[1] Romain Teyssier,[5] Gustavo Yepes,[6,7] Stefan Gottlöber,[3] Kyungjin Ahn,[8] Ilian T. Iliev[9] and Yehuda Hoffman[10]

[1] *CNRS, Observatoire astronomique de Strasbourg, Université de Strasbourg, UMR 7550, F-67000 Strasbourg, France*
[2] *CNRS, Centre de Recherche Astrophysique de Lyon, Université Lyon 1, ENS de Lyon, UMR 5574, F-69230 Saint-Genis-Laval, France*
[3] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[4] *Department of Astronomy, University Texas, Austin, TX 78712-1083, USA*
[5] *Institute for Theoretical Physics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[6] *Departamento de Física Teórica M-8, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[7] *Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[8] *Chosun University, 375 Seosuk-dong, Dong-gu, Gwangjiu 501-759, Korea*
[9] *Astronomy Center, Department of Physics & Astronomy, University of Sussex, Pevensey II Building, Falmer, Brighton BN1 9QH, UK*
[10] *Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel*

Accepted 2020 May 1. Received 2020 April 3; in original form 2019 October 30

ABSTRACT

Cosmic Dawn II (CoDa II) is a new, fully coupled radiation-hydrodynamics simulation of cosmic reionization and galaxy formation and their mutual impact, to redshift $z < 6$. With 4096^3 particles and cells in a 94 Mpc box, it is large enough to model global reionization and its feedback on galaxy formation while resolving all haloes above $10^8\,\mathrm{M}_\odot$. Using the same hybrid CPU–GPU code RAMSES–CUDATON as CoDa I in Ocvirk et al. (2016), CoDa II modified and re-calibrated the subgrid star formation algorithm, making reionization end earlier, at $z \gtrsim 6$, thereby better matching the observations of intergalactic Lyman α opacity from quasar spectra and electron-scattering optical depth from cosmic microwave background fluctuations. CoDa II predicts a UV continuum luminosity function in good agreement with observations of high-z galaxies, especially at $z = 6$. As in CoDa I, reionization feedback suppresses star formation in haloes below $\sim 2 \times 10^9\,\mathrm{M}_\odot$, though suppression here is less severe, a possible consequence of modifying the star formation algorithm. Suppression is environment dependent, occurring earlier (later) in overdense (underdense) regions, in response to their local reionization times. Using a constrained realization of lambda cold dark matter constructed from galaxy survey data to reproduce the large-scale structure and major objects of the present-day Local Universe, CoDa II serves to model both *global* and *local* reionization. In CoDa II, the Milky Way and M31 appear as individual islands of reionization, i.e. they were not reionized by the progenitor of the Virgo cluster, or by nearby groups, or by each other.

Key words: radiative transfer – methods: numerical – intergalactic medium – Local Group – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

The first billion years of the Universe is a key period for the evolution of the intergalactic medium (hereafter, 'IGM'), and the formation of galaxies. As the first stars form, their UV photons propagate, carving ionized regions which ultimately percolate, marking the end of the epoch of reionization (hereafter, 'EoR'). The EoR is considered as the next observational frontier, both for 21-cm radio telescopes such as LOFAR, MWA, SKA, NenuFAR, and infrared or near-infrared missions such as JWST. The recent tentative detection of a 21 cm from the dark ages by EDGES (Bowman et al. 2018) has stirred the community and prompted numerous teams to try to confirm the signal with other instruments.

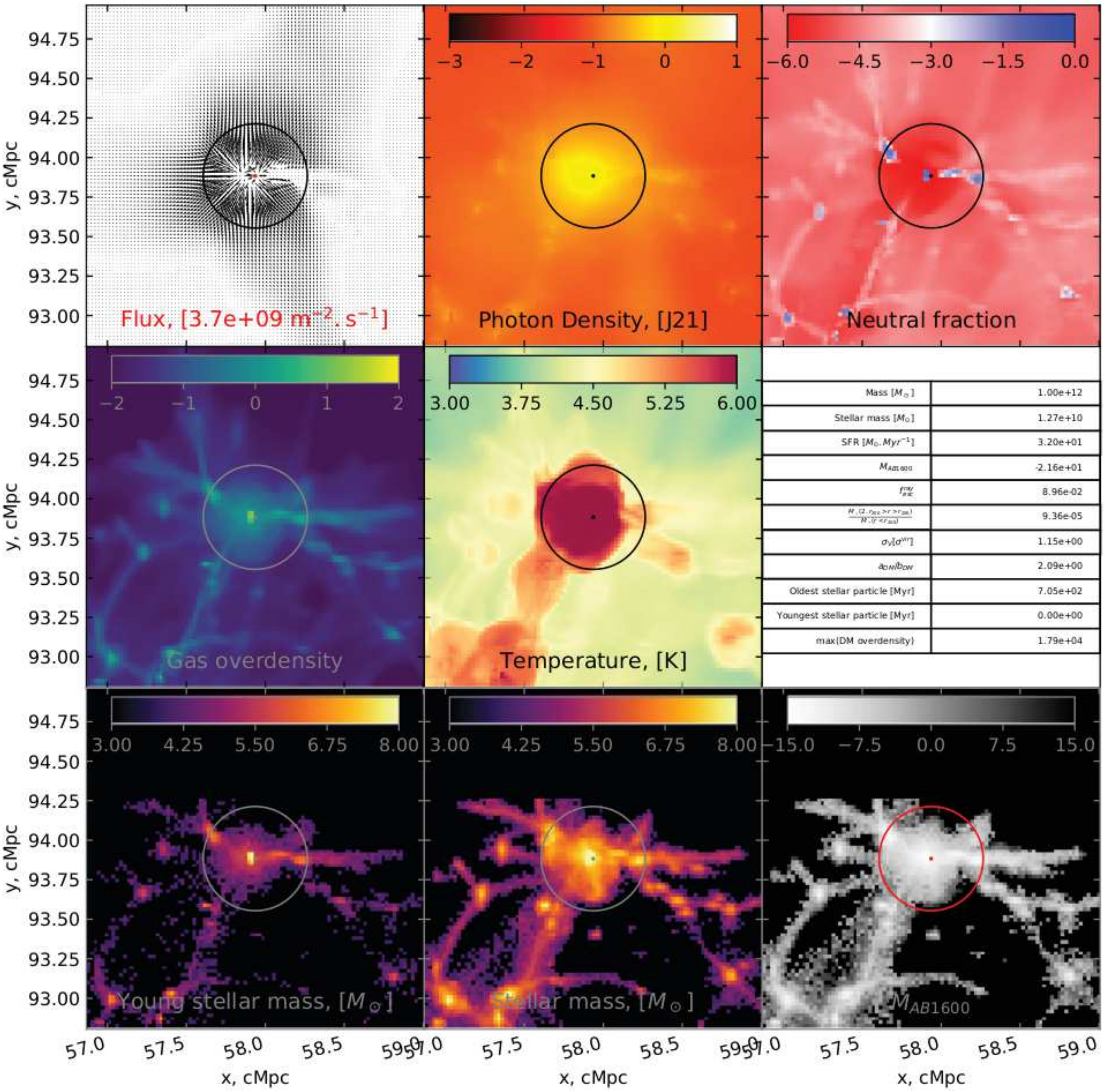
Meanwhile, a new breed of numerical simulations of galaxy formation has appeared, which aims at accounting for the rich physics, in particular the impact of the ionizing radiation, necessary to describe the EoR.[1] Faithful modelling of reionization, though, requires us to be able to describe simultaneously the coupled multiscale problem of global reionization and individual galaxy formation, with gravity, hydrodynamics and radiative transfer. To represent the complex geometry of reionization in a statistically

---

[1] See Dayal & Ferrara (2018) for a review of recent advances in numerical simulations of galaxy formation and reionization.

★ E-mail: pierre.ocvirk@astro.unistra.fr

Published by Oxford University Press on behalf of the Royal Astronomical Society



Mass [$M_\odot$]	1.00e+12
Stellar mass [$M_\odot$]	1.27e+10
SFR [$M_\odot.Myr^{-1}$]	3.20e+01
M_{AB1600}	-2.16e+01
f^{rav}_{esc}	8.96e-02
$\frac{M_*(2.r_{200}>r>r_{200})}{M_*(r<r_{200})}$	9.36e-05
$\sigma_V[\sigma^{vir}]$	1.15e+00
a_{DM}/b_{DM}	2.09e+00
Oldest stellar particle [Myr]	7.05e+02
Youngest stellar particle [Myr]	0.00e+00
max(DM overdensity)	1.79e+04

Figure 1.2: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\,M_\odot$).

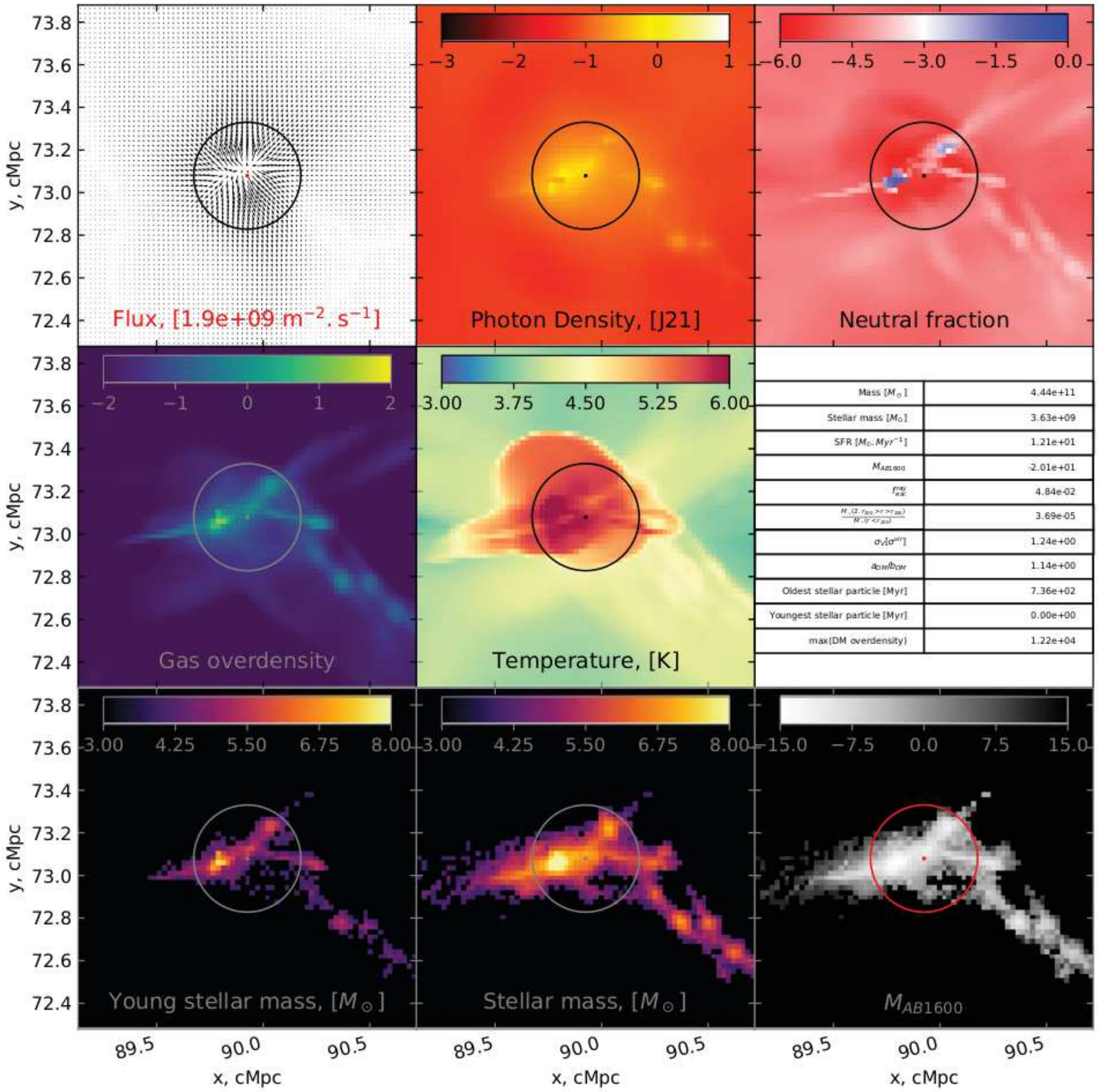
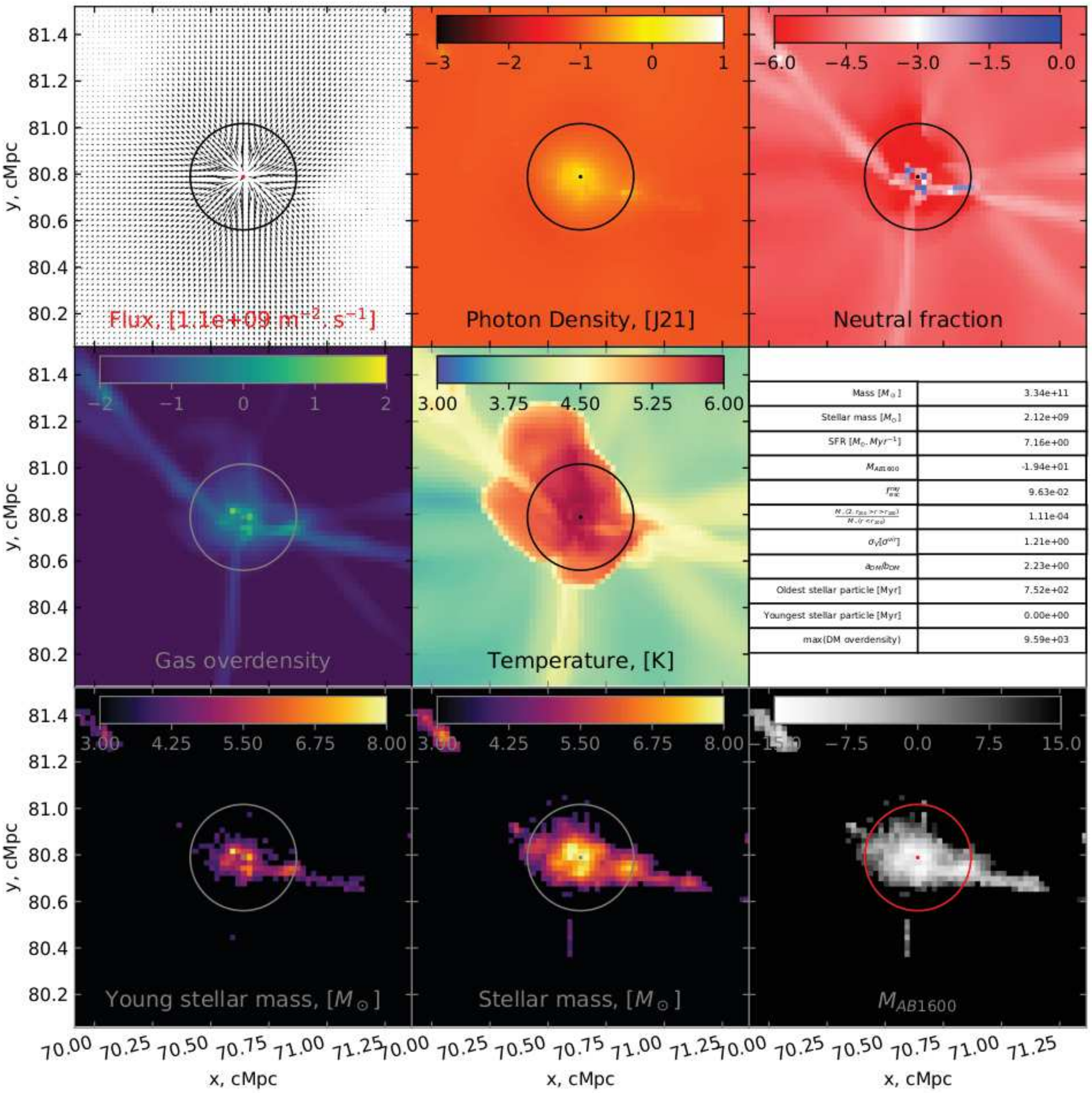
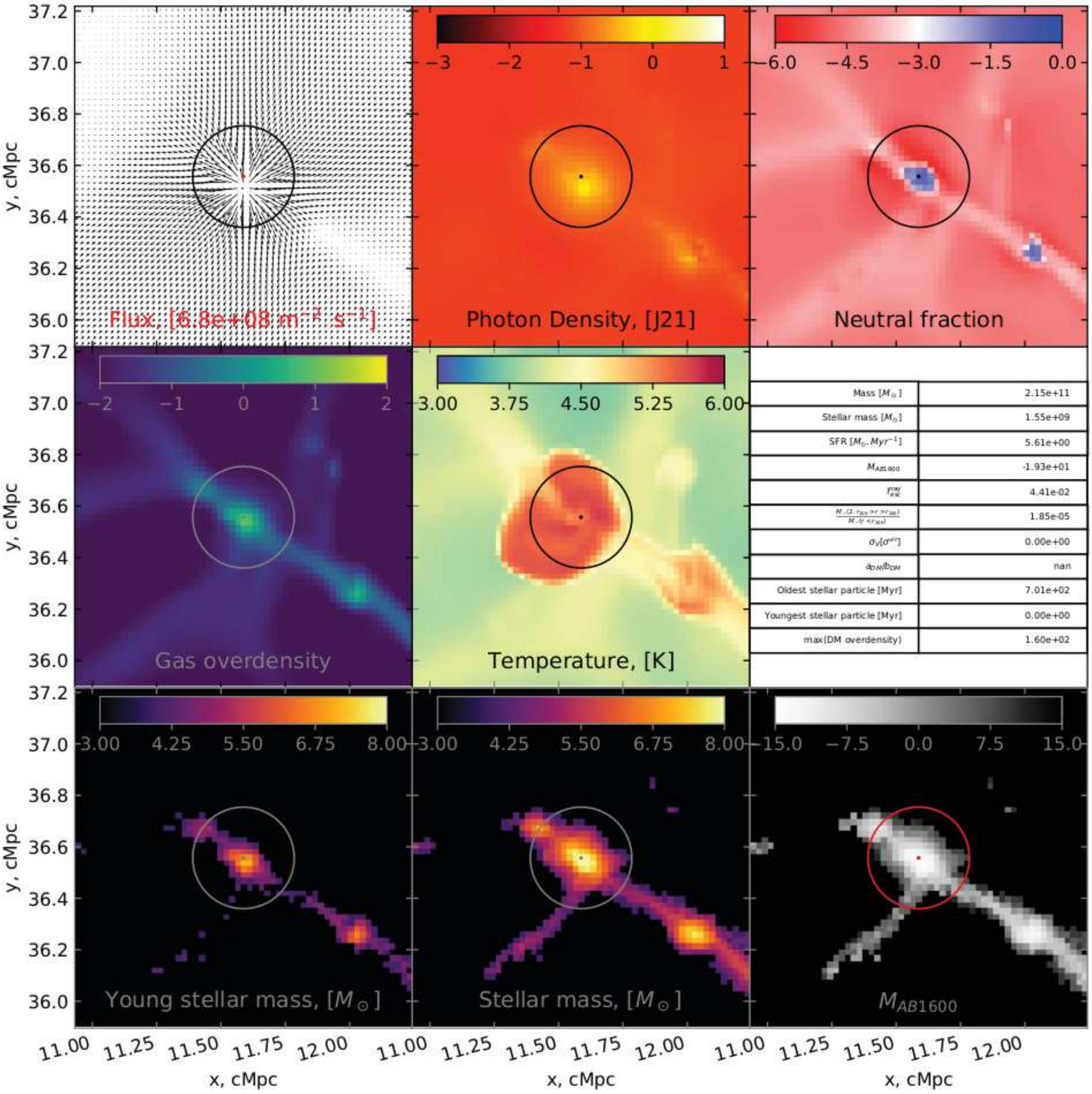


Figure 1.3: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\,M_\odot$).



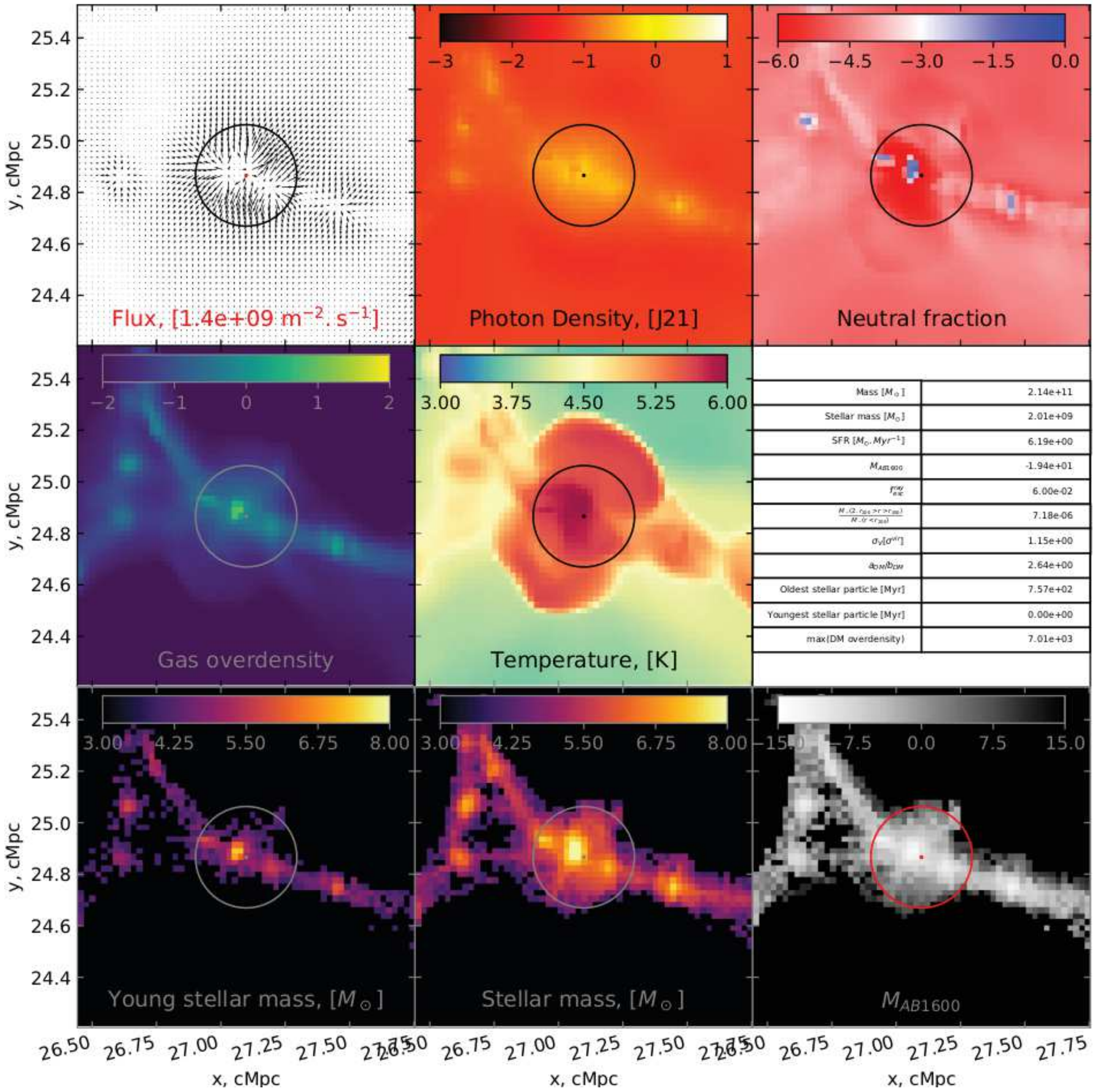
Mass [$M_\odot$]	3.34e+11
Stellar mass [$M_\odot$]	2.12e+09
SFR [$M_\odot.Myr^{-1}$]	7.16e+00
M_{AB1600}	-1.94e+01
f_{esc}^{ray}	9.63e-02
$\frac{M_*(2.r_{200}>r>r_{200})}{M_*(r<r_{200})}$	1.11e-04
$\sigma_V[\sigma^{vir}]$	1.21e+00
a_{DM}/b_{DM}	2.23e+00
Oldest stellar particle [Myr]	7.52e+02
Youngest stellar particle [Myr]	0.00e+00
max(DM overdensity)	9.59e+03

Figure 1.4: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\, M_\odot$).



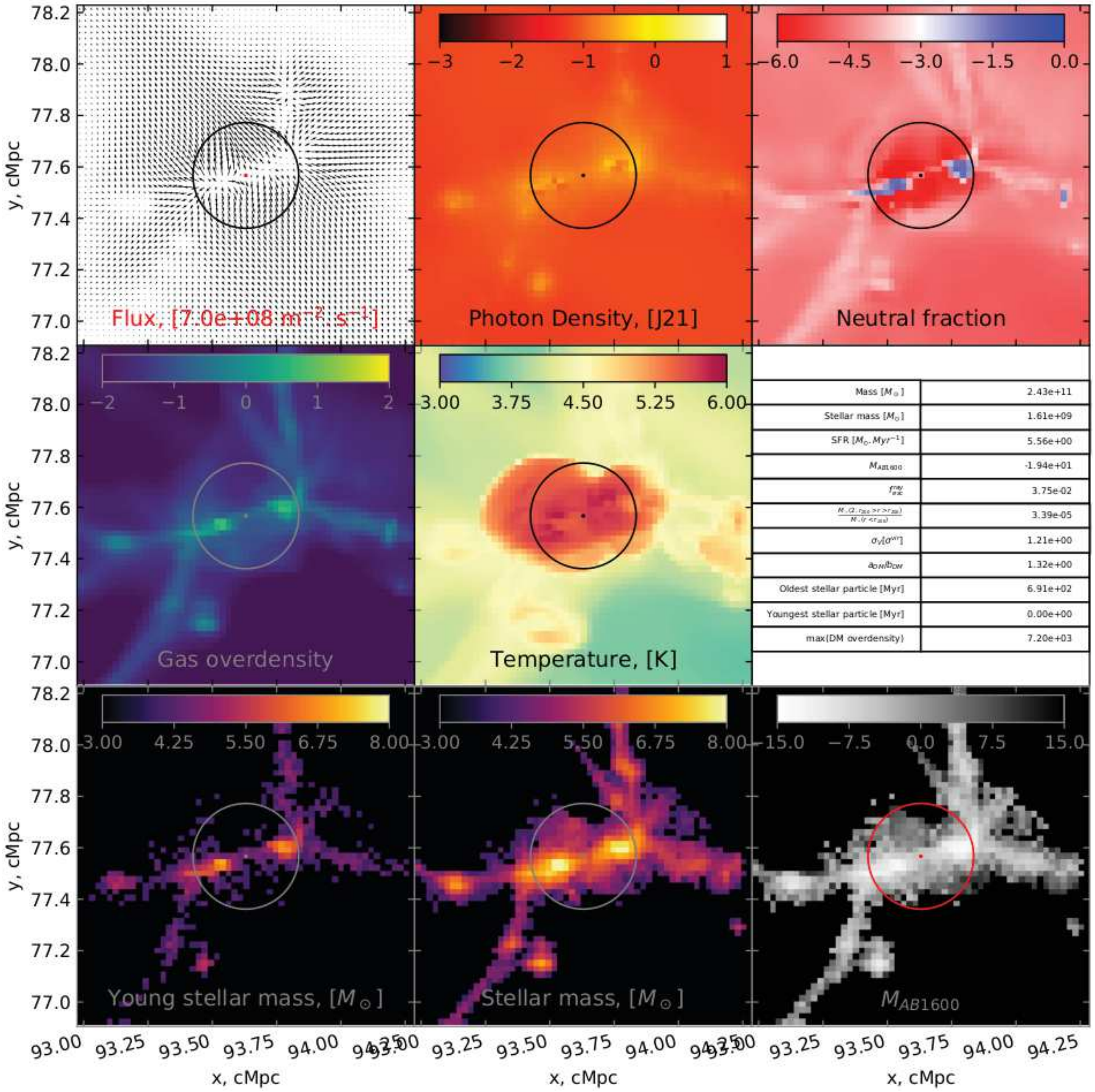
Mass [$M_\odot$]	2.15e+11
Stellar mass [$M_\odot$]	1.55e+09
SFR [$M_\odot.Myr^{-1}$]	5.61e+00
M_{AB1600}	-1.93e+01
f_{esc}^{ray}	4.41e-02
$\frac{M_*(2.r_{200}>r>r_{200})}{M_*(r<r_{200})}$	1.85e-05
$\sigma_V[\sigma^{vir}]$	0.00e+00
a_{DM}/b_{DM}	nan
Oldest stellar particle [Myr]	7.01e+02
Youngest stellar particle [Myr]	0.00e+00
max(DM overdensity)	1.60e+02

Figure 1.5: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\,M_\odot$).



Mass [$M_\odot$]	2.14e+11
Stellar mass [$M_\odot$]	2.01e+09
SFR [$M_\odot.Myr^{-1}$]	6.19e+00
M_{AB1600}	-1.94e+01
f_{esc}^{ray}	6.00e-02
$\frac{M_*(2.r_{200} > r > r_{200})}{M_*(r < r_{200})}$	7.18e-06
$\sigma_V[\sigma^{vir}]$	1.15e+00
a_{DM}/b_{DM}	2.64e+00
Oldest stellar particle [Myr]	7.57e+02
Youngest stellar particle [Myr]	0.00e+00
max(DM overdensity)	7.01e+03

Figure 1.6: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\,M_\odot$).



Mass [$M_\odot$]	2.43e+11
Stellar mass [$M_\odot$]	1.61e+09
SFR [$M_\odot.Myr^{-1}$]	5.56e+00
M_{AB1600}	-1.94e+01
f^{ray}_{esc}	3.75e-02
$\frac{M_*(2.r_{200}>r>r_{200})}{M_*(r<r_{200})}$	3.39e-05
$\sigma_V[\sigma^{vir}]$	1.21e+00
a_{DM}/b_{DM}	1.32e+00
Oldest stellar particle [Myr]	6.91e+02
Youngest stellar particle [Myr]	0.00e+00
max(DM overdensity)	7.20e+03

Figure 1.7: Sets of maps for two massive haloes ($M_{halo} \gtrsim 10^{11}\, M_\odot$).

meaningful way requires a comoving computational domain as large as ∼(100 Mpc)^3 e.g. Iliev et al. (2006, 2014). In order to follow the millions of galaxies in this volume over the full mass range of galactic haloes contributing to reionization,[2] while modelling the impact of reionization on these individual galaxies and the IGM, a physical resolution of a few kpc over the entire volume is needed.

The EoR is also important for its impact on low-mass galaxy formation. The baryonic gas content of low-mass galaxies is depressed by the hydrodynamical backreaction of their interstellar and surrounding intergalactic media, due to the photoheating that accompanies reionization by the rising ionizing UV background (Shapiro, Giroux & Babul 1994; Gnedin 2000; Hoeft et al. 2006). This, in turn, suppresses their star formation rates. The resulting 'fossil galaxies' may offer a credible solution to the missing satellite problem as formulated in the Local Group (see e.g. Bullock & Boylan-Kolchin 2017, and references, therein).

The faintest galaxies known, potentially susceptible to such radiative suppression of their gas and star formation, can only be detected nearby, in the Local Group. Therefore, to make direct comparisons possible between simulations and this observationally accessible sample, it is necessary to start from initial conditions preselected to produce the observed galaxies and large-scale structure of the Local Universe.

In the past few years, we have made significant progress in this direction, in our project of large-scale simulation of cosmic reionization and galaxy formation based upon fully coupled radiation hydrodynamics, the Cosmic Dawn ('CoDa') simulation project. The first of the CoDa simulations (hereafter, referred to as 'CoDa I'), described in Ocvirk et al. (2016) (hereafter, O16), with further results presented in Dawoodbhoy et al. (2018), was based upon the massively parallel, hybrid CPU–GPU code RAMSES–CUDATON and ran for 11 d on the Titan supercomputer at the Oak Ridge Leadership Computational Facility (OLCF), with 8092 GPUs and 8092 CPUs, utilizing 4096^3 particles and cells in a volume 94 comoving Mpc on a side. The second simulation, CoDa I-AMR, described in Aubert et al. (2018), was based upon the massively parallel, hybrid CPU–GPU code EMMA, with adaptive mesh refinement (AMR), which also ran on Titan, with 4096 GPUs and 32768 CPUs, utilizing 2048^3 particles, starting from a uniform grid of 2048^3 unrefined cells from which AMR increased the resolution locally, by up to a factor of 8, depending on the local overdensity, leading ultimately to 18 billion cells after refinement, in the same size box, starting from the same initial conditions (except coarsened to the initial, unrefined grid of CoDa I-AMR).

These CoDa simulations share four principal characteristics that make them special and a first of their kind, as follows. Their resolution is high enough to track the formation of the mass range of galactic haloes above $10^8\,\mathrm{M}_\odot$ believed to dominate reionization and the backreaction of reionization on their evolution as sources. Their volume is large enough to model the mean history of reionization and the hydrogen ionizing UV background, along with the inhomogeneity of reionization in time and space (Fig. 1 illustrates these aspects). They are based upon a constrained realization of the Gaussian random initial conditions for lambda cold dark matter (ΛCDM), derived from galaxy observations of the Local Universe, which are designed to reproduce the large-scale structure and familiar features observed in our Local Universe today when evolved to the present. By starting from this 'constrained realization' of initial conditions which reproduces a number of selected, familiar features of the Local Universe in a volume centred on the Local Group, such as the Milky Way, M31, and the Fornax and Virgo clusters, we are able to use these simulations to model both *global* and *local* reionization, simultaneously, including the impact of reionization on the formation of galaxies in the Local Group. Finally, by fully coupling the hydrodynamics to the radiative transfer of ionizing radiation, we move beyond simulations which adopt a pre-computed, uniform UV background, by computing the rise of the UV background, inhomogeneous in space and time, and its location-specific impact on galaxies and their emissivity, together, self-consistently.

In this paper, we present the Cosmic Dawn II ('CoDa II') simulation, which aims to improve upon our previous two simulations in several aspects. In CoDa I (O16), the end of global reionization was somewhat later than observed, at $z \lesssim 5$, a consequence of adopting a slightly lower efficiency for the subgrid star formation algorithm than necessary to end reionization by $z = 6$, despite our effort to calibrate this efficiency by a series of small-box simulations intended to predict the outcome for the large box. For CoDa II, the star formation algorithm parameters were recalibrated and the algorithm was also modified, to ensure that reionization ended earlier than CoDa I, by $z = 6$, in better agreement with the observations. While CoDa I-AMR (Aubert et al. 2018) was also successfully recalibrated to end reionization earlier than did CoDa I, both CoDa I and CoDa I-AMR started from the same generation of 'constrained realization' initial conditions developed by the *Constrained Local UniversE Simulations* ('CLUES') project which have since been updated and improved, including for instance more recent observational data. Those earlier initial conditions resulted in a Virgo replica less massive than observed, and were derived for a background universe with cosmological parameters for a WMAP5 cosmology (Hinshaw et al. 2009). The new initial conditions for CoDa II now feature a more realistic Virgo replica, and use a more recent set of cosmological parameters determined by the *Planck* collaboration (Planck Collaboration XVI 2014).

The main goal of this first CoDa II paper is to introduce the simulation and compare the results with current observational constraints on the global EoR and the high-redshift galaxies that caused it, such as the evolution of the cosmic neutral and ionized fractions, the mean intensity of the cosmic ionizing UV background, the cosmic star formation rate density, the Thomson-scattering optical depth through the intergalactic medium measured by cosmic microwave background fluctuation measurements, and the high-redshift galactic UV luminosity functions. In Section 2, we describe the RAMSES-CUDATON code on which the CoDa II simulation is based, and the simulation set-up, including the new 'constrained realization' initial conditions used here. Our results are presented in Section 3, including the global history of reionization compared with observational constraints, the star formation rates of galaxies and the net star formation rate density of the universe, and the UV continuum luminosity function of high-z galaxies during the EoR. We use the simulation to show how the star formation rate in low-mass haloes is affected by reionization and the environmental dependence of this process. Finally, we investigate the reionization history of the Local Group, and then finish with a short summary in Section 4.

---

[2] The so-called 'atomic-cooling haloes' (henceforth, 'ACHs'), those with virial temperatures above $\sim 10^4$ K, corresponding to halo masses above $\sim 10^8$ $\mathrm{M}_\odot$.

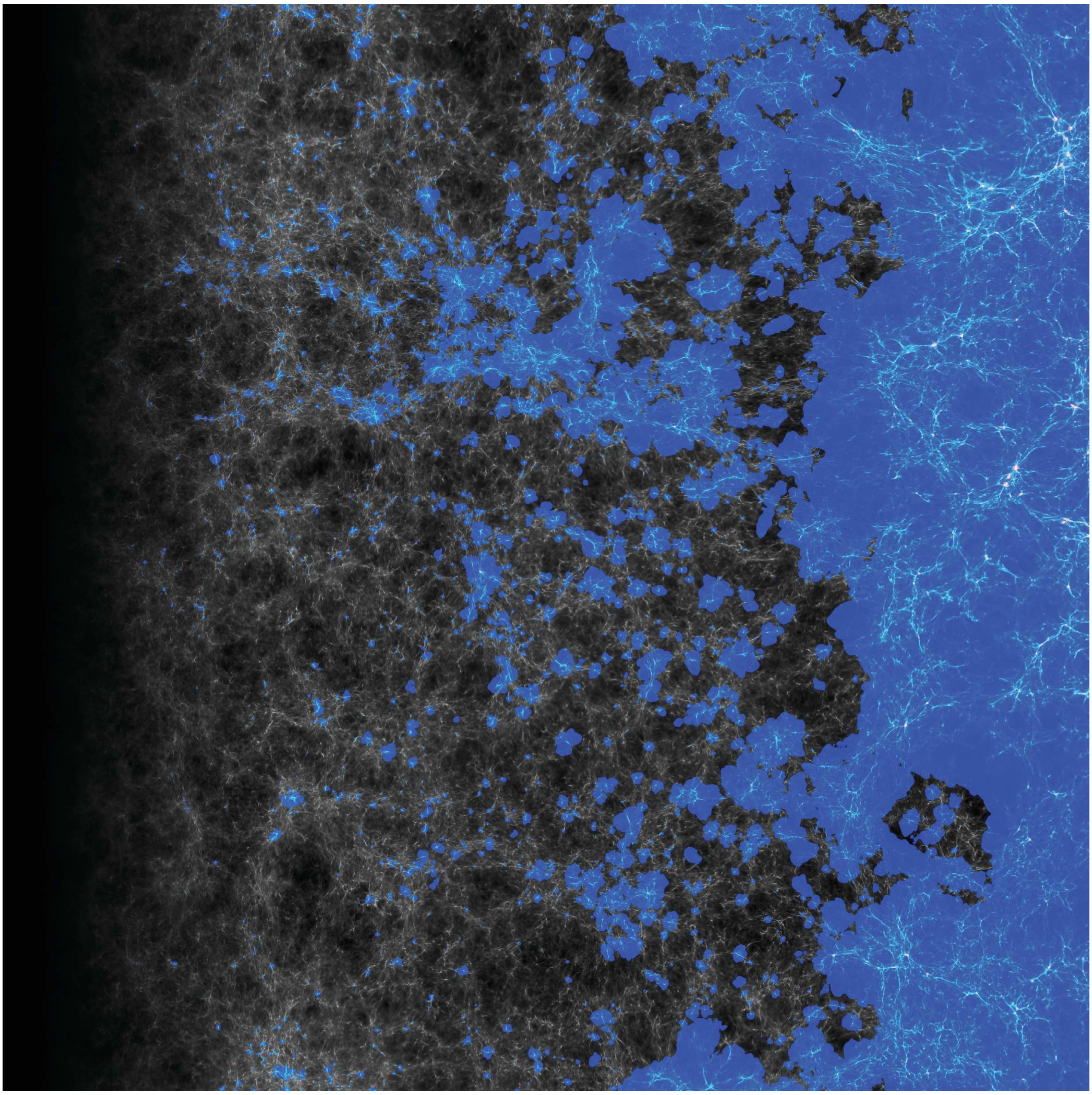


Figure 1. Illustration of the time history of reionization in the CoDa II simulation: a pseudo-light-cone image. A single, planar (x, y)-slice through the full $(94\,\mathrm{cMpc})^3$ simulation cube was chosen for display, and a sequence of $\sim$1024 closely spaced time-slices of the 4096^2 grid cells in that plane was outputted, to sample the time-evolution from redshift $z = 150$ to $z = 5.8$. The image is made by the concatenation of vertical stripes, 4-cells wide, each stripe selected from another time-slice in the sequence, to show the time evolution as time increases to the right along the x-axis, from redshift $z = 150$ (left) to $z = 5.8$ (right). Blue regions are photoheated, while small, bright red regions (which can be seen more clearly by zooming the high-resolution figure) correspond to regions heated by supernovae feedback and accretion shocks. The green colour, on the other hand, denotes regions where ionization is ongoing and incomplete, and temperature has not yet risen to the $\sim 10^4$ K typical of fully ionized regions. Brightness indicates the gas density contrast.

2 METHODOLOGY

The Cosmic Dawn simulation uses the fully coupled radiation hydrodynamics code RAMSES–CUDATON described in O16. This section outlines the basics of the code and the differences with CoDa I, as well as its deployment. The reader is referred to O16 for further detail. For quick reference, the parameters of the simulation are summarized in Table 1.

Table 1. Cosmic Dawn II simulation parameters summary.

Cosmology (Planck14)	
Dark energy density Ω_Λ	0.693
Matter density $\Omega_{\rm m}$	0.307
Baryonic matter density $\Omega_{\rm b}$	0.048
Hubble constant $h = {\rm H}_0/(100\,{\rm km\,s^{-1}})$	0.677
Power spectrum	
Normalization σ_8	0.8288
Index n	0.963
Setup	
Number of nodes (GPUs, cores used)	16 384 (16 384, 65 536)
Grid size	4096^3
Comoving box size $\rm L_{box}$	94.44 Mpc (64 h^{-1}Mpc)
Grid cells per node	$128 \times 128 \times 64$
Comoving force resolution dx	23.06 kpc
Physical force resolution at $z = 6$	3.3 kpc
DM particle number $N_{\rm DM}$	4096^3
DM particle mass $M_{\rm DM}$	4.07×10^5 M$_\odot$
Average cell gas mass	0.75×10^5 M$_\odot$
Initial redshift $z_{\rm start}$	150
End redshift $z_{\rm end}$	5.8
Star formation	
Density threshold $\delta_\star$	50 $\langle\rho_{\rm gas}\rangle$
Efficiency $\epsilon_\star$	0.02
Stellar particle birth mass $M_\star$	11732 M$_\odot$
Feedback	
Massive star lifetime $t_\star$	10 Myr
Supernova	
Mass fraction $\eta_{\rm SN}$	10 per cent
Energy $E_{\rm SN}$	10^{51} erg
Radiation	
Stellar ionizing emissivity	4.32×10^{46} ph s^{-1} M$_\odot$$^{-1}$
Stellar particle escape fraction $\rm f_{esc,\star}$	0.42
Effective photon energy	20.28 eV
Effective H I cross-section σ_E	2.493 x 10^{-22} m^2
Speed of light c	299 792 458 m s^{-1}

2.1 RAMSES–CUDATON

Since, the code is already described in detail in O16, we here only give a brief summary of it and highlight the differences in implementation / parameter choices between CoDa I and CoDa II. RAMSES–CUDATON results from the coupling of RAMSES (Teyssier 2002) and ATON (Aubert & Teyssier 2008). RAMSES handles gravity, hydrodynamics, and star formation and its mechanical feedback, while ATON handles photon propagation, hydrogen-photon thermochemistry, and computes the cooling terms. Finally, CUDATON results from the porting of ATON to CUDA (Aubert & Teyssier 2010) to enable it to take advantage of hybrid architectures featuring GPGPU (general purpose graphics processing unit). Thanks to this unique feature, we were able to perform CoDa II using the full speed of light, therefore circumventing potential artefacts arising with reduced speed of light approaches, highlighted by Deparis et al. (2019) and Ocvirk et al. (2019), or arising with infinite speed of light approaches in many ray-tracing codes. However, with this optimization, we are only able to use RAMSES in unigrid mode, i.e. without adaptive mesh refinement (hereafter, 'AMR'). This is one of the shortcomings that the EMMA code (Aubert, Deparis & Ocvirk 2015) and the RAMSES-RT code (Rosdahl et al. 2013) attempt to address.

2.2 Differences with CoDa I

2.2.1 *Star formation and calibration*

As in O16, we consider star formation using a phenomenological approach. In each cell with gas density larger than a gas overdensity $\delta_\star = 50$, we spawn new star particles at a rate given by

$$\dot{\rho}_\star = \epsilon_\star \frac{\rho_{\rm gas}}{t_{\rm ff}}, \tag{1}$$

where $t_{\rm ff} = \sqrt{\frac{3\pi}{32G\rho}}$ is the free-fall time of the gaseous component and $\epsilon_\star = 0.02$ is the star formation efficiency, instead of 0.01 in O16. The higher star formation efficiency of CoDa II is intended to achieve the end of reionization by redshift 6, whereas CoDa I finished reionization by $z = 4.6$. In practice, we calibrate this subgrid star formation efficiency by running a suite of reionization simulations in much smaller boxes than our final production run, but at the same grid and particle resolution as the latter, and then test the best-choice efficiency parameter by another suite of simulations of increasing box size. We adjusted the value of $\epsilon_\star$ so as to achieve a good agreement with the evolution of the cosmic star formation rate density inferred from observations of high-redshift galaxies (discussed in Section 3.3).

Unlike the first Cosmic Dawn simulation reported in O16, we do not require the cell temperature to be lower than $T_\star = 2 \times 10^4$ K in order to form stars: all cells, no matter their temperature, are eligible to forming stars above the density threshold. This is to account for the fact that, at higher resolution, such ionized cells may still host cold, neutral regions, which may still form stars. We will see in Section 3 that this choice has an impact on the strength of the radiative suppression of star formation.

The star particle mass at birth depends on the cell gas density, but is always a multiple of a fixed elementary mass $M_\star^{\rm birth}$, chosen to be a small fraction of the baryonic mass resolution. In this framework, with the box size and resolution of CoDa (see Section 2.3), we have $M_\star^{\rm birth} = 11732\,{\rm M}_\odot$. This mass is small enough to sample adequately the star formation history of even low-mass galaxies, and still large enough to mitigate stochastic variations in the number of massive stars per star particle. It is larger than the elementary mass used in CoDa I, which had $M_\star^{\rm birth} = 3194\,{\rm M}_\odot$. This change reduces the number of star particles in the simulation, which in turn reduces their computational cost, in memory and processing.

2.2.2 *Ionizing source model*

The radiative transfer considered here is of H ionizing radiation released by stellar sources associated with the star particles described in the previous section, subject to absorption by H I bound-free opacity. As with our subgrid star formation algorithm, the rate of release of ionizing photons by star particles into the grid cells in which they are located must be parametrized, so as to represent the absorption taking place at the subpc scales of the interstellar molecular clouds where stars are formed inside each galaxy.

Hence, we must add to our subgrid star formation algorithm described above a parametrized ionizing photon efficiency (henceforth, 'IPE'), the number of ionizing photons released per unit stellar baryon per unit time, into the host grid cell of each star particle. This IPE is based on an assumed stellar IMF and, since star particles are assumed to form embedded in high-density molecular clouds which are unresolved by our grid cells, we also adopt a subgrid stellar-birthplace escape fraction. We define this IPE, therefore, as $\xi_{\rm IPE} \equiv f_{{\rm esc},\star}\xi_{\rm ph,IMF}$, where $f_{{\rm esc},\star}$ is the stellar-birthplace escape fraction and $\xi_{\rm ph,IMF}$ is the number of ionizing photons emitted per Myr per stellar baryon. The subgrid stellar-birthplace escape fraction $f_{\rm esc,\star}$ should not be confused with the *galactic* escape fraction, which is not a subgrid quantity put in 'by hand' but is rather the self-consistent outcome of our radiative transfer module CUDATON between grid cells, for all the cells associated with a given galactic halo. There is no need for us to adopt a galactic escape fraction, therefore. Our RT calculation finds the actual time-varying escape fraction for each galactic halo, which differs from one galaxy to the next.

Each stellar particle is considered to radiate for one massive star lifetime $t_\star = 10$ Myr, after which the massive stars die (triggering a supernova explosion) and the particle becomes dark in the H-ionizing UV. We adopted an emissivity $\xi_{\rm ph,IMF} = 1140$ ionizing photons/Myr per stellar baryon. This is similar to the ionizing emissivity of a BPASS model (Eldridge et al. 2017) for a $Z = 0.001$ binary population with Kroupa initial mass function (Kroupa 2001), as presented in Rosdahl et al. (2018), averaged over the first 10 Myr. Although we do not follow chemical enrichment in CoDa II, Pawlik et al. (2017) showed that the metallicity in gas at the average density of the ISM is indeed close to $Z = 0.001$ during the epoch of reionization, which validates the choice of this metallicity. We used a mono-frequency treatment of the radiation with an effective frequency of 20.28 eV, as in Baek et al. (2009).

Finally, we calibrated the stellar-particle escape fraction $f_{\rm esc,\star}$ by adjusting the value in our set of smaller box simulations as mentioned above, so as to obtain a reionization redshift close to $z = 6$, which led us to adopt a value of $\mathrm{f}_{\rm esc,\star} = 0.42$, close to the value $\mathrm{f}_{\rm esc,\star} = 0.5$ used in O16. We note that the *net* escape fraction from the cell in which the star particle is located is less than or equal to this subgrid value adopted for the unresolved stellar-birthplace, since there is additional absorption of ionizing photons inside the cell, by the bound-free opacity of the simulated gas in that cell, which depends upon that cell's neutral H fraction and gas density, as tracked self-consistently by RAMSES–UDATON. The overall *galactic* escape fraction will be even smaller, since galactic haloes typically involve multiple cells, and one cell can absorb the photons released in another. Our radiative transfer module CUDATON explicitly accounts for this.

As in Coda I, we neglect here the possible release of hard, ionizing UV radiation associated with a phase in the evolution of our simulated galaxies in which they host active galactic nuclei (hereafter, AGNs). Some AGN sources may arise during the epoch of reionization, e.g. in rare, massive protoclusters (Dubois et al. 2012). However, they are very rare and have long been thought to be minor contributors to the cosmic budget of ionizing photons responsible for the reionization of hydrogen (e.g. Shapiro & Giroux 1987; Giroux & Shapiro 1996; Haardt & Madau 2012; Haardt & Salvaterra 2015), although they could be important for explaining the line-of-sight variations of the properties of the Ly α forest just after reionization (Chardin et al. 2015). The question of the possible AGN contribution to reionization has recently been revisited following claims by Giallongo et al. (2015) that the observed AGN luminosity function at high redshift may have been underestimated (Madau & Haardt 2015). Further studies indicate that the AGN contribution to reionization must be subdominant or else would violate other observational constraints (e.g. Worseck et al. 2016; D'Aloisio et al. 2017; Oñorbe et al. 2017; Qin et al. 2017; Mitra, Choudhury & Ferrara 2018), while subsequent updates of the observed quasar luminosity function also conclude that the AGN contribution to reionization is very small (e.g. Parsa, Dunlop & McLure 2018; Kulkarni, Worseck & Hennawi 2019).

2.3 Simulation setup

2.3.1 Initial conditions

Our initial conditions are a constrained realization of the ΛCDM universe, intended to reproduce the observed features of our Local Universe in a box centred on the Milky Way if evolved to the present, while at the same time serving as a representative sample of the universe-at-large with which to model global reionization. We generated specifically the initial conditions for this CoDa II simulation as part of the CLUES (*Constrained Local UniversE Simulations*) project. These initial conditions are constrained by observational data on the positions and velocities of galaxies in the Local Universe described in Tully et al. (2013) (referred to as *Cosmicflows-2*) to result in a simulation that resembles the Local Universe within the framework of *Planck* cosmology ($\Omega_{\rm m} = 0.307$, $\Omega_\Lambda = 0.693$, $H_0 = 67.77$ km s^{-1} Mpc^{-1}, $\sigma_8 = 0.829$, Planck Collaboration XVI 2014), in a comoving box 94 (i.e. 64 h^{-1}) cMpc on a side. It contains 4096^3 dark matter particles on a cubic-lattice with the same number of cells, which are also the grid cells for the gas and radiation. The mass of each dark matter particle is then 4.07×10^5 M$_\odot$. The CoDa II initial conditions have the average universal density for this chosen cosmology.

Sorce et al. (2016a) describe in detail the steps of the method used to build these constrained initial conditions. The main steps are summarized as follows:

(i) We started with a set of data points, the radial peculiar velocities of observed galaxies, i.e. radial peculiar velocities v_r at discrete positions r at $z = 0$. Galaxies in the radial peculiar velocity catalogue were then grouped (e.g. Tully 2015a, b) to produce a data set that traces the coherent large-scale velocity field, with non-linear virial motions removed that would affect the reconstruction obtained with the linear method as shown by Sorce, Hoffman & Gottlöber (2017) and Sorce & Tempel (2017).

(ii) Biases inherent to any observational radial peculiar velocity catalogue were minimized (Sorce 2015).

(iii) The three-dimensional peculiar velocity field and associated cosmic displacement field were then reconstructed by applying the Wiener-filter (henceforth, WF) technique – a linear minimum variance estimator (Zaroubi et al. 1995; Zaroubi, Hoffman & Dekel 1999) – to the grouped and bias-minimized radial peculiar velocity constraints.

(iv) To account for the displacement of mass elements located at the positions of the galaxies in the catalogue at $z = 0$, away from the positions of their Lagrangian patch in the early universe, caused by the growth of structure over time, the positions of the constraints were relocated to the positions of their progenitors using the Reverse Zel'dovich Approximation (Doumler et al. 2013). Additionally, noisy radial peculiar velocities are replaced by their WF 3D reconstructions (Sorce et al. 2014).

(v) Linear density and velocity fields were then produced, constrained by the modified observational peculiar velocities combined with a random realization of ΛCDM to restore statistically the missing structures using the constrained realization technique (Hoffman & Ribak 1991, 1992).

(vi) These density and velocity fields, extrapolated to $z = 0$ according to linear theory, have to be rescaled to an actual starting redshift ($z_i = 150$ for the CoDa II simulation) to build its constrained initial conditions, and the resolution was increased (MUSIC; Hahn & Abel 2011). Increasing the resolution implies adding higher frequency modes to the initial conditions, which introduce random small-scale features. For this purpose, we created 200 random realizations of the initial conditions with these higher frequency modes added, and tried them out by dark matter-only N-body simulations [using the GADGET-2 code (Springel 2005)] evolved to $z = 0$, in order to select that with the best match to the observed features of the Local Universe at $z = 0$, including the Local Group[3] (according to its mass, the separation between the Milky Way and M31, their mass ratio, and very nearby environment with M33 and Centaurus A counterparts). The production of other prominent features of the Local Universe, like the Virgo cluster (with close to the right mass at close to the right position), was also used as a selection criterion.

The CoDa II initial conditions were thus chosen so as to reproduce familiar features of the observed Local Universe within a volume 94 cMpc on a side, centred on the LG, as closely as possible. Towards this end, all 200 of the dark matter-only N-body simulations from these 200 constrained realizations (as described above) were examined at $z = 0$, and, amongst these, we selected a reduced set which best reproduced the Local Group, then further selected amongst these to meet our additional requirements, as follows:

(a) A pair of galaxies is located at the centre (by construction) with masses, M_{200}, between 5.5×10^{11} and $2 \times 10^{12}\,h^{-1}\,M_\odot$.
(b) There is no other halo more massive than $5.5 \times 10^{11}\,h^{-1}\,M_\odot$ within a sphere of radius $2.5\,h^{-1}$ Mpc.
(c) Their separation is smaller than $1.5\,h^{-1}$Mpc.
(d) Their mass ratio is smaller than 2.
(e) They are located between $10\,h^{-1}$ and $14\,h^{-1}$ Mpc away from the Virgo cluster replica.
(f) There are haloes that could stand for M33 and Centaurus A (by far the most restrictive criterion).

Left with a dozen simulations with halo pairs, we selected the pair (and, thus, the simulation) that satisfied both requirements, that of a small separation and a small mass ratio: at $z = 0$, the distance between the two haloes of the pair is $0.85\,h^{-1}$ Mpc and their mass ratio is 1.2 (with masses of 1.55×10^{12} and $1.3 \times 10^{12}\,h^{-1}\,M_\odot$).

These N-body simulations were at lower mass-resolution than our final CoDa II simulation but at a resolution high enough (at 512^3 particles, with particle mass $1.7\times10^8\,h^{-1}\,M_\odot$) to ensure that the resulting Local Group candidates would be stable to increasing the resolution still further when the same initial conditions were resimulated in our CoDa II production run.[4] The resolution of the best initial conditions thereby selected was then increased still further, to 4096^3 particles and cells, adding additional modes at higher frequency, still with MUSIC. In order to provide a companion dark matter-only N-body simulation to $z = 0$ from these constrained-realization initial conditions, we also coarsened the final, high-resolution CoDa II initial conditions from 4096^3 down to 2048^3, to run a GADGET-2 N-body simulation to $z = 0$, called CoDa II-DM2048.[5] The CoDa II-DM2048 simulation was then used to make a final comparison of the simulated LG and other features of the Local Universe by our CoDa II initial conditions with the observed Local Universe at $z = 0$. A comparison of the observed Local Universe with the outcome of simulating the CoDa II initial conditions to $z = 0$ by CoDa II-DM2048, is shown in Fig. 2.

The advantage of these new initial conditions with respect to those used for the first generation of CoDa simulations, CoDa I and CoDa I-AMR, is a large-scale structure that matches the local one down to the linear threshold ($\sim 3\,h^{-1}$ Mpc) on larger distances (the entirety of the box if not for the periodic boundary conditions) and with more accurate positions ($\sim$3–4 h^{-1} Mpc). In particular, this simulation contains a Virgo cluster at the proper position and with a mass in better agreement with recent observational mass estimates (Sorce et al. 2016b). While the Virgo replica had a mass of $7\times10^{13}\,h^{-1}\,M_\odot$ in the CoDa I simulation, its mass is now $2.2\times10^{14}\,h^{-1}\,M_\odot$, i.e. the cluster is 3 times more massive.[6] These properties will be important for the planned follow-up projects.

Finally, the baryonic initial conditions at the initial redshift of $z = 150$ were generated assuming a uniform temperature equal to that of the CMB at that time, with identical gas and dark matter velocity fields. The initial value of the H ionized fraction was taken to be the homogeneous, post-recombination-era freeze-out value at $z = 150$, as computed following standard recipes such as in RECFAST (Seager et al. 1999).

2.3.2 Code deployment

Like CoDa I (O16) and CoDa I-AMR (Aubert et al. 2018), CoDa II was performed on the massively hybrid CPU–GPU supercomputer Titan, at OLCF. The code was deployed on 16 384 Titan nodes and 16 384 GPUs (1 GPU per node), with each GPU coupled to 4 cores, for a total of 65 536 cores. Each node hosted 4 MPI processes that each managed a volume of $64 \times 128 \times 128$ cells.

2.4 Post-processing: friends-of-friends halo catalogues

We used the massively parallel friends-of-friends (hereafter, FoF) halo finder of Roy, Bouillot & Rasera (2014) with a standard linking length of $b = 0.2$ to detect dark matter haloes in the CoDa II simulation. As shown in O16, they are reliably detected down to $\sim 10^8\,M_\odot$.

2.5 Online data publication

We plan to make a subset of the data and higher level products publicly available through the VizieR data base at CDS Strasbourg.[7] The latter link will be active only after acceptation of the paper. The CoDa II catalogues will also be made available through the cosmosim data base hosted by Leibniz Institut für Astrophysik Potsdam.[8] They will contain halo properties such as dark matter halo mass, position, stellar mass, the age of the oldest and youngest stellar particles in the halo, the star formation rate, and the AB magnitude at 1600 Angstrom.

---

[3] The Local Group belongs to the non-linear regime and is thus not directly constrained but is induced by the local environment (Carlesi et al. 2016).

[4] The resolution of each of these N-body simulations must indeed be large enough for their candidate Local Group to be affected only barely when resimulated with even higher resolution, when yet higher frequency random modes are introduced to the initial conditions, thereby adding even smaller scale features.

[5] It corresponds to the simulation ESMDPL_2048 from the MultiDark project. It is fully available at www.cosmosim.org

[6] Higher masses in prefect agreement with observational estimates for Virgo have been obtained in realizations of $500\,h^{-1}$ Mpc, but with the fixed 4096^3 grid size used here, such a box size results in a spatial and mass resolution unsuitable for this study.

[7] http://vizier.u-strasbg.fr/viz-bin/VizieR?-source = VI/146

[8] https://www.cosmosim.org/cms/simulations/cosmic-dawn/

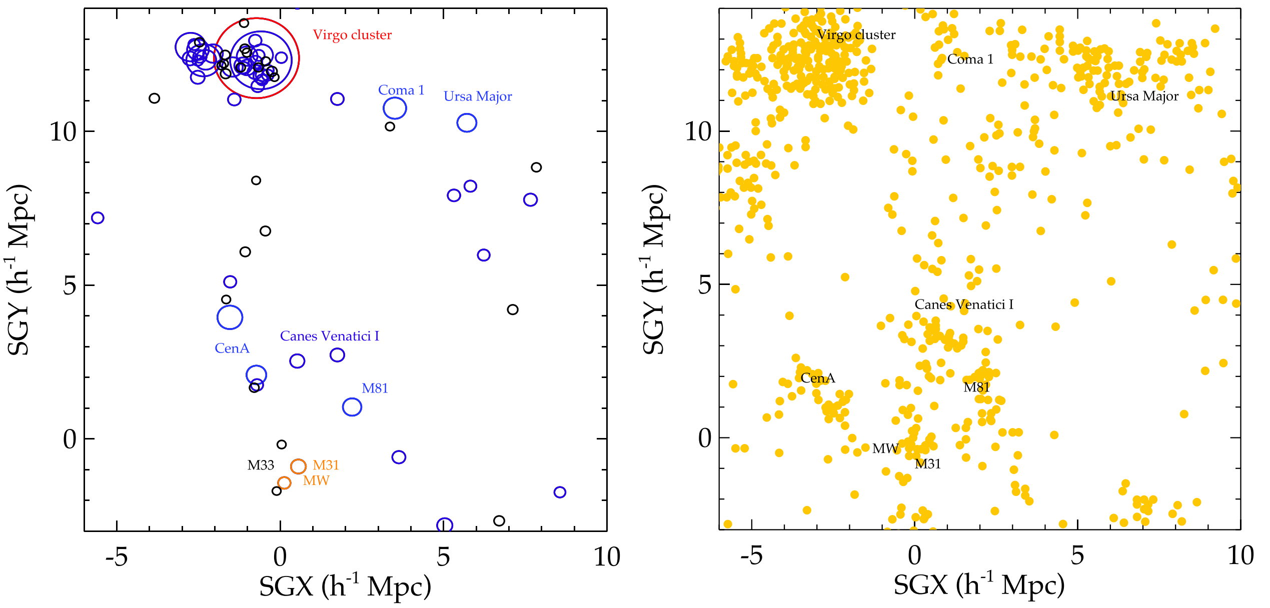


Figure 2. Comparison of the *N*-body haloes with masses $M > 2 \times 10^{11}$ M$_{\odot}$ in CoDa II-DM2048 at $z = 0$ in the Supergalactic Plane, in a subregion 16 h^{-1} Mpc on a side and 4 h^{-1} Mpc thick (left-hand panel) with the observed galaxies, groups, and clusters of the nearby Local Universe for the same region, a thin slice centred on the supergalactic equator, 4 h^{-1} Mpc thick (right panel), based upon the observed galaxies as plotted in fig. 4 of Courtois et al. (2013), taken from the Extragalactic Distance Database (http://edd.ifa.hawaii.edu). Circles of radius equal to the virial radii of the simulated haloes are shown on the left-hand panel. Features of the Local Universe and their analogues in the simulation are labelled. The correspondence between figure labels and Simbad (http://simbad.u-strasbg.fr/simbad/) names is given in Table 2.

Table 2. Correspondence table between names used in the article and Simbad names.

Name	Simbad name
Virgo cluster	Vir I
Coma I	Coma I group
Ursa Major	Ursa Major cluster
Canes Venatici I	CVn group
CenA	Cen A group
M81	M81 group

3 RESULTS

3.1 Global properties

In this section, we study the mean, global reionization history of the CoDa II simulation and compare it with available observational constraints. We demonstrate that the simulation is in agreement with most expectations from observations.

The most basic quantities to consider when gauging the success of a global EoR simulation are the evolution of the volume-weighted cosmic means of the neutral and ionized fractions, and the mean intensity of the ionizing radiation background. These are shown in Fig. 3, along with several observational constraints from Fan et al. (2006), Ouchi et al. (2010), and the *Planck* CMB Thomson-scattering optical depth τ (Planck Collaboration VI 2018).

Intergalactic H II regions grow in number and size over time until they overlap fully to finish reionizing the universe, before redshift 6. As shown in the plot of mean ionized hydrogen fraction versus redshift in Fig. 2, the ionized fraction was $x_{\rm H\,II} \sim 10^{-1}$, 0.5, 0.9, and 0.99 at z = 8.5, 7, 6.5, and 6.2, respectively, compatible with observationally inferred values by, e.g. Greig et al. (2017), Davies et al. (2018), Mason et al. (2018), and Hoag et al. (2019). At this point, the neutral fraction plot shows a characteristic, very steep decrease, down to $x_{\rm H\,I} \sim 10^{-4.6}$, after which the slope becomes slightly shallower. This transition marks the end of the EoR. According to this definition, reionization is complete in the simulation at redshift z = 6.1, in agreement with the brisk drop in neutral fraction between z = 7 and z = 6 indicated by the combined measurements of Ouchi et al. (2010) and Fan et al. (2006). The global timing of CoDa II is therefore in better agreement with these observations than was CoDa I. The latter ended reionization somewhat late, below redshift 5, although we found we could relabel the simulation redshifts of CoDa I by uniformly rescaling them by multiplying the simulation redshift by a factor of 1.3, in order to compare those results with observations or make observational predictions, with good agreement with these same observational constraints after this rescaling. Thanks to the new calibration and choice of parameters, our CoDa II simulation ended reionization earlier, by redshift 6, just as required by the observations, so we do not need to perform any such rescaling; we can compare its results here directly with observations and make direct predictions.

The post-reionization neutral hydrogen fraction in CoDa II is about 1 dex lower than observed. Such an offset is not uncommon and the literature shows that simulations of the EoR can exhibit a variety of such departures from the observed evolution of the ionized fraction measured from quasar lines of sight (e.g. Aubert & Teyssier 2010; Petkova & Springel 2011; Gnedin & Kaurov 2014; So et al. 2014; Zawada et al. 2014; Aubert et al. 2015; Bauer et al. 2015; Ocvirk et al. 2016; Aubert et al. 2018; Ocvirk et al. 2019; Wu et al. 2019). This issue is likely related to the evolution of the ionizing UV background in CoDa II, shown in Fig. 3. The surge between z = 7 and z = 6 correlates with the drop of the neutral fraction, and CoDa

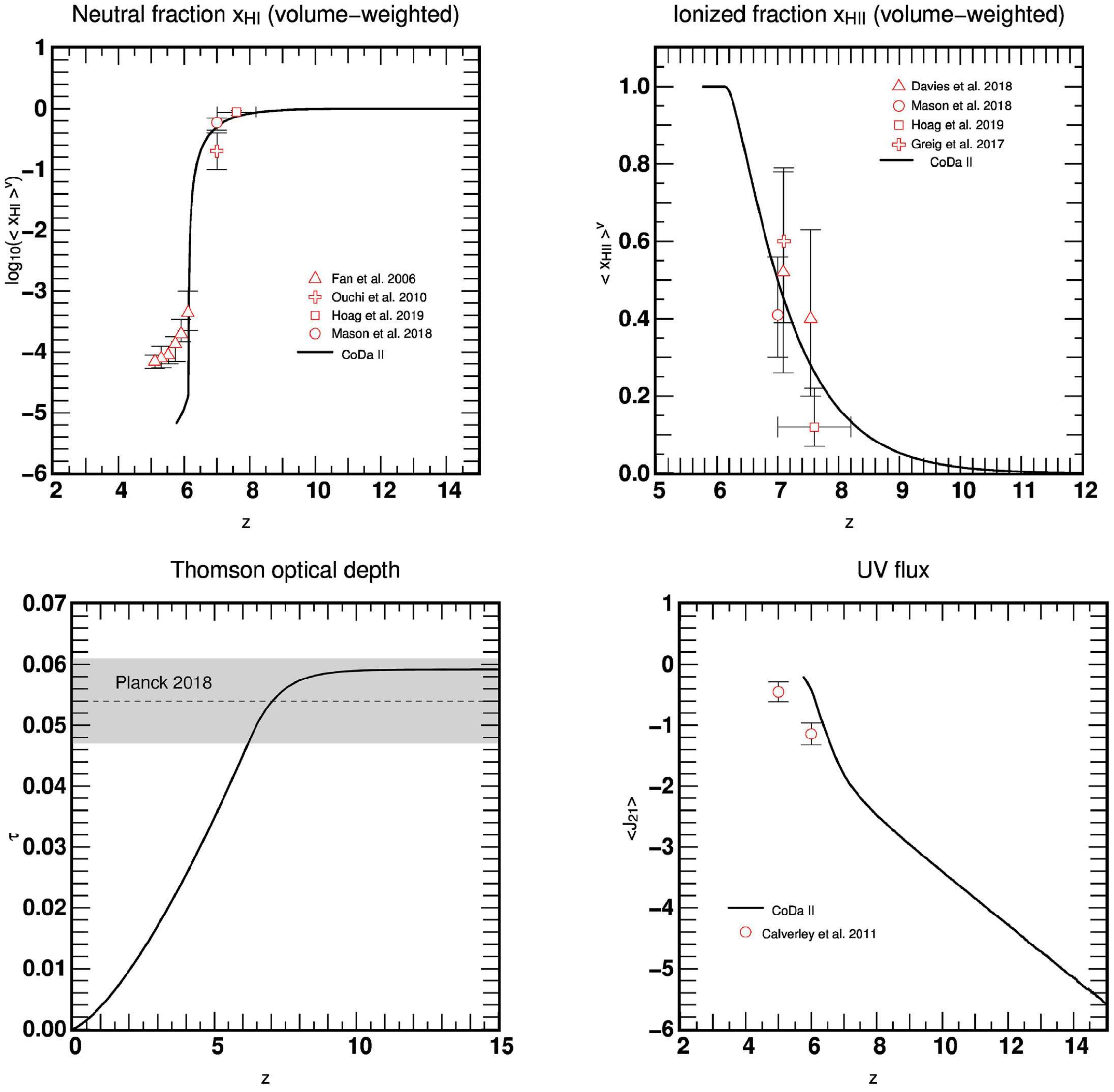


Figure 3. CoDa II simulation global reionization history, compared to observations: the evolution of the volume-weighted average of the neutral hydrogen fraction (top left) and ionized hydrogen fraction (top right), the Thomson scattering optical depth caused by free electrons in the IGM (integrated thru the simulated universe from $z = 0$ to each z, using the average ionized fraction in CoDa II at each z thru the end of the EOR and assuming the IGM was fully ionized thereafter), compared to the value integrated thru the entire post-recombination universe according to observations of CMB polarization anisotropy by Planck 2018 (bottom left), and the evolution of the average ionizing UV flux density (bottom right).

II overshoots the observations of Calverley et al. (2011), resulting in a lower-than-observed neutral fraction.

It is commonly accepted that resolving Lyman-limit systems is necessary to describe properly the population of absorbers in the IGM and predict the correct average neutral hydrogen fraction (Miralda-Escudé, Haehnelt & Rees 2000; Shukla et al. 2016). In particular, Rahmati & Schaye (2018) state a physical size for Lyman-limit systems (hereafter LLSs) of 1–10 kpc in their simulation. While this is compatible with the spatial resolution of CoDa II (∼3.3 physical kpc at $z = 6$ and less at higher redshifts), LLSs may be only marginally resolved and this could have an impact on our predicted neutral fraction after overlap. This is sufficient to resolve the scale of Jeans-mass filtering in the photoionized IGM after it is photoheated to $\sim 10^4$ K. However, according to Emberson, Thomas & Alvarez (2013) and Park et al. (2016), still smaller scale structure, on the scale of the Jeans mass in the cold, neutral IGM before it is reionized, must be resolved in order to capture fully this missing opacity of self-shielded regions, on the minihalo scale. A suite of higher resolution simulations with RAMSES–CUDATON are required to understand this issue, and are beyond the scope of this paper.

3.2 UV luminosity function

The UV luminosity function is the other important observable simulations of the EoR must match. To compute magnitudes, we first estimate the halo virial radius as R_{200}, the radius within which

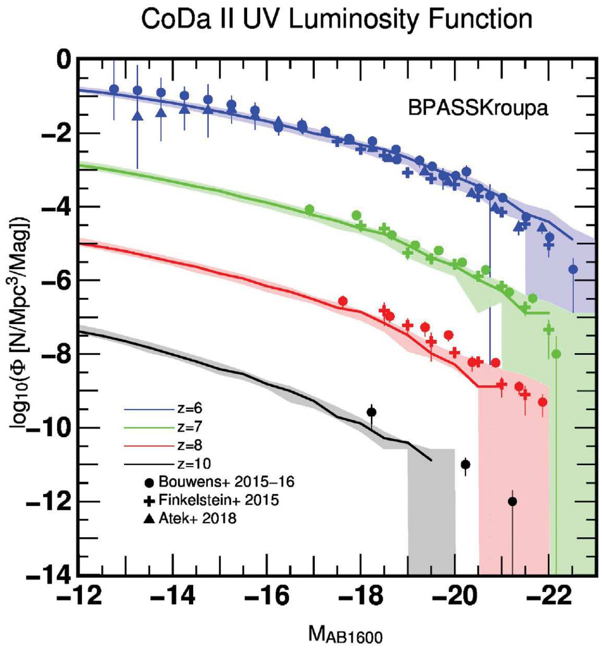


Figure 4. CoDa II UV luminosity functions, and comparison with observations. The full circles, crosses, and triangles with error bars are the observations from Bouwens et al. (2015, 2017), Finkelstein et al. (2015), and Atek et al. (2018), respectively, at $z = [6,7,8,10]$, from top to bottom, while the shaded area and the thick line show the envelope and the median of the LFs of 5 equal, independent, rectangular sub-volumes taken in the CoDa II simulation. For clarity, the LFs have been shifted downwards by 0, 2, 4, and 6 dex for redshifts $z = [6,7,8,10]$, respectively.

the average density of the dark matter is 200 times the cosmic mean dark matter density

$$R_{200} = \left(\frac{3\, M_{\rm FoF}}{4\pi \times 200\, \langle \rho_{\rm DM} \rangle} \right)^{1/3}, \tag{2}$$

where $M_{\rm FoF}$ is the halo mass given by FoF and $\langle \rho_{\rm DM} \rangle$ is the cosmic mean dark matter density. We then computed the $M_{\rm AB1600}$ magnitudes using our BPASS $Z = 0.001$ binary population model, assuming no dust extinction, for all the star particles associated with each individual halo, evolving the stellar populations inside each star particle over time, according to the birth-times of each particle.

The results are shown in Fig. 4, along with observational constraints. The latter are taken from Bouwens et al. (2015, 2017) and Finkelstein et al. (2015), which have been shown to be in broad agreement with a number of other studies including Oesch et al. (2013) and Bowler et al. (2015). We also show the $z = 6$ data of Atek et al. (2018), to highlight the current uncertainties of the very faint-end obtained using gravitational lensing in cluster fields. The luminosity functions (hereafter LF) have been shifted vertically for clarity. The shaded area shows the envelope of the LFs obtained for 5 non-overlapping, rectangular subvolumes of the CoDa II simulation. Each of these subvolumes spans ∼168 000 cMpc^3 ($\Delta x = {\rm L_{box}}/5$, $\Delta y = \Delta z = {\rm L_{box}}$, i.e. 1/5 of the full box volume), which is similar to the volume probed by CANDELS-DEEP at $z = 6$. The resulting envelope therefore illustrates the expected effect of cosmic variance at $M_{\rm AB1600} > -20$. The thick solid line shows the median of these 5 LFs. The CoDa II LF is in good agreement with the observations at all redshifts, although a small shift (i.e. in the sense of undershooting) can be seen for $z > 6$, which increases at higher redshifts. We note, however, that given the error bars and the spread of the observed LFs, our results remain compatible with the observations.

3.3 Cosmic star formation rate density

Observations of high-redshift galaxies have been used to infer the global star formation rate density (‘SFRD’) of the universe at early times, indirectly, by measuring the UV LF of galaxies and assuming an IMF for the stars responsible for the UV starlight that escapes from each galaxy at 1600 Å, subject to a possible correction for an unknown amount of attenuation inside each galaxy by interstellar dust. In fact, there is considerable uncertainty that results from the fact that the bulk of the UV luminosity per galaxy is released by the most massive stars in the galaxy’s IMF, as long as star formation continues, while the bulk of the mass in stars formed is generally assumed to be dominated by the unseen lower mass stars in the conventionally assumed IMFs. Furthermore, once star formation ceases or slows down significantly inside a galaxy, its stellar population will evolve passively and its UV luminosity decline, after its most massive stars have died a fiery death as SNe and no longer contribute their UV starlight. In a fully coupled radiation-hydrodynamical simulation of reionization and galaxy formation like CoDa II, in which the release of ionizing UV starlight and SN energy inside galaxies are very important in affecting the evolution both inside and between galaxies, it is only those massive stars that matter, and any inference about the total mass of stars formed along with those massive ones is similarly subject to an assumption about the shape of the IMF over the mass range which is well below the level of 10’s of $M_\odot$ which dominates the production of ionizing photons. Nevertheless, with these caveats, we can compare our predictions for the SFRD in CoDa II with the observational constraints by making the same underlying assumption about the IMF for both the simulated SFR and the inferred one from the observed galaxies, as follows.

We computed CoDa II’s cosmic SFRD as the global star formation rate of the whole CoDa II volume divided by the box volume. The results are shown in Fig. 5. As a first remark, we note that the cosmic SFRD in CoDa II, as in CoDa I (O16), increases at all times, unlike the simulation of So et al. (2014), which shows a decline at late times. In the latter, the authors suggested this could be due to the small box size they used. This may indeed be the case, since CoDa II is more than 95 times larger in volume. The fact that the global SFRD in CoDa II increases continuously over time right through the end of the EOR was noted previously for CoDa I in O16. As pointed out there, this conflicts with the prediction by Barkana & Loeb (2000) that the global SFR would suffer a sharp drop as the end of reionization approached, caused by a jump in the local IGM Jeans-mass filter scale inside H II regions as reionization overtook a significant fraction of the volume of the universe. While the SFR dependence on halo mass does reflect this suppression effect at lower mass, the global SFR integrated over halo mass in CoDa II, like in CoDa I, is apparently dominated by galaxies more massive than those whose SFR is suppressed by reionization feedback.

In order to compare this integrated SFRD of CoDa II with the observations of high-redshift galaxies, we must put the latter on a self-consistent footing with the assumptions of the CoDa II stellar emissivity and IMF. As described in Section 2.2.2, our adopted value of the emissivity in ionizing photons per stellar baryon, $\xi_{\rm ph,IMF} = 1140$, assigned to each star particle when they form, is consistent with the ionizing emissivity of a BPASS model for a $Z = 0.001$ binary population with Kroupa IMF, averaged over the first 10 Myr. In practice, this means that, to compare the mass in star particles formed per time in the simulation volume with that

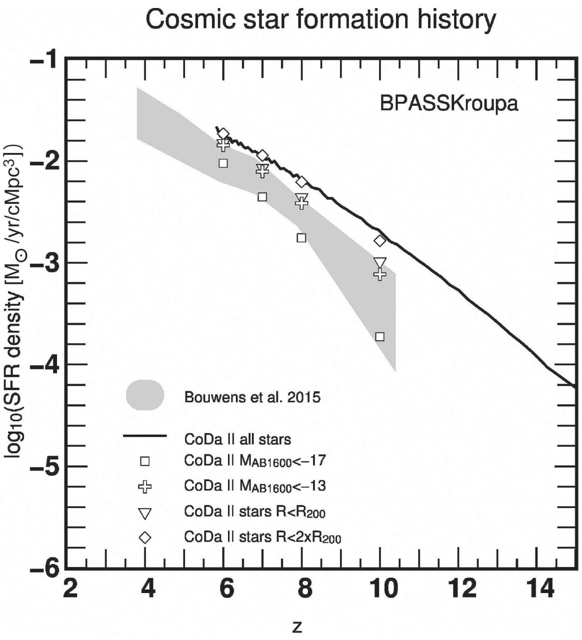


Figure 5. Evolution of the cosmic star formation rate density (SFRD) as simulated by CoDa II, compared with that inferred from observations: (1) Using all star particles in the simulation box (thick line), (2) accounting for haloes above realistic magnitude cuts (crosses and squares), and (3) using all star particles within the R_{200} and $2R_{200}$ spheres of haloes (triangles and diamonds). Observational result is taken from table 7 of Bouwens et al. (2015), but with the observed SFR decreased by 0.34 dex to account for the different M_{AB1600} – SFR conversion assumed here, resulting from our BPASS source model, rather than the Salpeter IMF adopted by Bouwens et al. (2015).

inferred from observations of the UV LF of high-redshift galaxies, those observations must be interpreted in terms of the same assumed IMF when inferring the observed SFRD and its cosmic evolution. Also shown in Fig. 5 are the observational constraints taken from Bouwens et al. (2015): the grey area shows the envelope bracketing the dust-corrected and dust-uncorrected SFRD's. We decreased this observed SFRD by 0.34 dex to account for the different M_{AB1600} – SFR conversions resulting from our assumption of the BPASS model, rather than the Salpeter IMF for single stars assumed by Bouwens et al. (2015). We computed this offset by applying the methodology of Madau, Pozzetti & Dickinson (1998) to our BPASS source model. This involves computing the 1600 Å luminosities of a population undergoing an exponentially decaying burst of star formation, resulting in a conversion factor:

$$L_{1600} = \mathrm{const}\,\frac{\mathrm{SFR}}{\mathrm{M_{\odot}yr^{-1}}}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}, \tag{3}$$

where const = 1.75×10^{28} with our source model, and const = 8×10^{27} in B15.

The CoDa II total SFRD (thick solid line) overshoots the observations by about 0.1 dex at $z = 6$. However, the cosmic SFRD inferred by B15 only includes galaxies with $M_{AB1600} < -17.7$, and does not account for fainter (undetected) objects. Therefore, to compare our CoDa II results to B15 in a fair way, we need to apply similar magnitude cuts to both. With such cuts (squares in Fig. 5), CoDa II falls within the observationally favoured SFRD at $z = 6$, and very close to the observations at higher redshifts.

At $z = 6$, the SFRD inferred from the CoDa II sample with $M_{\mathrm{AB1600}} < -17$ is 2.5 times smaller than the total SFRD in the box, and the offset grows with increasing redshift. However, not all that offset is due to the magnitude cuts: the triangles in Fig. 5 show the cosmic SFRD obtained when considering only those star particles which are located within a sphere of radius R_{200} centred on the centre of each halo, as opposed to considering all the star particles in the box. Even when considering all haloes, the R_{200}-bounded SFRD is always smaller than the full-box SFRD, showing that some of the star particles are located either beyond R_{200} for haloes detected by the FoF halo finder or else are not associated with haloes because their haloes are below our assumed lower halo mass limit of 10^8 $\mathrm{M}_{\odot}$ (246 dark matter particles). Haloes, in particular at the high redshifts studied here, are strongly triaxial, so that their particles and stars may extend beyond a sphere of radius R_{200}. This aspect is illustrated in more detail in Appendix C, where we show that stars beyond R_{200} are still associated with the most nearby dark matter halo. We further confirm this by considering all stars within $2R_{200}$ (diamonds in Fig. 5), yielding an SFRD which matches fhe total full box SFRD at redshifts below $z = 8$. Only at $z = 10$, some of the stars may be associated to small haloes below the assumed limit.

Finally, comparing the 'detectable' CoDa II galaxy population with the total SFRD within the R_{200} spheres of all haloes, we find that the CoDa II sample with $M_{\mathrm{AB1600}} < -17$ at $z = 6$ accounts for about 63 per cent of the halo $R < R_{200}$ SFRD, and only 18 per cent at $z = 10$. On the other hand, in the case of deeper detection limits such as in lensing cluster fields where lensing magnification makes it possible to detect fainter haloes with $M_{\mathrm{AB1600}} < -13$, these fractions increase to 94 per cent and 76 per cent, respectively.

3.4 The very faint end of the luminosity function

We can also use CoDa II to investigate the LF at magnitudes fainter than currently observed. We show in Fig. 6 the whole LF, going down to $M_{\mathrm{AB1600}} = -5$, for 4 redshifts. The progressive build-up of the LF with cosmic time is clearly apparent, both at the bright end, with more and more massive haloes hosting vigorous star formation, and at the faint end, made of passively evolving galaxies. The maximum of the LF is reached at $M_{\mathrm{AB1600}} \sim -11$ at all redshifts. At redshifts lower than 10, a plateau develops faintwards of -11, extending to fainter magnitudes as time goes by. At $z = 6$, the plateau extends down to $M_{\mathrm{AB1600}} = -6$. We show in Fig. 7 that this $[-11, -6]$ magnitude range corresponds to masses of 10^8–10^9 $\mathrm{M}_{\odot}$, i.e. the mass range where SFR suppression by ionizing radiation takes place (see Section 3.5). Indeed, if haloes below $\sim 10^9$ $\mathrm{M}_{\odot}$ have their SFR's suppressed after they are overtaken by reionization and, thereafter, have only the luminosity associated with the passive evolution of their low-mass stars (after the high-mass stars finished their lives as SNe), we expect the number density of galaxies with magnitudes fainter than $M_{\mathrm{AB1600}} = -11$ to grow continually, as more and more volume is overtaken by reionization, as shown by our figure.

The full LFs for these 4 redshifts are given in the appendix in tabular form.

3.4.1 *Halo mass and UV luminosity*

The results of CoDa II show that the UV continuum luminosity of a galaxy is correlated with the mass of its galactic halo. The average luminosity of haloes of the same mass at a given redshift increases with halo mass, consistent with the correlation found for Coda I (O16). This is shown for CoDa II in the panels of Fig. 7,

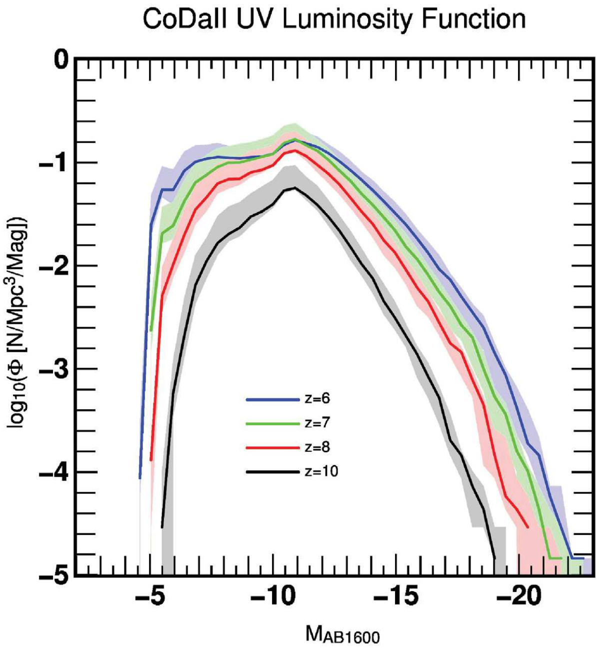


Figure 6. CoDa II UV luminosity functions, including the very faint end, at 4 epochs. The shaded area and the thick line show the envelope and the median of the LFs of 5 equal, independent, rectangular sub-volumes taken in the CoDa II simulation. The full LFs are given in a table in appendix.

which correspond to the galaxy populations of the $z = 6$ and $z = 10$ UV LF's in Fig. 4. Moreover, for each halo of a given mass, the luminosity can fluctuate significantly over time, which is reflected in the vertical spread of the distributions in Fig. 7. At $z = 6$, the vertical dispersion increases with decreasing halo mass. Moreover, the spread at low mass is larger at $z = 6$ than $z = 10$. The increase in dispersion is not driven by the brightest galaxies of the low-mass haloes. Indeed, the brightest galaxies residing in a haloes of mass a few times 10^8 $\mathrm{M}_{\odot}$ have $M_{\mathrm{AB1600}} \sim -12$, in both snapshots, and contain a couple of 1–2 Myr old elementary stellar particles in CoDa II. By contrast, the faintest galaxies of the $z = 10$ snapshot have $M_{\mathrm{AB1600}} \sim -5.5$, while the faintest galaxies of the $z = 6$ snapshot are as faint as $M_{\mathrm{AB1600}} \sim -5$. This is the signature of passively evolving galaxies: their star formation has stopped, and their stellar populations progressively fade, making them fainter as redshift goes down.

3.5 Impact of radiative feedback on galaxy formation

As described in Section 1, the CoDa suite of fully coupled radiation-hydrodynamics simulations was designed to address the long-standing question of how galaxies that released the UV H-ionizing radiation that reionized the universe were affected by the photoheating that accompanied photoionization during the EoR, acting back on their gas content and star formation rates, and on the progress of reionization, itself. In O16, we summarized the globally averaged effect of this feedback on the star formation rates of haloes of different mass at different redshifts in CoDa I, showing that the SFR in low-mass haloes below a few times $10^9\mathrm{M}_{\odot}$ was reduced towards the end of reionization, as more and more of the volume was reionized. In Dawoodbhoy et al. (2018), we studied these CoDa I results in more detail, going beyond the globally averaged results to identify the causal connection between the suppression of star formation in low-mass haloes and reionization by showing that SFR suppression followed the local time of arrival of reionization at the location of each galaxy. In that case, we showed, reionization was not limited to the late stage of reionization, as it appeared from the globally averaged results in O16, but was happening throughout the EoR, at different times in different places.

In Dawoodbhoy et al. (2018), we summarized the literature that had attempted to characterize this feedback effect and establish the mass range of haloes subject to suppression, by variety of approximations over the years, including simulations. We refer the reader to that summary and the references therein, for this background. In the context of 3D galaxy formation simulations, a number of authors have attempted to address this impact of radiative feedback on galaxy formation, by adopting a pre-determined, uniform UV background like that modelled by Haardt & Madau (1996, 2012), and more recently with fully coupled RHD simulations (Pawlik, Milosavljević & Bromm 2013; Jeon et al. 2014; Wise et al. 2014; Aubert, Deparis & Ocvirk 2015; Pawlik, Schaye & Dalla Vecchia 2015; Rosdahl et al. 2015; Ocvirk et al. 2016; Dawoodbhoy et al. 2018; Rosdahl et al. 2018; Katz et al. 2019; Wu et al. 2019), investigating a possible suppression or reduction of star formation in low-mass galaxies. Here, we will highlight some of the results of the CoDa II simulation on this issue and compare them with those in CoDa I.

We computed the instantaneous SFR of CoDa II haloes as the stellar mass formed within a sphere of radius R_{200} centred on the dark matter halo centre of mass, during the last 10 Myr, divided by a duration of 10 Myr. Fig. 8 shows the instantaneous SFR that results, as a function of the instantaneous mass of the dark matter halo, for several redshifts.

There is a general trend for more massive haloes at all epochs to form more stars, as seen in Ocvirk, Pichon & Teyssier (2008) and O16. Here, we find a trend: SFR $\propto M^{\alpha}$ for $M > 10^{10}$ $\mathrm{M}_{\odot}$ with a slope $\alpha \sim 5/3$, just as in CoDa I. This slope is compatible with the values found in the literature i.e. $1 < \alpha < 2.5$, in numerical and semi-analytical studies, such as Tescari et al. (2009), Hasegawa & Semelin (2013), Yang et al. (2013), Gong et al. (2014), Aubert et al. (2015), and Ocvirk et al. (2016). Moreover, Fig. 8 shows a smooth decrease in SFR for the low-mass haloes, around $\sim 2 \times 10^9$ $\mathrm{M}_{\odot}$, relative to this slope at higher mass. Below this mass, the SFR–mass relation is slightly steeper, and gets even steeper at lower redshifts, as reionization advances. This steepening at late times during the EoR is a signature of suppression by reionization feedback, also seen in the globally averaged results from CoDa I. However, the strength of the suppression is stronger in CoDa I than in CoDa II.

In order to understand the origin of this evolution in the SFR–mass relation, in particular with respect to the relative importance of supernova feedback and radiative feedback, we performed two additional control simulations, with the same parameters as CoDa II, but in a smaller box, 8 h^{-1}cMpc on a side, isolating the feedback physics:

(i) *SN, RT*: this simulation has the exact same physics as CoDa II, i.e. SNe and radiation.

(ii) *SN, no RT*: this simulation was run with SNe but without radiation (and therefore does not undergo global reionization).

Both simulations ran to redshift $z = 5$. Fig. 9 shows the star formation history of haloes present in the final snapshot at $z = 5$, binned according to their final masses at this redshift. Again, the trend of increasing SFR with halo mass is clear. Also, expectedly, since $z = 5$ haloes were less massive at, say, $z = 15$, their SFR is lower at higher redshift. For all mass bins, the SFR of the run with radiation is smaller than without radiation. While the effect is very small

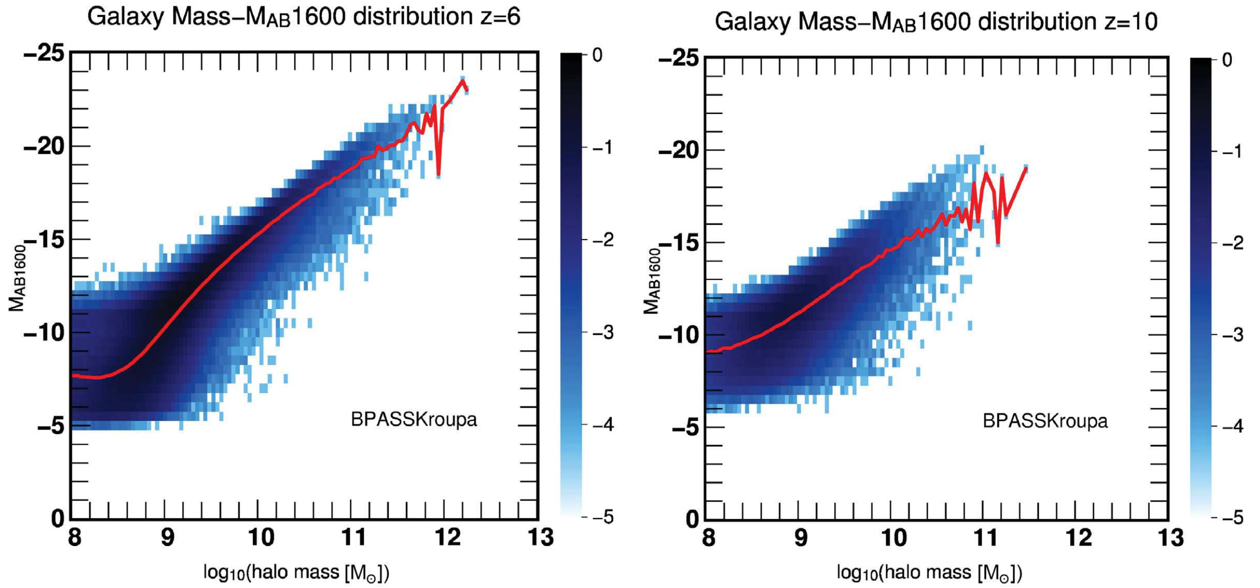


Figure 7. CoDa II galaxy mass – magnitude distributions at $z = 6$ and $z = 10$. The colour indicates the galaxy comoving number density in N/Mpc^3/Mag/log(M$_\odot$). The red line indicates the average magnitude for each mass bin.

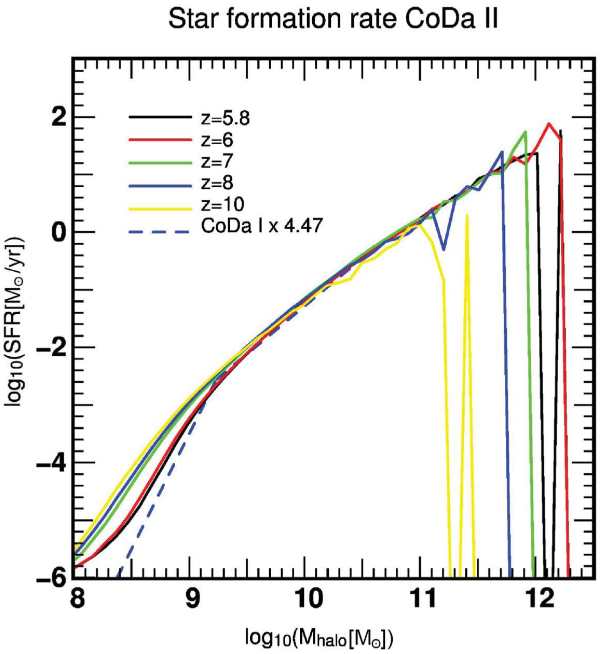


Figure 8. Instantaneous star formation rate per halo as a function of instantaneous halo mass, for various redshifts. The instantaneous SFR is computed as the stellar mass formed within an R_{200} radius sphere centred on the dark matter halo centre of mass, during the last 10 Myr, divided by a duration of 10 Myr. The solid lines show CoDa II data while the dashed line shows the post-overlap CoDa I data, multiplied by a factor to make it match the high-mass regime end, for easier comparison.

in the highest mass bin, it gets stronger the lower the mass. The lowest mass bin, in particular, shows a striking behaviour: the SFR increases with time, and then starts to decrease around $z = 6.5$, i.e. during overlap and shortly before the end of reionization. This is markedly different from the run without radiation, in which the low-

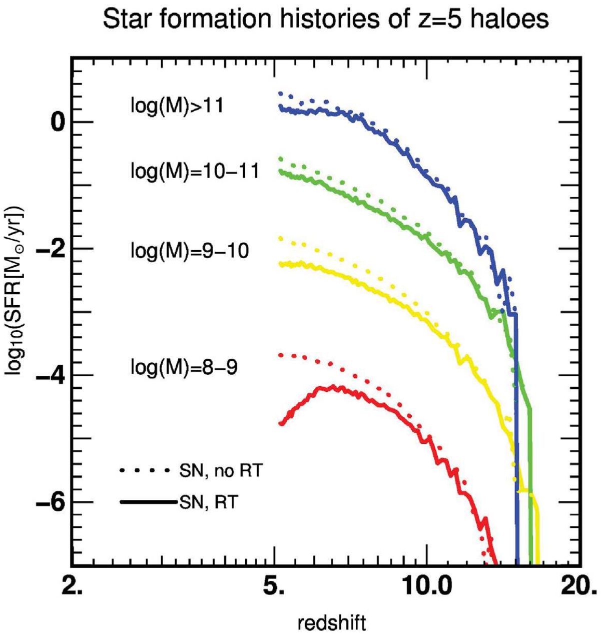


Figure 9. Star formation histories of 4 halo mass bins for two simulations in a test box 8 h^{-1}cMpc on a side. *Solid line*: with full fiducial physics (supernovae and radiative transfer), *dotted line*: with supernovae only, no radiation. The mass bins correspond to the haloes final mass (i.e. halo mass measured in the final simulation time-step, at $z \sim 5$.).

mass haloes have a monotonically increasing SFR. This demonstrates that the UV ionizing radiation suffusing the Universe during the EoR is indeed the cause of the withering SFR of the low-mass haloes in CoDa II.

This gradual suppression reflects a radiation-driven reduction in the baryonic fraction of low-mass haloes (Dawoodbhoy et al. 2018; Sullivan, Iliev & Dixon 2018). In contrast, more massive haloes are

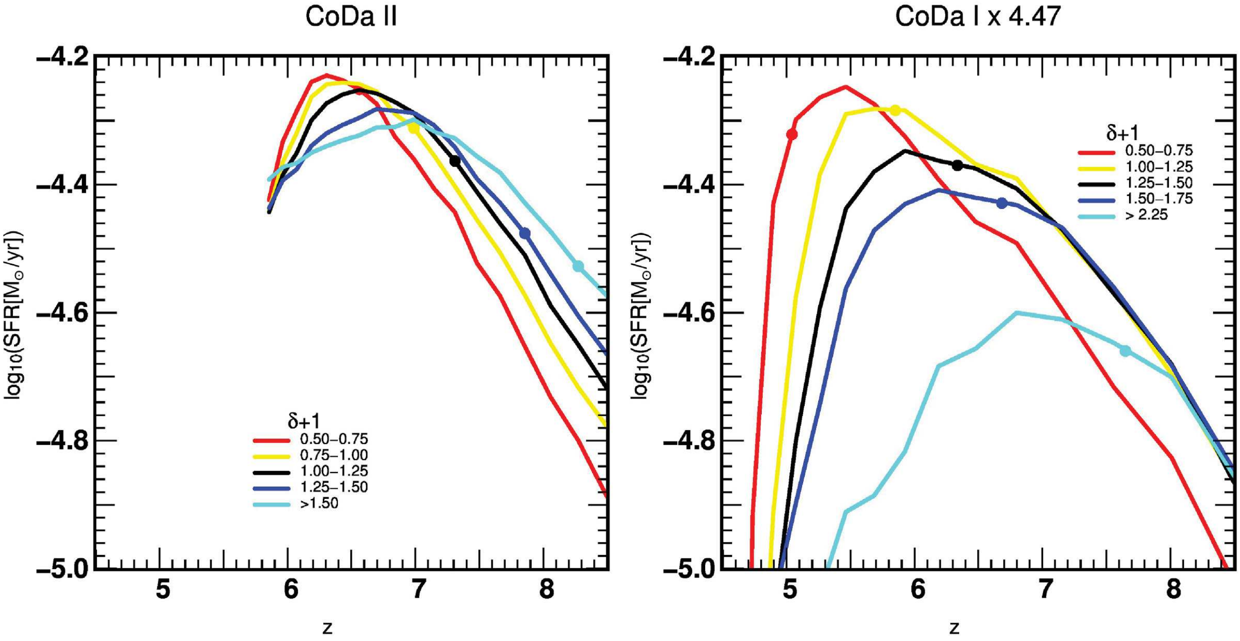


Figure 10. Average star formation histories of low-mass haloes ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$) in CoDa II (left) and CoDa I (right), for haloes present in that mass bin just after global reionization ended ($z = 5.8$ for CoDa II and $z = 4.44$ for CoDa I) for different environments, quantified by their overdensity $\delta + 1 = \rho/\bar{\rho}$ computed on 4 h^{-1}cMpc sub-boxes. The filled circle on each curve shows the average reionization redshift (when ionized fraction was 0.5) of each overdensity bin. Note that the overdensity bins we used are slightly different between the two simulations, because the lower reionization redshift of CoDa I results in a higher maximum overdensity.

able to retain their gas and keep on forming stars at a very similar rate pre- and post-overlap. We note that the reduction in the SFR of low-mass haloes due to global photoheating during reionization is less severe than in CoDa I. Indeed, the sub-grid model for star formation in CoDa I included an additional criterion, allowing stars to form only in cells colder than $T = 20\,000$ K, thereby making a smaller amount of cells in the simulation eligible for star formation after local reionization occurs. This more stringent requirement led to a more severe suppression of star formation in low-mass haloes (dashed curve in Fig. 8). This requirement was relaxed in CoDa II, where star formation is purely based on a gas overdensity threshold and an efficiency parameter (see Table 1). It is interesting to note that suppression nevertheless still clearly happens in Coda II at low masses, although in a slightly less severe form.

We performed a resolution study, presented in Appendix A, where we show that increasing the spatial resolution of our setup by up to a factor 4 in space and 64 in mass (although in a smaller test box), results in similar star formation suppression in the low-mass halo range (10^8–10^9 M$_\odot$) after reionization. A quick comparison with the literature, e.g. Pawlik et al. (2015) and Wu et al. (2019), shows reasonable agreement. Therefore, star formation suppression in CoDa II is not likely to be an artefact due to numerical resolution.

3.6 Environment and the timing of SFR suppression in low-mass galaxies

Reionization is a patchy process, as shows Fig. 1. Therefore in the context of the radiative feedback demonstrated in Section 3.5, we may expect different environments (galaxy cluster or group progenitor, galaxy neighbourhood, voids), to affect low-mass galaxies' star formation rates with different strengths and at different times. The CoDa II simulation is well suited to such investigations thanks to the variety of environments it contains and its self-consistent modelling of the inhomogeneous progress of reionization. Here, we wish to illustrate this aspect, while leaving a more complete, detailed study in the spirit of Dawoodbhoy et al. (2018) for a later paper.

We computed the overdensity in which CoDa II galaxies reside at the end of the simulation, as $\delta + 1 = \rho/\bar{\rho}$ averaged in sub-boxes of 4h^{-1}cMpc on a side. This is the typical scale of a galaxy group progenitor, and therefore this quantifies the environment at an intermediate, meta-galactic scale. We divide our galaxy population into 5 overdensity bins and compute the average star formation histories of CoDa II low-mass haloes ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$ at $z = 5.8$. The results for CoDa II are shown in the left-hand panel of Fig. 10 for the 5 overdensity bins. Each curve shows the same characteristic parabolic shape seen for low-mass haloes in Fig. 9, with a timing and amplitude clearly correlated with the local overdensity.

In particular, haloes in dense environments have a more intense early ($z = 8$) star formation history. This is likely caused by the inherently faster formation of their dark matter haloes and therefore larger mass accretion rates, as compared to underdense regions, as shown in Maulbetsch et al. (2007).

The filled circle overplotted on each curve shows the average redshift at which the cells in that overdensity bin reached the point of ionized fraction 0.5.

In the highest density environments, the star formation in low-mass haloes starts early, culminates around $z = 7$, and then decreases.

In underdense regions, the same scenario unfolds, with a delay: star formation starts later, reionization happens later, and therefore the average star formation rate peaks and decreases later as well.

Strikingly, at $z \sim 6.2$, low-mass haloes in underdense regions form stars at a higher rate than their counterpart in overdense regions. However, this episode is short-lived, and as the epoch of reionization comes to an end, the average star formation rates of low-mass haloes seem to converge to a similar value independent of their environment.

These results for CoDa II are consistent with the results found for CoDa I by Dawoodbhoy et al. (2018), both for the correlation of the local reionization times (or redshifts) with local overdensity and the suppression of SFR in low-mass haloes following their local reionization time. We note that in Dawoodbhoy et al. (2018), the definition used for reionization time (or redshift) was the 90 per cent-ionized time, which is somewhat lower redshift than the 50 per cent-ionized time used here. Hence, the delay found here for CoDa II between the half-ionized redshift and the redshift at which the SFR peaks and then declines for a given overdensity bin, due to suppression, is larger here due to this choice of half-ionized reionization redshift.

To make a direct comparison here between CoDa II and CoDa I, we show in the right-hand panel of Fig. 10 the same analysis, performed this time on CoDa I data, for the same range of halo masses ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$) just after the end of global reionization, i.e. $z = 4.4$. The SFRs for CoDa I have been multiplied by 4.47 as in Fig. 8 to make them comparable in amplitude to CoDa II results. Because of the lower redshift, the maximum overdensity of the simulation is higher. The behaviour of low-mass galaxies share several similarities between the two simulations: in the mass range considered, the SFRs of all overdensity bins rise, culminate, and decay, correlated with the average reionization times of their bin. Also, the earlier suppression taking place in overdense regions is confirmed in CoDa II, although in a less severe form, as was already noted in Section 3.5. We note however that we are unable to predict what would happen if CoDa II had run to lower redshift. It could be that at $z = 5$, CoDa II low-mass haloes would end up with very low SFRs, comparable to those of CoDa I. Further simulations will be needed to address this aspect conclusively.

3.7 Mapping the inhomogeneous timing of reionization in the Local Universe: 'islands in the stream'

The effects of reionization arrive at different times in different places, not only globally, but within the Local Universe, as well. The CoDa II simulation, with its constrained initial conditions, was designed to capture this aspect of the reionization of the Local Universe. To illustrate it, we computed the reionization redshift of CoDa II gas cells as the redshift at which the ionized fraction of the cells reaches 50 per cent for the first time. This was done using the level 10 ionized fraction of the simulation, yielding a reionization redshift field of 1024^3 cells (4096^3 is level 12). Using the level 10 instead of level 12 reduces the memory requirements of this process by 64, but does not impact our results. A reionization map is then created by a slice through this reionization redshift field 1 h^{-1}cMpc thick, shown in Fig. 11. The progenitors of several major Local Universe objects analogues were traced back from $z = 0$ to $z = 6$ using the CoDa II-DM2048 simulation. They are labelled on Fig. 11 to allow us to gauge their reionization history and radiative influence region during the epoch of reionization in the Local Universe.

The map shows local maxima of z_r, which are 'islands' of early reionization surrounded by a 'sea' in which reionization happened at a later time. Such regions are internally reionized, in an inside-out pattern.

As can be seen in Fig. 11, the MW and M31 are examples of such islands, not only surrounded by the sea, but separated from each other by a channel, so they appear not to have been ionized from outside or to have significantly influenced each other during their reionization. For comparison, An example of a reionization map in a scenario where M31's progenitor partially or completely reionizes the MW progenitor is shown in fig. 4 of Ocvirk et al. (2013). However, the latter scenario is disfavoured because it requires an extreme feedback model, as discussed in Ocvirk et al. (2013).

Investigating further out, we see that the Virgo cluster's progenitor, is also separated from the LG progenitors by later-to-reionize regions. It is therefore not responsible for the reionization of the LG in the CoDa II simulation. A similar result was obtained by the CoDa I-AMR simulation with the EMMA code, as described in Aubert et al. (2018), starting from CoDa I initial conditions, which were an earlier version of CLUES constrained realization initial conditions of the Local Universe.

Finally, the CenA and M81 groups' progenitors also produced their own self-reionized islands. Despite their proximity, they do not appear to have affected the reionization of the Local Group, or each other, as the respective islands of CenA, M81, MW, and M31 are well-separated from each other.

4 CONCLUSIONS

CoDa II is a state-of-the-art, fully coupled radiation-hydrodynamics simulation of reionization and galaxy formation, in a volume large enough to model global reionization and with a uniform spatial resolution high enough to resolve the formation of all the galactic haloes in that volume above $\sim 10^8\,\mathrm{M}_\odot$. Based upon the massively parallel, hybrid CPU–GPU code RAMSES–CUDATON, CoDa II utilizes 4096^3 dark matter particles on a uniform cubic lattice of 4096^3 grid cells, which is also used for the gas and radiation field, in a comoving box 94 Mpc on a side, centered on the Local Group in a constrained realization of the Local Universe. CoDa II is a successor to our first such simulation, CoDa I, described in O16 and Dawoodbhoy et al. (2018), but with new initial conditions, updated parameters for the background universe, and a recalibration and modification of our subgrid model for star formation, chosen to ensure that CoDa II finished reionizing the universe before $z \sim 6$. For both CoDa I and II, our initial conditions were based upon a constrained realization of the Gaussian random density fluctuations for the ΛCDM model, chosen by 'reverse engineering' the structure of the present-day Local Universe, to yield initial conditions in the linear regime that, when evolved over cosmic time, will reproduce that observed structure. This makes CoDa I and II both suitable for modelling both global and local reionization and its impact on galaxy formation, including the progenitors of the Local Group. The galaxy data and background cosmology on which the constrained realization for CoDa II is based were updated and improved relative to those used for CoDa I.

CoDa II was performed on the hybrid CPU–GPU supercomputer Titan at OLCF, using 16 384 nodes and 16 384 GPUs (one GPU per node, 4 cores per GPU), for a total of 65 536 cores. This is 2 times the node and GPU count of CoDa I, and an 8-fold increase in CPU count. This is the first time a GPU-accelerated, fully coupled radiation-hydrodynamics galaxy formation code has been used on such a scale.

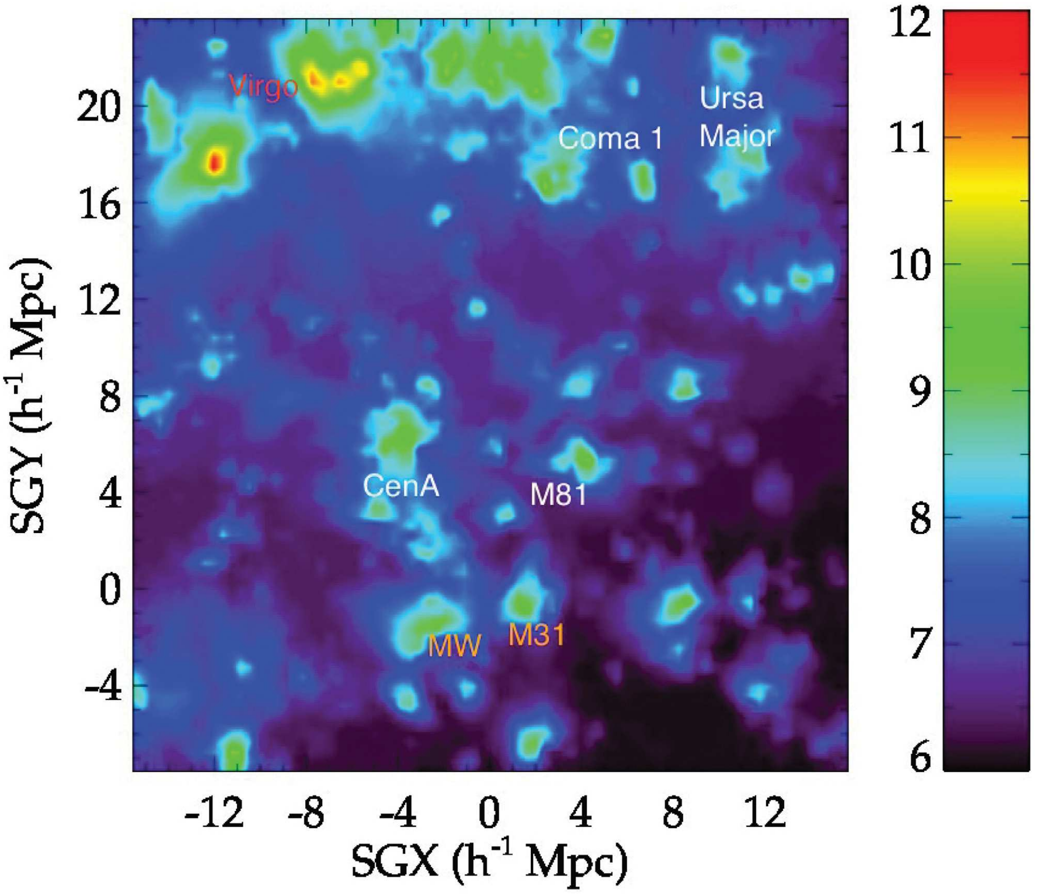


Figure 11. Reionization map of the Local Universe: Islands in the stream. The redshift at which each of the 1024^3 grid cells first reached an ionization of 50 per cent is displayed in a planar-slice 1 h^{-1} cMpc thick through the CoDa II simulation volume, labelled with the names of progenitors of familiar objects in the Local Universe today at their respective locations in the map.

The simulation accurately describes the properties of the gas and its interaction with ionizing radiation, self-consistently, including the growth of typical butterflyshaped ionized regions around the first stars and first galaxies, accompanied by photoheating and the subsequent dynamical back-reaction of the gas in response to pressure-gradients. This back-reaction includes the progressive smoothing of small-scale gas structures and the resistance of ionized intergalactic gas to its gravitational capture by galactic haloes too small to produce potential wells deep enough to overcome these pressure forces.

The simulation is in broad agreement with several observational constraints on the EoR, such as the Thomson scattering optical depth integrated through the IGM, as inferred from measurements of fluctuations in the cosmic microwave background by the *Planck* satellite, and the redshift by which reionization ended, as estimated from measurements of the Lyman α opacity of neutral H atoms in the absorption spectra of high-redshift quasars, due to the intervening IGM. While the timing of the evolution of the ionizing flux density is reasonable in CoDa II, its rapid rise at the end of reionization overshoots the observed value inferred from observations of quasar absorption spectra at $z = 6$ by a factor ~10. This in turn causes the neutral fraction in the post-reionization IGM to be too small compared to observations of the Lyman α forest at this redshift, by a similar factor. This may be a consequence of the existence of additional small-scale structure in the real universe, responsible for additional H I bound-free opacity that limits the mean-free-path of ionizing photons when reionization ends, that was missed by the simulation because it was unresolved.

We compute the UV continuum luminosity functions for our simulated galaxy population, and find them to be in very good agreement with high-redshift galaxy observations, in particular at $z = 6$. At higher redshifts, a small offset appears, increasing to ~1 mag at $z = 10$. We apply magnitude cuts to mimic observational detection limits, and find that the cosmic star formation rate associated with haloes with $M_{\mathrm{AB1600}} < -17$ only captures 63 per cent (18 per cent) of the total SFR in haloes at $z = 6$ ($z = 10$). Integrating down to fainter magnitudes, $M_{\mathrm{AB1600}} < -13$, as attempted in deep cluster fields, captures a much larger fraction of the total SFR in haloes, 94 per cent (76 per cent) at $z = 6$ ($z = 10$). This suggests that calibrating a numerical simulation of galaxy formation by comparing its total star formation rate density to observations without applying the magnitude cut may lead to non-optimal calibration, and in particular too small a simulated star formation rate.

The average star formation rate of individual galaxies increases with their dark matter halo mass, yielding a typical SFR = $M_{\mathrm{halo}}^{\alpha}$

relation, when the SFR is globally averaged, with a slope $\alpha = 5/3$. At low-halo masses (below $\sim 2 \times 10^9$ M$_\odot$), the slope becomes steeper, as galaxies become less efficient at forming stars. As time goes by and reionization progresses, the average slope for these low-mass haloes becomes even steeper, reflecting the suppression of the SFR in low-mass haloes caused by reionization feedback.

We show unambiguously that this behaviour at low-halo mass is the result of the spreading, rising ionizing UV background, by comparing our results with a simulation without radiative transfer. We note that this radiative reduction of the SFR appears to be somewhat less severe than in CoDa I due to different sub-grid physics, but it is still robustly measured.

In contrast, the gas core of high-mass haloes is dense enough to remain cool and/or cool down fast enough to keep forming stars, even if in bursts.

Furthermore, we show that environment has a strong impact on the star formation histories of low-mass galaxies: galaxies in overdense regions stop forming stars earlier than galaxies in underdense regions, supporting the analysis of CoDa I by Dawoodbhoy et al. (2018).

The CoDa II UV luminosity function peaks at $M_{\rm AB1600} = -11$ at all the studied redshifts, and develops into a plateau at fainter magnitudes, populated by radiation-suppressed, fading galaxies. We provide the full CoDa II UV luminosity functions for 4 epochs in tabular form, across 8 decades in luminosity, from $M_{\rm AB1600} = -3$ to $M_{\rm AB1600} = -24$, in the hope that these predictions will be useful for comparison with other simulations, but also to observers analysing high-redshift galaxy observations or designing new deep surveys. We analyse CoDa II's Local Universe analogue region, and find that the Milky Way and its neighbour M31 appear as individual islands in the reionization redshift map. This means that both galaxies reionized in isolation, i.e. they were not reionized by the progenitor of the Virgo cluster, or by nearby groups, or by each other.

ACKNOWLEDGEMENTS

This study was performed in the context of several French ANR (Agence Nationale de la Recherche) projects. PO acknowledges support from the French ANR funded project ORAGE (ANR-14-CE33-0016). ND and DA acknowledge funding from the French ANR for project ANR-12-JS05-0001 (EMMA). The CoDa II simulation was performed at Oak Ridge National Laboratory/Oak Ridge Leadership Computing Facility on the Titan supercomputer (INCITE 2016 award AST031). Processing was performed on the Eos and Rhea clusters. Resolution study simulations were performed on Piz Daint at the Swiss National Supercomputing Center (PRACE Tier 0 award, project id pr37). The authors would like to acknowledge the High Performance Computing center of the University of Strasbourg for supporting this work by providing scientific support and access to computing resources. Part of the computing resources were funded by the Equipex EquipMeso project (Programme Investissements d'Avenir) and the CPER Alsacalcul/Big Data.

The CoDaII-DM2048 / ESMDPL_2048 simulation has been performed at Leibniz Rechenzentrum Munich within the project pr74no.

ITI was supported by the Science and Technology Facilities Council [grant number ST/L000652/1]. JS acknowledges support from the 'l'Oréal-UNESCO Pour les femmes et la Science' and the 'Centre National d'Études Spatiales (CNES)' postdoctoral fellowship programs. YH has been partially supported by the Israel Science Foundation (1013/12). KA was supported by National Research Foundation grants NRF-2016R1D1A1B04935414 and NRF-2016R1A5A1013277

GY also acknowledges support from Ministerio de Economia Industria y Competitividad and Fondo Europeo de Desarrollo Regional (MINECO–FEDER) under research grants AYA2012-31101, AYA2015-63810-P, and AYA2015-63810-P, as well as Ministerio de Ciencia, Innovacion y Universidades (MICIU/FEDER) under grant PGC2018-094975-C21.

PRS was supported in part by U.S. National Science Foundation (NSF) grant AST-1009799, National Aeronautics and Space Administration (NASA) grant NNX11AE09G, NASA/Jet Propulsion Laboratory grant RSA Nos. 1492788 and 1515294, and supercomputer resources from NSF XSEDE grant TG-AST090005 and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. PO thanks Y. Dubois, F. Roy, and Y. Rasera for their precious help dealing with SN feedback in RAMSES and various hacks in pFoF. This work made use of v2.1 of the Binary Population and Spectral Synthesis (BPASS) models as last described in Eldridge et al. (2017).

REFERENCES

Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, MNRAS, 479, 5184
Aubert D., Teyssier R., 2008, MNRAS, 387, 295
Aubert D., Teyssier R., 2010, ApJ, 724, 244
Aubert D., Deparis N., Ocvirk P., 2015, MNRAS, 454, 1012
Aubert D. et al., 2018, ApJ, 856, L22
Baek S., Di Matteo P., Semelin B., Combes F., Revaz Y., 2009, A&A, 495, 389
Barkana R., Loeb A., 2000, ApJ, 539, 20
Bauer A., Springel V., Vogelsberger M., Genel S., Torrey P., Sijacki D., Nelson D., Hernquist L., 2015, MNRAS, 453, 3593
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, ApJ, 843, 129
Bowler R. A. A. et al., 2015, MNRAS, 452, 1817
Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, MNRAS, 412, 2543
Carlesi E. et al., 2016, MNRAS, 458, 900
Chardin J., Haehnelt M. G., Aubert D., Puchwein E., 2015, MNRAS, 453, 2943
Courtois H. M., Pomarède D., Tully R. B., Hoffman Y., Courtois D., 2013, AJ, 146, 69
D'Aloisio A., Upton Sanderbeck P. R., McQuinn M., Trac H., Shapiro P. R., 2017, MNRAS, 468, 4691
Davies F. B. et al., 2018, ApJ, 864, 142
Dawoodbhoy T. et al., 2018, MNRAS, 480, 1740
Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1
Deparis N., Aubert D., Ocvirk P., Chardin J., Lewis J., 2019, A&A, 622, A142
Doumler T., Hoffman Y., Courtois H., Gottlöber S., 2013, MNRAS, 430, 888
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, MNRAS, 420, 2662
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058
Emberson J. D., Thomas R. M., Alvarez M. A., 2013, ApJ, 763, 146
Fan X. et al., 2006, AJ, 132, 117
Finkelstein S. L., Ryan R. E., Jr, Papovich C., Dickinson M., Song M., Somerville R. S., Ferguson H. C., 2015, ApJ, 810, 71
Giallongo E. et al., 2015, A&A, 578, A83
Giroux M. L., Shapiro P. R., 1996, ApJS, 102, 191
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., Kaurov A. A., 2014, ApJ, 793, 30
Gong Y., Silva M., Cooray A., Santos M. G., 2014, ApJ, 785, 72
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
Haardt F., Madau P., 1996, ApJ, 461, 20

Haardt F., Madau P., 2012, ApJ, 746, 125
Haardt F., Salvaterra R., 2015, A&A, 575, L16
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Hasegawa K., Semelin B., 2013, MNRAS, 428, 154
Hinshaw G., Weiland J. L., Hill R. S., Odegard N., 2009, ApJS, 180, 225
Hoag A. et al., 2019, ApJ, 878, 12
Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hoffman Y., Ribak E., 1992, ApJ, 384, 448
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625
Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725
Jeon M., Pawlik A. H., Bromm V., Milosavljević M., 2014, MNRAS, 444, 3288
Katz H. et al., 2020, MNRAS, 494, 2200
Kroupa P., 2001, MNRAS, 322, 231
Kulkarni G., Worseck G., Hennawi J. F., 2019, MNRAS, 488, 1035
Madau P., Haardt F., 2015, ApJ, 813, L8
Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
Maulbetsch C., Avila-Reese V., Colín P., Gottlöber S., Khalatyan A., Steinmetz M., 2007, ApJ, 654, 53
Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1
Mitra S., Choudhury T. R., Ferrara A., 2018, MNRAS, 473, 1416
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Ocvirk P., Aubert D., Chardin J., Knebe A., Libeskind N., Gottlöber S., Yepes G., Hoffman Y., 2013, ApJ, 777, 51
Ocvirk P. et al., 2016, MNRAS, 463, 1462 (O16)
Ocvirk P., Aubert D., Chardin J., Deparis N., Lewis J., 2019, A&A, 626, A77
Oesch P. A. et al., 2013, ApJ, 773, 75
Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63
Ouchi M. et al., 2010, ApJ, 723, 869
Park H., Shapiro P. R., Choi J.-h., Yoshida N., Hirano S., Ahn K., 2016, ApJ, 831, 86
Parsa S., Dunlop J. S., McLure R. J., 2018, MNRAS, 474, 2904
Pawlik A. H., Milosavljević M., Bromm V., 2013, ApJ, 767, 59
Pawlik A. H., Schaye J., Dalla Vecchia C., 2015, MNRAS, 451, 1586
Pawlik A. H., Rahmati A., Schaye J., Jeon M., Dalla Vecchia C., 2017, MNRAS, 466, 960
Petkova M., Springel V., 2011, MNRAS, 412, 935
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration VI, 2018, preprint (arXiv:1807.06209)
Qin Y. et al., 2017, MNRAS, 472, 2009
Rahmati A., Schaye J., 2018, MNRAS, 478, 5123
Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188
Rosdahl J., Schaye J., Teyssier R., Agertz O., 2015, MNRAS, 451, 34
Rosdahl J. et al., 2018, MNRAS, 479, 994
Roy F., Bouillot V. R., Rasera Y., 2014, A&A, 564, A13
Seager S., Sasselov D. D., Scott D., 1999, ApJ, 523, L1
Shapiro P. R., Giroux M. L., 1987, ApJ, 321, L107
Shapiro P. R., Giroux M. L., Babul A., 1994, ApJ, 427, 25
Shukla H., Mellema G., Iliev I. T., Shapiro P. R., 2016, MNRAS, 458, 135
So G. C., Norman M. L., Reynolds D. R., Wise J. H., 2014, ApJ, 789, 149
Sorce J. G., 2015, MNRAS, 450, 2644
Sorce J. G., Tempel E., 2017, MNRAS, 469, 2859
Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014, MNRAS, 437, 3586
Sorce J. G. et al., 2016a, MNRAS, 455, 2078
Sorce J. G., Gottlöber S., Hoffman Y., Yepes G., 2016b, MNRAS, 460, 2015
Sorce J. G., Hoffman Y., Gottlöber S., 2017, MNRAS, 468, 1812
Springel V., 2005, MNRAS, 364, 1105
Sullivan D., Iliev I. T., Dixon K. L., 2018, MNRAS, 473, 38
Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411
Teyssier R., 2002, A&A, 385, 337
Tully R. B., 2015a, AJ, 149, 54
Tully R. B., 2015b, AJ, 149, 171
Tully R. B. et al., 2013, AJ, 146, 86
Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560
Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, ApJ, 825, 144
Wu X., Kannan R., Marinacci F., Vogelsberger M., Hernquist L., 2019, MNRAS, 488, 419
Yang X., Mo H. J., van den Bosch F. C., Bonaca A., Li S., Lu Y., Lu Y., Lu Z., 2013, ApJ, 770, 115
Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446
Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ, 520, 413
Zawada K., Semelin B., Vonlanthen P., Baek S., Revaz Y., 2014, MNRAS, 439, 1615

APPENDIX A: RESOLUTION STUDY ON SUPPRESSION OF STAR FORMATION

We conducted a resolution study to determine the impact of spatial resolution on the suppression of star formation in low-mass haloes due to radiative feedback. To this end we performed a series of three simulations of a test box 4 h^{-1}cMpc on a side, with the full CoDa II physics, using exactly the same parameters as given in Table 1, except for grid cell size and dark matter particle mass, which therefore impact spatial resolution and mass resolution. We explored spatial (mass) resolutions 2 (8) times and 4 (64) times higher than the fiducial CoDa II setup. These simulations are described in Table A1. The average star formation histories of $z = 5$ haloes were then computed for three different bins of dark matter halo mass and are shown in the left-hand panel of Fig. A1. The global shapes of the star formation histories for each mass bin are similar to those of Fig. 9 in the full physics case ('SN, RT'). The highest mass bin of Fig. 9 is missing in this resolution study because it was performed in a 4 h^{-1}cMpc instead of 8 h^{-1}cMpc box, to keep the computational cost of the study in check. This is not a problem, however, since it is in low-mass haloes that suppression is strongest and matters the most. That is the focus of our resolution study.

The left-hand panel of Fig. A1 shows that for the two most massive bins, i.e. between 10^9 and 10^{11} $M_{\odot}$, where the suppression is almost non-existent, increasing spatial resolution by a factor of 2 yields an increase of star formation rate of a factor ∼2. A similar increase is also seen at early times in the 10^{8-9} $M_{\odot}$ halo mass bin, where star formation rises earlier in the high-resolution runs. However, in all 3 runs, the low-mass haloes experience a star formation plateau at a similar level of $\sim 10^{-4}$ $M_{\odot}$ yr^{-1}. Reionization ends for all three simulations at redshifts in the range $z = 6$–7 (shown by the vertical solid, dashed, and dotted lines), and the post-reionization epochs are marked, as in Fig. A1, by a decline of SFR in the low-mass haloes. The decrease is seen at all resolutions, and the timing and slope of the suppression are similar across all three runs.

To account for the slight differences in reionization redshifts between the three runs, we shifted the runs *in time* to align their reionization histories, so that they have the same reionization redshift $z = 6$. The resulting average halo star formation histories and J_{21} are shown in Fig. A2.

With these small time shifts applied to each, the curves of J_{21} lie virtually on top of each other, especially after ∼550 Myr, demonstrating that the 3 runs have very similar levels of irradiation at each stage of their respective reionization histories.

Table A1. Summary table of the simulations of our resolution study, giving the number of particles, the dark matter particle mass, and the spatial resolution in comoving and physical kpc at $z = 5$. All simulations are performed with full CoDa II physics.

Simulation	N_{DM}	M_{DM} ($\mathrm{M}_{\odot}$)	dx (ckpc)	$\mathrm{dx}_{\mathrm{phys}}$ (kpc, $z = 5$)	Comment
L04N0256	256^3	4.07×10^5	23.06	3.84	Fiducial
L04N0512	512^3	6.21×10^4	11.53	1.92	$\mathrm{dx}_{\mathrm{fiducial}}/2$
L04N1024	1024^3	7.77×10^3	5.76	0.96	$\mathrm{dx}_{\mathrm{fiducial}}/4$

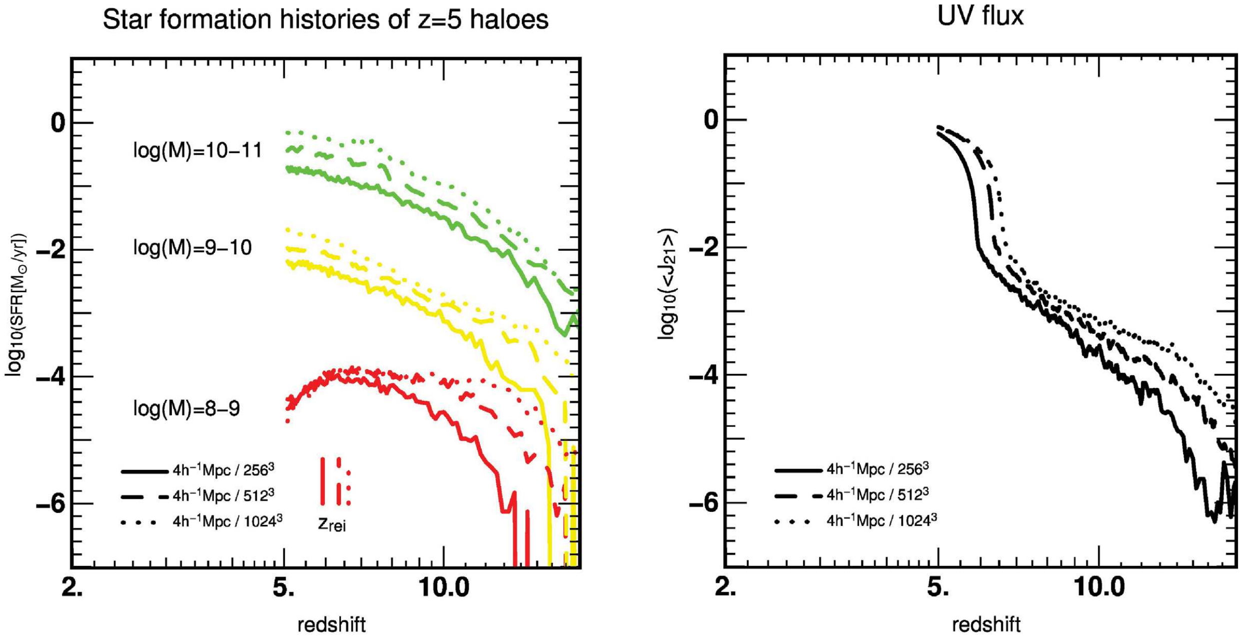


Figure A1. Results of the resolution study simulations with full CoDaII physics listed in Table A1. The test box is 4 h^{-1}cMpc on a side, at different spatial resolutions. *Solid line*: fiducial CoDa II resolution (256^3), *dashed line*: 512^3, *dotted line*: 1024^3. *Left:* Average star formation histories for 3 halo mass bins. The vertical lines above the 'z_{rei}' label indicate the reionization redshift of each simulation. The mass bins correspond to the haloes final mass (i.e. halo mass measured in the final simulation time-step, at $z = 5$. Right: Volume-averaged ionizing UV flux J_{21} for these simulations.

We can now take a closer look at the SFRs of the low-mass bin, as shown in the zoomed left-hand panel of Fig. A2. The slightly earlier onset and rise of SFR at higher resolutions is confirmed, as is the quasi-plateau at an SFR = 10^{-4} $\mathrm{M}_{\odot}$ yr^{-1}, from 600 Myr to $z = 6$, shared by all 3 resolution runs. During this plateau, the SFRs of all 3 runs are very consistent, within 0.1 dex of each other. The same sharp decline of SFR takes place at $z = 6$ for all 3 runs.

We already explored the impact of numerical resolution on star formation suppression in CoDa I in Ocvirk et al. (2016), and found little impact when increasing resolution by a factor 2. The new results of this appendix confirm that, despite the revised subgrid model for star formation for CoDaII, which results in a milder suppression than CoDa I, suppression is still seen in $10^{8–9}$ $\mathrm{M}_{\odot}$ haloes, even at a resolution 4 times higher.

Some other efforts have also been made to address the question of radiative suppression of star formation in low-mass haloes based upon cosmological hydrodynamical simulations with RT. Although their methodology and metrics for measuring suppression may differ from ours, there is reasonable agreement with at least some of them. For instance, Wu et al. (2019) is particularly valuable because of the use of a completely different code, AREPO-RT. They found that photoheating reduces the SFR by 50 per cent in haloes below $10^{8.4}$ $\mathrm{M}_{\odot}$ at $z = 6$ and $10^{8.8}$ $\mathrm{M}_{\odot}$ at $z = 5$. At both redshifts, their suppression mass is within the $10^{8–9}$ $\mathrm{M}_{\odot}$ mass bin where suppression can be seen to take place in Fig. A1.

A similar agreement is found with Pawlik et al. (2015), also using a completely different methodology (smoothed particles hydrodynamics and a form of adaptive ray-tracing). They found a break at the low-mass end of the SFR-halo mass relation, which evolves with redshift. Their simulations reionize earlier than CoDa II, at $z \sim 8$. At this redshift, they find a break in the low-mass end of the SFR-halo mass relation, at 5.10^9 $\mathrm{M}_{\odot}$, characteristic of suppression of star formation by photoheating and supernova feedback. Altough this mass increases with decreasing redshift, the overlap between

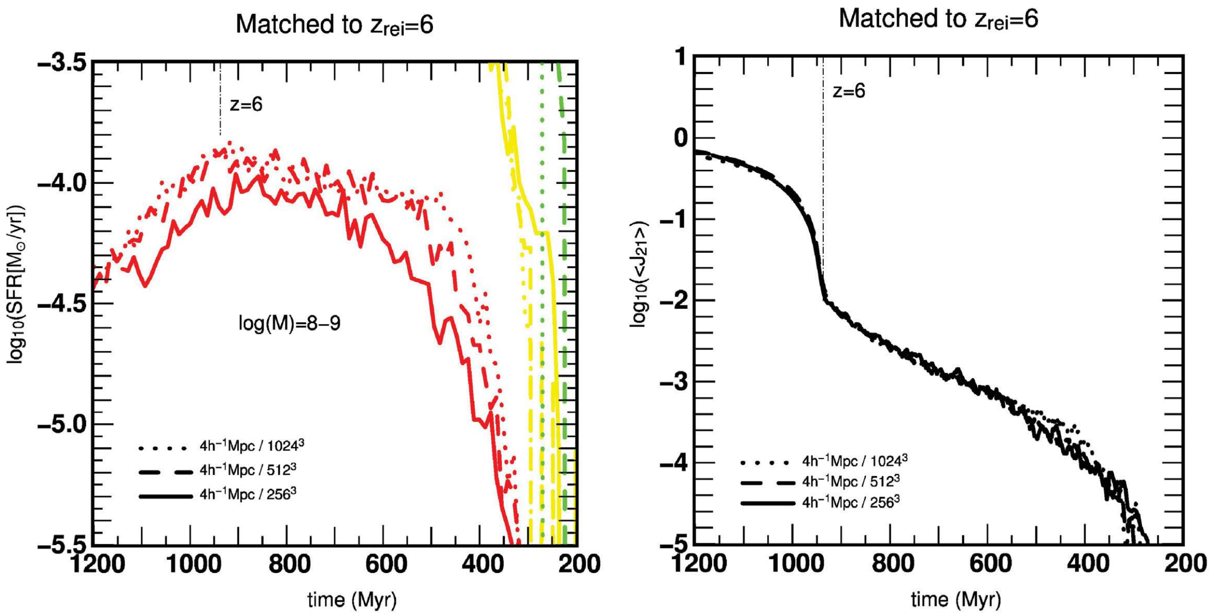


Figure A2. Same as Fig. A1, but the simulations have been shifted in time to align them, so that all 3 runs have $z_{\mathrm{rei}} = 6$. This results in a shift of −17, 66, and 112 Myr for the 256^3, 512^3, and 1024^3 run, respectively. The x-axis now shows time since big bang, with $z = 6$ marked as a vertical dot–dashed line for reference. The left-hand panel is zoomed on the low-mass bin.

the halo mass range where suppression takes place in CoDa II and Pawlik et al. (2015) is encouraging.

The rough agreement between these examples from the literature and our own results, along with the resolution study presented here, suggests that the suppression in star formation seen in CoDa II is not an artefact due to numerical resolution.

APPENDIX B: LUMINOSITY FUNCTIONS

We give the CoDa II luminosity functions at 4 epochs in Table B1, in the hope that they may be of use to the community and in the design, analysis, and interpretation of future surveys.

Table B1. CoDa II UV luminosity functions at 4 different epochs. Each group of 4 columns corresponds to a different redshift. Within each group, the 4 columns give the base-10 logarithm of (1): the median LF, (2): the average LF, (3): the minimum LF, (4): the maximum LF. The LF statistics have been determined using 5 independent sub-boxes each spanning 1/5 of the total CoDa II volume.

M_{AB1600}	$\log_{10}(\Phi[\mathrm{Mag}^{-1}\mathrm{Mpc}^{-3}]$ median (1), average (2), min (3), max (4))															
	$z = 6$				$z = 7$				$z = 8$				$z = 10$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
−24.0																
−23.5		−5.58		−4.88												
−23.0		−5.58		−4.88		−5.58		−4.88								
−22.5	−4.88	−4.80		−4.58		−5.28		−4.88								
−22.0	−4.41	−4.41		−4.11	−4.88	−4.98		−4.58		−5.28		−4.88				
−21.5	−4.18	−4.17	−4.41	−3.93	−4.88	−4.74		−4.41		−5.28		−4.58				
−21.0	−3.74	−3.80	−4.18	−3.65	−4.28	−4.26	−4.58	−4.04	−4.88	−4.74		−4.41				
−20.5	−3.44	−3.47	−3.74	−3.28	−3.98	−4.04	−4.88	−3.80	−4.88	−4.50	−4.88	−4.18				
−20.0	−3.18	−3.21	−3.50	−3.06	−3.60	−3.62	−3.88	−3.50	−4.28	−4.17	−4.41	−3.98		−5.28		−4.58
−19.5	−2.95	−3.00	−3.36	−2.87	−3.38	−3.38	−3.65	−3.25	−3.98	−3.89	−4.11	−3.71	−4.88	−4.88		−4.58
−19.0	−2.67	−2.70	−2.86	−2.55	−3.04	−3.07	−3.38	−2.90	−3.49	−3.50	−3.93	−3.35	−4.41	−4.44	−4.58	−4.41
−18.5	−2.45	−2.50	−2.73	−2.39	−2.74	−2.77	−3.00	−2.67	−3.16	−3.17	−3.39	−3.06	−4.28	−4.24	−4.58	−4.11
−18.0	−2.32	−2.33	−2.51	−2.22	−2.59	−2.61	−2.73	−2.49	−2.85	−2.89	−3.08	−2.81	−3.88	−3.90	−4.11	−3.74
−17.5	−2.13	−2.15	−2.31	−2.07	−2.38	−2.41	−2.60	−2.30	−2.74	−2.73	−2.89	−2.60	−3.71	−3.60	−3.77	−3.44
−17.0	−2.01	−1.98	−2.16	−1.87	−2.23	−2.22	−2.37	−2.11	−2.51	−2.51	−2.64	−2.43	−3.27	−3.25	−3.47	−3.15
−16.5	−1.84	−1.83	−1.98	−1.73	−2.05	−2.05	−2.22	−1.95	−2.31	−2.32	−2.48	−2.20	−3.00	−3.02	−3.30	−2.87
−16.0	−1.68	−1.68	−1.84	−1.59	−1.88	−1.88	−2.03	−1.78	−2.16	−2.13	−2.31	−2.01	−2.81	−2.81	−2.93	−2.72
−15.5	−1.55	−1.53	−1.69	−1.44	−1.74	−1.73	−1.89	−1.62	−1.95	−1.95	−2.11	−1.84	−2.54	−2.57	−2.70	−2.45
−15.0	−1.42	−1.41	−1.55	−1.32	−1.57	−1.57	−1.71	−1.48	−1.81	−1.79	−1.93	−1.69	−2.41	−2.37	−2.56	−2.24

Table B1 – *continued*

M_{AB1600}	$\log_{10}(\Phi[\mathrm{Mag}^{-1}\mathrm{Mpc}^{-3}]$ median (1), average (2), min (3), max (4))															
	$z = 6$				$z = 7$				$z = 8$				$z = 10$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
−14.5	−1.30	−1.30	−1.41	−1.20	−1.45	−1.43	−1.57	−1.34	−1.63	−1.62	−1.79	−1.53	−2.20	−2.16	−2.32	−2.05
−14.0	−1.18	−1.18	−1.30	−1.10	−1.32	−1.31	−1.43	−1.22	−1.50	−1.47	−1.59	−1.38	−2.01	−1.96	−2.11	−1.86
−13.5	−1.08	−1.08	−1.19	−0.99	−1.19	−1.17	−1.29	−1.08	−1.35	−1.34	−1.45	−1.24	−1.82	−1.80	−1.97	−1.68
−13.0	−0.98	−0.98	−1.08	−0.89	−1.07	−1.06	−1.17	−0.97	−1.21	−1.19	−1.31	−1.10	−1.65	−1.62	−1.75	−1.51
−12.5	−0.90	−0.89	−0.98	−0.81	−0.96	−0.94	−1.05	−0.84	−1.09	−1.06	−1.18	−0.96	−1.50	−1.46	−1.58	−1.34
−12.0	−0.84	−0.82	−0.88	−0.74	−0.87	−0.84	−0.93	−0.75	−1.00	−0.95	−1.06	−0.84	−1.39	−1.33	−1.43	−1.20
−11.5	−0.81	−0.78	−0.83	−0.70	−0.81	−0.78	−0.86	−0.67	−0.92	−0.87	−0.96	−0.75	−1.31	−1.24	−1.33	−1.11
−11.0	−0.79	−0.76	−0.80	−0.69	−0.77	−0.72	−0.79	−0.61	−0.88	−0.81	−0.89	−0.68	−1.24	−1.16	−1.26	−1.02
−10.5	−0.88	−0.85	−0.89	−0.77	−0.87	−0.81	−0.88	−0.70	−0.97	−0.90	−0.97	−0.77	−1.32	−1.25	−1.34	−1.09
−10.0	−0.95	−0.93	−0.97	−0.85	−0.95	−0.90	−0.97	−0.78	−1.07	−1.01	−1.10	−0.87	−1.48	−1.41	−1.53	−1.24
−9.5	−0.93	−0.92	−0.97	−0.84	−0.96	−0.91	−1.00	−0.79	−1.10	−1.03	−1.13	−0.88	−1.51	−1.45	−1.58	−1.28
−9.0	−0.97	−0.96	−1.02	−0.86	−1.01	−0.97	−1.07	−0.83	−1.15	−1.10	−1.21	−0.94	−1.63	−1.56	−1.70	−1.38
−8.5	−0.95	−0.94	−1.01	−0.83	−1.00	−0.97	−1.08	−0.83	−1.17	−1.13	−1.26	−0.97	−1.70	−1.63	−1.78	−1.44
−8.0	−0.95	−0.93	−1.01	−0.82	−1.05	−1.02	−1.14	−0.86	−1.22	−1.18	−1.32	−1.02	−1.79	−1.74	−1.92	−1.54
−7.5	−0.97	−0.96	−1.08	−0.83	−1.14	−1.11	−1.25	−0.94	−1.35	−1.32	−1.48	−1.14	−2.01	−1.97	−2.17	−1.76
−7.0	−1.04	−1.03	−1.16	−0.88	−1.26	−1.24	−1.39	−1.07	−1.55	−1.52	−1.71	−1.32	−2.32	−2.28	−2.55	−2.03
−6.5	−1.14	−1.12	−1.28	−0.95	−1.45	−1.42	−1.60	−1.23	−1.80	−1.76	−1.96	−1.55	−2.90	−2.81	−3.05	−2.51
−6.0	−1.16	−1.12	−1.26	−0.95	−1.56	−1.51	−1.66	−1.32	−2.09	−2.03	−2.21	−1.81	−3.84	−3.74	−4.18	−3.47
−5.5	−1.56	−1.49	−1.65	−1.28	−2.50	−2.43	−2.65	−2.18	−3.60	−3.60	−3.84	−3.39				
−5.0	−3.23	−3.08	−3.45	−2.76												
−4.5																
−4.0																
−3.5																
−3.0																

APPENDIX C: STARS BEYOND R_{200}

Following our finding in Section 3.3 that a fraction of the stars are found beyond R_{200}, we provide a few maps to illustrate this aspect and gain insight into the relative distributions of stars and dark matter haloes in CoDa II.

The left column of Fig. C1 shows a gas density map overplotted with the R_{200} spheres of dark matter haloes more massive than $M > 10^8$ M$_\odot$ (black circles), along with stars, within and beyond R_{200}, for 3 typical star-forming regions of the simulation. The slice is 30 cells thick, i.e. about 470 h^{-1}ckpc. The fraction of stars beyond R_{200} increases from top to bottom, and is given in the title of each panel. In most cases, star particles are contained within the R_{200} of their parent haloes (blue circles), and the latter are well aligned along gas filaments, forming strings of haloes. Stars beyond R_{200} are mostly located between pairs of neighbour haloes within these strings. This is confirmed in the bottom plot, which contains a higher density region and more massive haloes, and where the fraction of stars beyond R_{200} reaches 21 per cent, comparable to what we found globally in Section 3.3 at $z = 6$. Here, we show that the fraction of stars beyond R_{200} is primarily the result of approximating FoF halo shapes by a sphere. The right column of Fig. C1 shows the particles of the FoF dark matter haloes of each region. They are clearly not spherical, always triaxial, often elongated in the direction of the filament, and often feature bridge-like structures linking them. Stars beyond R_{200} fall along these aspherical features. At higher redshifts, halo shapes become even more aspherical, accentuating this effect and increasing the number of stars found beyond R_{200}. If one considers more complex halo shapes (using convex hulls for instance) one would expect that the census of stars within haloes would be close to complete. Still a small fraction of stars might be formed in small haloes below the assumed minimum halo mass.

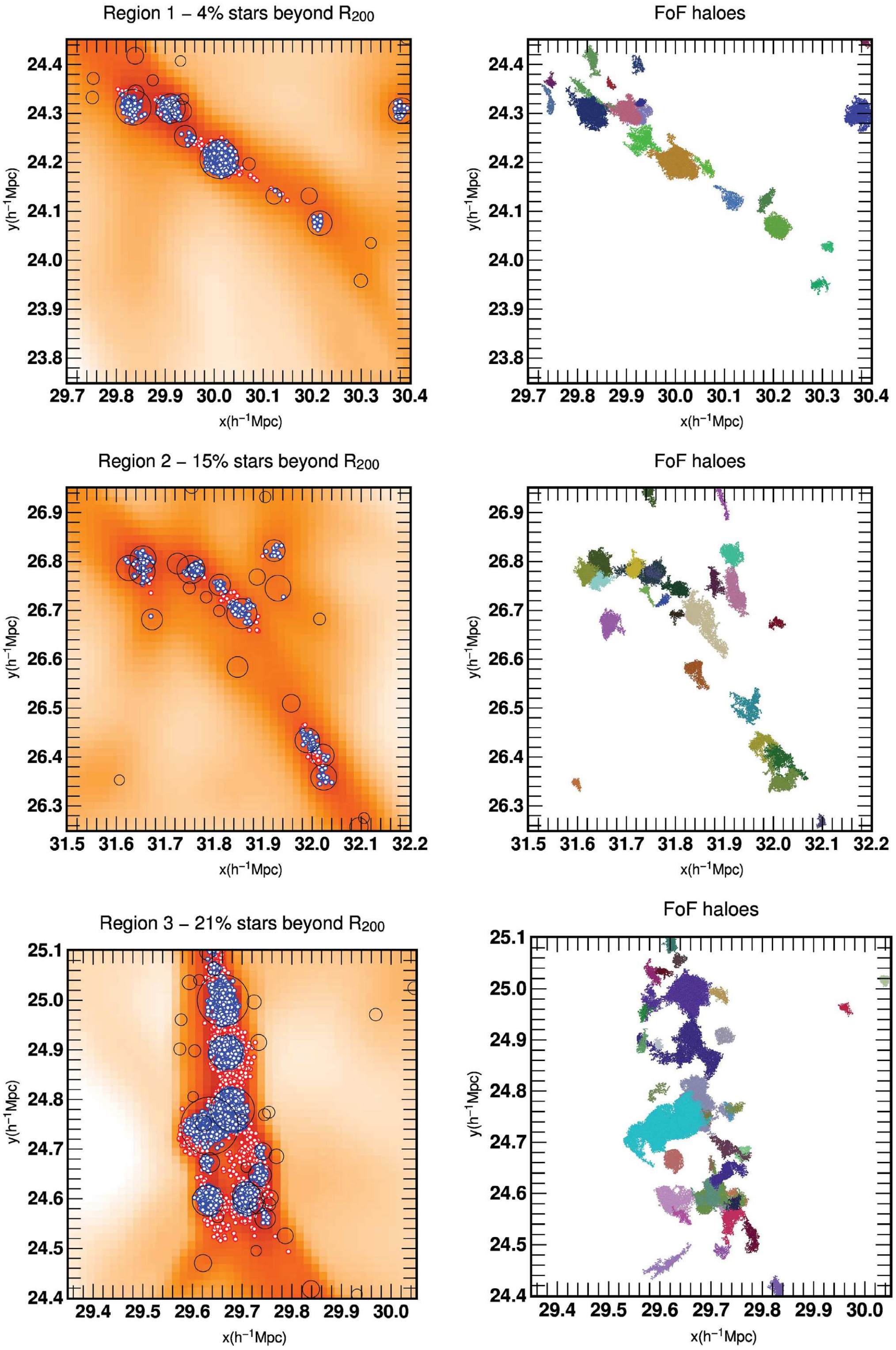


Figure C1. *Left:* Maps of the gas density for three different example regions in CoDa II, overplotted with the R_{200} spheres of dark matter haloes more massive than $M > 10^8$ M$_\odot$ (black circles). Also shown are the stellar particles, in blue if the particle is inside of the R_{200} sphere of a dark matter halo and red otherwise. The percentage of stars beyond R_{200} is indicated in the title of each panel. The most massive halo in the top, middle, and bottom panel is 5.5×10^9, 2.6×10^9, and 1.8×10^{10} M$_\odot$, respectively. *Right:* Maps of the dark matter particles of the FoF haloes of the corresponding left-hand panel. A colour is chosen randomly for each halo.

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

1.2 Galactic ionising photon budget in CoDa II

The following 5 pages give a brief summary of my first accepted paper as first author. The reader can use this section to avoid reading the full paper, although the analysis is much more detailed in the latter.

1.2.1 Why study the photon budget in CoDa II?

Investigating the ionising photon budget of galaxies could lead to insight into the way Reionization unfolds, and is linked to many essential questions that are key to understanding the formation and evolution of the first galaxies, as well as observations of Reionization. The large box size (94.44 Mpc) of the fully coupled RHD cosmological CoDa II simulation, and its massive sample of galaxies, will prove crucial for investigating the link between galactic evolution and properties, and f_{esc}, whilst resolving the least massive star forming galaxies. Further, using a reduced speed of light can affect ionisation front speeds [Deparis et al., 2019], and has been shown to alter the post overlap average ionisation in [Ocvirk et al., 2019], and may affect the f_{esc} and the photon budget by impacting ionisation rates, and the local UVBG. The use of c=1 in CoDa II bypasses these potential caveats. To determine the galactic ionising photon budget, I first computed the escape of ionising radiation from star forming haloes to the IGM, and their ionising photon production.

1.2.2 Escape fractions

Computation

To compute the fraction of ionising photons that reach the IGM (f_{esc} or since we use ray tracing here : f_{esc}^{ray}), I used the gas ionisation and density fields to compute the angular average LyC transmission along 768 rays from each halo star forming cell to the halo's r_{200}[6]. The final f_{esc}^{ray} for each halo is then obtained by performing the average of each star forming cell's angular average, weighted by the photon production in the cell (Lewis et al. [2020] contains more details). We elaborate on why we picked this method (as opposed to measuring f_{esc} directly using the flux fields of CoDa II) in Sec. Computation by integrating the photon flux.

Average trends

Fig. 1.8 shows the average trends of f_{esc}^{ray} with mass and with redshift. The f_{esc}^{ray} of the lowest mass star forming galaxies is close to unity, more than doubling from z=14.9 till z=6, when the post Reionization overlap of ionised IGM regions occurs. Lewis et al. [2020] show that before the Universe is reionised at z=6, low mass haloes that have just formed their first stellar particle have lower f_{esc}^{ray} values than those that are older. This partly explains the strong evolution of the average f_{esc}^{ray} of low mass haloes with redshift. On average, the f_{esc}^{ray} of haloes decreases with increasing mass, particularly for haloes more massive than $\approx 10^9\,M_\odot$. The slope of this trend remains relatively constant throughout Reionization. This means that at all times, the most massive haloes, which are also the brightest as they are the

---

[6] Defined in Sec. Ionising photon budget of galaxies.

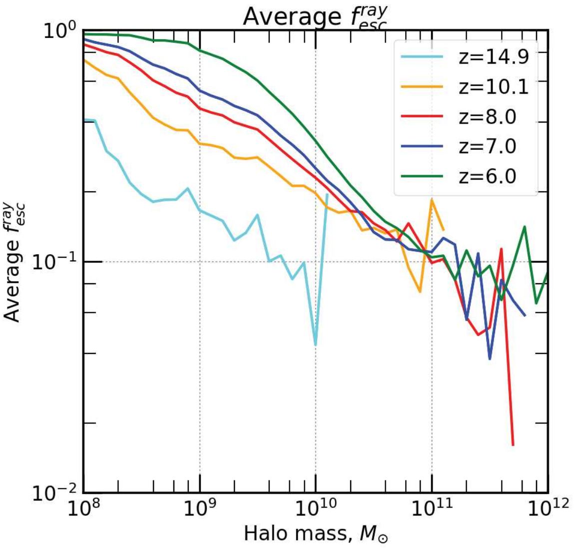


Figure 1.8: Average LyC escape fractions as a function of halo mass for several redshifts.

most efficient star formers, have the lowest escape fractions. The average f_{esc} of the highest mass haloes (>4 × $10^{10}\,M_\odot$) appears to be constant with redshift. However, it is difficult to interpret the averages of the haloes with masses >$10^{11}\,M_\odot$ as they are few, and the resulting average is very noisy.

Explaining the low escape fractions in high mass haloes

Fig. 1.9 shows the relative contribution of cells to the total SFR per mass bin, and rate of photons escaping to the IGM (L_{esc}) per mass bin, as a function of cell transparency to LyC photons at z=6. In the low mass haloes (red curves), a large fraction (>0.5) of star forming cells is transparent. The higher the halo mass, the greater the fraction of opaque (T<0.1) star forming cells is. In the highest mass bin, more than 80% of star forming cells have transmissions less than 0.1. The right panel shows that all the star forming cells that contribute significantly to the total escaping luminosity have LyC cell transmissions greater than 0.1 in the low mass bin. In the higher mass bins, most of the contribution to L_{esc} comes from cells that are more transmissive than 0.01. In combination with the left panel, this shows that the cells that contain about 80% of star formation in the highest mass bin, contribute close to no photons to the IGM. This explains the low f_{esc}^{ray} values found in the higher mass haloes, as well as the decrease in average f_{esc}^{ray} with mass. f_{esc}^{ray} is set by the capacity of higher mass haloes to recombine in their densest regions, hiding a fraction of their star formation in neutral UV opaque cells. Lewis et al. [2020] show that the discussed low LyC transmissions are primarily driven by high neutral fractions. This is because as haloes' masses increase, their ability to recombine increases more than their star formation rate[7], and the fraction of stars in UV opaque cells increases. In fact, at z=6 more than half of the star formation of the most massive (<$10^{11}\,M_\odot$) haloes occurs within cells that have neutral fractions (x_{HI})>63%, whereas for the haloes within 10^9 to $10^{10}\,M_\odot$, no such neutral star

---

[7] Within the mass range of galaxies that we consider, as halo mass increases galaxies accrete more gas thereby increasing the gas density. At the same time, the rate of recombination is $\propto \rho_{gas}^2$, whereas the rate of star formation is $\propto \rho_{gas}^{1.5}$ by definition in RAMSES-CUDATION. Hence, the recombination rate scales faster with halo mass than the star formation rate does.

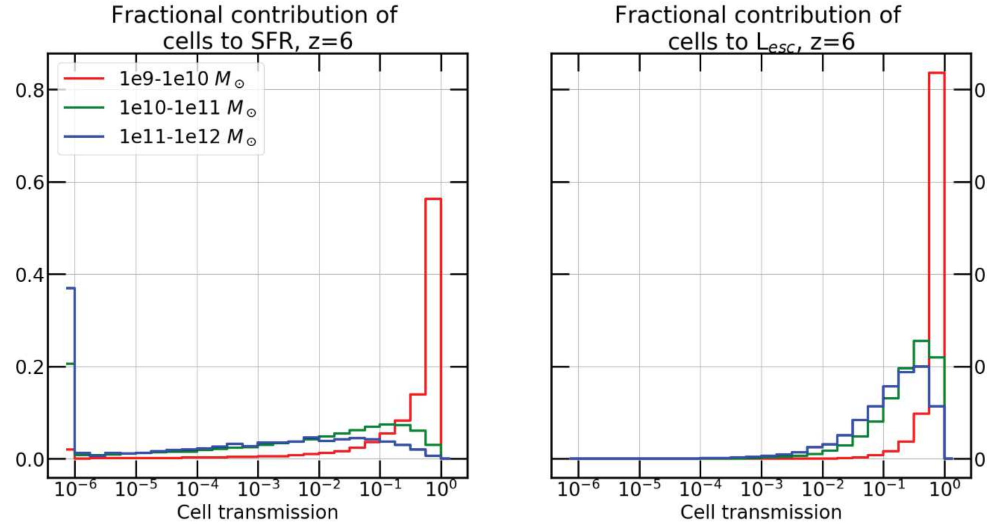


Figure 1.9: *Left* : Fractional contribution of halo star forming cells to total star formation, as a function of their LyC transmission. The distribution is split into 3 bins according to the masses of the haloes that contain the cells. *Right*: Same as on the left, but we measure the contribution to the total escaping ionising luminosity instead of total SFR.

forming cells exist.

1.2.3 Escaping luminosities

Fig. 1.10 shows the average halo L_{esc} as a function of halo mass for several redshifts. Both the full sample average (dotted lines) and the star forming sample average (full lines) show that on average, L_{esc} is an increasing function of halo mass, despite the fact that f_{esc}^{ray} decreases over the same period. This implies that the star formation efficiency increases faster with mass than f_{esc}^{ray} decreases. In the star forming sample, the imprint of the stellar mass resolution of CoDa II is apparent in the plateau that forms until $\approx 10^9\,M_\odot$. Above this mass, both samples have very similar average behaviours. Below, the full sample averages are considerably lower than the star forming averages as they include SFR=0(L_{intr},L_{esc}) haloes. In fact, because of this the imprint of star formation suppression is visible in these curves, that decrease markedly towards the end of Reionization (z≈6) and overlap.

1.2.4 Ionising photon budget of galaxies

Fig. 1.11 shows the escaping ionising photon budget in CoDa II. The photon budget is dominated by haloes from 6×10^8 to 3×10^{10} $M_\odot$[8] during Reionization (they produce ≈ 80% of the photons that reach the IGM and contribute to Reionization at z=7, when the Universe is undergoing very rapid Reionization) . The lesser importance of the least massive haloes in spite of their great numbers is due to their inefficient star formation, and its suppression as Reionization progresses (as shown by Ocvirk et al. [2020]). Even though the average halo luminosity escaping to the IGM is a monotonically increasing

---

[8] This corresponds to a M_{AB1600} range of -12 to -19

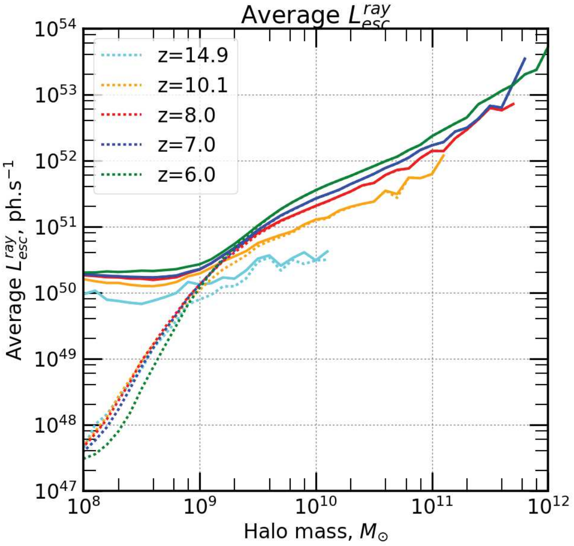


Figure 1.10: Average halo escaping luminosity as a function of halo mass. Full lines show the average for the star forming sub-sample of haloes. Dotted lines show the average for the full sample. Note that the lowest transmission bins also show the fraction of cells below the lowest transmission bin.

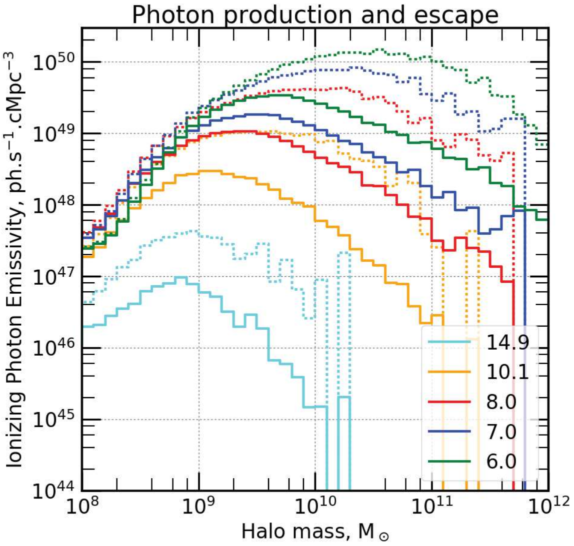


Figure 1.11: Ionising galactic photon budget (full lines), and the intrinsic photon production (dotted lines) in $\mathrm{ph} \cdot \mathrm{s}^{-1} \cdot \mathrm{cMpc}^{-3}$.

function of halo mass, the most massive haloes do not contribute significantly to Reionization, due to their low number densities and their low f_{esc}. Their small f_{esc} are reflected in the drastic difference between the intrinsic photon production curves (dotted) and the photon budget curves (full lines) of high mass halos. As time and Reionization advance, the mass range of haloes that dominate the photon budget shifts towards the more massive, brighter galaxies. This is simply attributable to a build up in their numbers, and to the increase of the suppression of star formation in low mass galaxies.

1.2.5 Lewis et al. 2020

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **496,** 4342–4357 (2020)
Advance Access publication 2020 June 18

doi:10.1093/mnras/staa1748

Galactic ionizing photon budget during the epoch of reionization in the Cosmic Dawn II simulation

Joseph S. W. Lewis ,[1★] Pierre Ocvirk,[1] Dominique Aubert,[1] Jenny G. Sorce,[2,3,4] Paul R. Shapiro,[5] Nicolas Deparis,[1] Taha Dawoodbhoy,[5] Romain Teyssier,[6] Gustavo Yepes,[7,8] Stefan Gottlöber,[4] Kyungjin Ahn,[9] Ilian T. Iliev [10] and Jonathan Chardin[1]

[1]*Observatoire Astronomique de Strasbourg, Université de Strasbourg, CNRS UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France*
[2]*Univ. Lyon, ENS de Lyon, Univ. Lyon I, CNRS, Centre de Recherche Astrophysique de Lyon, UMR5574, F-69007 Lyon, France*
[3]*Univ. Lyon, Univ. Lyon I, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon, UMR5574, F-69230 Saint-Genis-Laval, France*
[4]*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[5]*Department of Astronomy, University Texas, Austin, TX 78712-1083, USA*
[6]*Institute for Theoretical Physics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[7]*Departamento de Física Teórica M-8, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[8]*Centro de Investigación Avanzada en Física Fundamental (CIAFF), Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[9]*Chosun University, 375 Seosuk-dong, Dong-gu, Gwangjiu 501-759, Korea*
[10]*Astronomy Center, Department of Physics & Astronomy, Pevensey II Building, University of Sussex, Falmer, Brighton BN1 9QH, United Kingdom*

Accepted 2020 June 11. Received 2020 June 9; in original form 2020 January 17

ABSTRACT

Cosmic Dawn II yields the first statistically meaningful determination of the relative contribution to reionization by galaxies of different halo mass, from a fully coupled radiation-hydrodynamics simulation of the epoch of reionization large enough ($\sim$100 Mpc) to model global reionization while resolving the formation of all galactic haloes above $\sim 10^8\,\mathrm{M}_\odot$. Cell transmission inside haloes is bi-modal – ionized cells are transparent, while neutral cells absorb the photons their stars produce – and the halo escape fraction $f_{\rm esc}$ reflects the balance of star formation rate (SFR) between these modes. The latter is increasingly prevalent at higher halo mass, driving down $f_{\rm esc}$ (we provide analytical fits to our results), whereas halo escape luminosity, proportional to $f_{\rm esc}$ $\times$ SFR, increases with mass. Haloes with dark matter masses within $6 \times 10^8\,\mathrm{M}_\odot < M_{\rm halo} < 3 \times 10^{10}\,\mathrm{M}_\odot$ produce $\sim$80 per cent of the escaping photons at $z = 7$, when the universe is 50 per cent ionized, making them the main drivers of cosmic reionization. Less massive haloes, though more numerous, have low SFRs and contribute less than 10 per cent of the photon budget then, despite their high $f_{\rm esc}$. High-mass haloes are too few and too opaque, contributing $<$10 per cent despite their high SFRs. The dominant mass range is lower (higher) at higher (lower) redshift, as mass function and reionization advance together (e.g. at $z = 8.5$, $x_{\rm H\,I} = 0.9$, $M_{\rm halo} < 5 \times 10^9\,\mathrm{M}_\odot$ haloes contributed $\sim$80 per cent). Galaxies with UV magnitudes $M_{\rm AB1600}$ between -12 and -19 dominated reionization between $z = 6$ and 8.

Key words: Galaxy: formation – galaxies: high-redshift – dark ages, reionization, first stars.

1 INTRODUCTION

Current observations are consistent with the hypothesis that the Universe was reionized when UV starlight from massive stars escaped from the early galaxies in which they formed, creating intergalactic H II regions that grew in size and number until they overlapped to fully ionize the intergalactic medium (hereafter IGM) by $z = 6$. [For reviews and references, see, for instance, Dayal & Ferrara 2018 and Barkana & Loeb 2007]. During the epoch of reionization (EOR), the globally averaged ionized fraction was equivalent to the volume filling factor of these H II regions, which increased monotonically in an evolving patchwork of fully ionized and fully neutral zones. How fast this volume filling factor grew was determined primarily by the average balance between the rate of release of ionizing photons by galaxies and the recombination rate of H atoms in their surrounding IGM. The release rate in a given patch depended upon the galaxy formation rate there, the star formation rates (hereafter 'SFRs') inside each galaxy, the spectra and luminosities of those stars, and the galactic escape fractions $f_{\rm esc}$ of their ionizing photons. The recombination rate in the surrounding IGM depended upon its evolving inhomogeneous density field. All these ingredients varied in space and time in a complex way. First, since structure formation was inhomogeneous. Secondly, because Reionization and the energy release associated with the star formation that drives it exerted hydrodynamical feedback therefore the ingredients were interdependent.

To predict their coupled evolution in the context of ΛCDM cosmology, to test the latter against observations, we must model the gravitational and gas dynamics of dark and baryonic matter and

---

★ E-mail: j.lewis@astro.unistra.fr

Published by Oxford University Press on behalf of the Royal Astronomical Society

the radiative transfer of ionizing radiation in and between galaxies as they form. To capture the large-scale structure of inhomogeneous reionization, this must be done in a representative volume large enough ($\sim$100 Mpc), and with enough resolving power to form all the galaxies in that volume which contribute to reionization. As the dominant contributors are thought to be the 'atomic-cooling haloes' (hereafter 'ACHs') – those of virial temperatures above 10^4 K and masses above $10^8\,\mathrm{M}_\odot$ – which means we must be able to resolve the formation of all the millions of haloes above $10^8\,\mathrm{M}_\odot$ in that large volume.[1] In principle, to capture the full details of star formation within each galaxy, we would also have to resolve the interstellar medium (ISM) of each galaxy down to the sub-parsec scale on which molecular clouds fragment into protostars that then collapse into stars. The latter is currently out-of-reach computationally, however, even in the highest resolution simulations to date of a single galaxy. This means star formation and its local energy release must generally be treated as a 'sub-grid' process.

We have developed the Cosmic Dawn (hereafter 'CoDa') Project, to simulate reionization and galaxy formation together, self-consistently, with fully coupled, radiation-hydrodynamics, on a large-enough scale and with sufficient mass resolution to satisfy these requirements (Ocvirk et al. 2016; Aubert et al. 2018; Dawoodbhoy et al. 2018; Ocvirk et al. 2020). CoDa I (91 Mpc box), described in Ocvirk et al. (2016) and Dawoodbhoy et al. (2018), and CoDa II (94.5 Mpc box), described in Ocvirk et al. (2020), both used the massively parallel, hybrid CPU-GPU code RAMSES-CUDATON on a uniform grid of 4096^3 cells for the baryons and the radiation field, with 4096^3 N-body particles for the dark matter. CoDa I-AMR (91 Mpc), on the other hand, used another massively parallel, hybrid CPU-GPU code EMMA (Aubert, Deparis & Ocvirk 2015), with adaptive mesh refinement (AMR), with 2048^3 particles on a grid of 2048^3 coarse cells from which AMR increased the resolution locally, by up to a factor of 8, to follow the increasing local overdensity, leading to 18 billion cells after refinement. All three CoDa simulations were in volumes large enough to model the inhomogeneity and globally averaged time history of reionization, while also serving to model reionization and galaxy formation in the Local Universe, by adopting 'constrained realizations' of the Gaussian random noise initial conditions that were derived from galaxy survey data so as to reproduce the familiar structures of the Local Universe, such as the MW, M31, and the Virgo cluster, when evolved forward to $z = 0$ (Sorce et al. 2016). We refer the reader to the papers cited above to describe our CoDa simulations and their relative differences in more detail. Our purpose here is to use the most recent of them, CoDa II, to find the ionizing luminosities of all the galaxies that formed in it during the EOR, to make the first statistically meaningful determination of the relative contribution to reionization by galaxies of different halo mass, over the full range of masses that contribute significantly, in a fully coupled radiation hydrodynamical numerical simulation.

The escape fraction f_esc of galaxies is difficult to observe directly. Indeed, the individual galaxies must be bright enough to detect, but also to compare their fluxes and spectral information at different wavelengths (above and below the H Lyman limit). This must then be interpreted in terms of a model in which stars are assumed to have some initial mass function (hereafter 'IMF') and an SFR, which determines their spectral energy distribution (hereafter 'SED') over time. The radiation that emerges from the galaxy at wavelengths longwards of the Lyman limit is then assumed to be a combination of this SED and the nebular emission that results from re-processing the absorbed fraction of ionizing starlight by the interstellar gas, and may be partially attenuated by internal dust. Starlight emitted bluewards of the H Lyman limit is attenuated by photoionizing H atoms in the ISM of the galaxy and possibly attenuated further by dust. These processes are reflected in the net absorbed fraction $(1 - f_\mathrm{esc})$, which may also include attenuation by the bound-free opacity of foreground Lyman limit absorbers along the line of sight. Observations of galaxies at different redshifts face different challenges, as they involve different spectral regions depending on z. Moreover, the foreground opacity is also a strong function of increasing redshift. As a consequence, observational determinations of f_esc are few and still uncertain. A review of this subject is well beyond the scope of this paper; the reader is referred to, e.g. Izotov et al. (2016), and the review by Dayal & Ferrara (2018) and references therein for a summary.

On the theory side, results are also rather limited. Some attempt to derive an empirical f_esc, one-size-fits-all. For an assumed form and amplitude of the UV luminosity function of galaxies above redshift 6, they proceed by adjusting f_esc to release enough ionizing photons to finish reionization in time to satisfy various observational constraints (we refer to these efforts as 'one-zone' models, e.g. Robertson et al. 2015). In doing so values such as $f_\mathrm{esc} = 10$ or 20 per cent are sometimes reported, but this depends strongly on the underlying assumptions that led to it. These models often requires some redshift evolution (for instance, in Haardt & Salvaterra 2015; Puchwein et al. 2019). Other attempts to determine a global value for f_esc employ semi-analytical or semi-analytical models of reionization. Again, these are similarly adjusted to match observational constraints, but use the model's own statistical determinations of the rate of formation of galactic haloes from cosmological initial conditions, and some assumption about the SFR per halo (e.g. Ferrara & Loeb 2013; Dayal et al. 2020). A variety of galaxy formation simulations also exist that attempt to predict the f_esc and SFR from their simulated galaxies. However, these are mostly without radiative transfer or only post-processed with radiative transfer (such as in Razoumov & Sommer-Larsen 2010; Yajima, Choi & Nagamine 2011; Paardekooper, Khochfar & Dalla Vecchia 2015; Ma et al. 2015; Anderson et al. 2017). Recent studies do account for fully coupled radiation-hydrodynamics, with a focus on spatial resolution, considering a single galaxy or a fraction thereof (Kimm & Cen 2014; Trebitsch et al. 2017, 2018; Kimm et al. 2019; Trebitsch et al. 2020; Yoo, Kimm & Rosdahl 2020). This focus on resolving internal galactic structure precludes the follow-up of ionizing radiation propagation at cosmological scales.

A review of this subject, too, is well beyond the scope of this paper; the reader is again referred to Dayal & Ferrara (2018) for a summary and references, but we will describe some of these results in what follows, as we compare with our own.

The relative contribution of different mass haloes to the total ionizing photon budget released into the IGM during the EOR depends, not only upon the values of f_esc for each galaxy, but on their SFRs and and the evolution of their population, as well. These aspects combine to determine the ionizing luminosity function of galaxies. In this work, we shall investigate this galactic ionizing photon budget. Since star formation efficiency typically rises with halo mass within the range of masses we represent (not only in

---

[1] Although the first stars form in mini-haloes (hereafter 'MHs'; i.e. those with halo mass $\mathrm{M} < 10^8\,\mathrm{M}_\odot$ and virial temperatures $\mathrm{T} < 10^4$ K) that can cool gas by H_2 molecular cooling (some stars may even form in metal-cooling MHs, as shown in Wise et al. 2014), the rising UV background during the EOR limits their contribution to the earliest stages of reionization. As a result, their relative contribution when compared to the more massive ACHs appears small ($\sim$10–20 per cent), (Ahn et al. 2012; Kimm et al. 2017).

CoDa II, but as generally expected, see Moster, Naab & White 2013; Legrand et al. 2019), while the abundance of haloes decreases with halo mass (e.g. see the halo mass functions of Sheth, Mo & Tormen 2001; Watson et al. 2013), the mass range of contributing haloes may, in principle, be broad, with a maximum contribution from haloes that, at different redshifts, may be anywhere within the broad range $10^8\,\mathrm{M}_\odot < \mathrm{M}_\mathrm{halo} < 10^{12}\,\mathrm{M}_\odot$.

Previous work on the role of simulated galaxies in ionizing the IGM has sometimes been difficult to reconcile. However, it seems that with the recent advent of higher resolution and of fully coupled simulations, a few elements of consensus have begun to emerge. Studies such as Anderson et al. (2017) and Yajima et al. (2011) find that reionization is driven by the more numerous, low-mass galaxies ($M_\mathrm{halo} < 10^{9.5}\,\mathrm{M}_\odot$), which broadly agrees with the conclusions of Kimm & Cen (2014), who find that the photon budget is dominated by masses $10^8\,\mathrm{M}_\odot < \mathrm{M}_\mathrm{halo} < 10^9\,\mathrm{M}_\odot$ before $z = 8$, after which more massive haloes take over. Similarly, a more recent effort by Katz et al. (2018) seems to favour haloes within the range $10^9\,\mathrm{M}_\odot < \mathrm{M}_\mathrm{halo} < 10^{10}\,\mathrm{M}_\odot$ during the EOR, and those of higher mass at $z = 6$.

Most of the previous simulations are in volumes that are not large enough (most are boxes smaller than 25 Mpc across) to fully represent the halo mass function above $\sim 10^{10}\,\mathrm{M}_\odot$. They may therefore be missing some of the largest haloes and galaxies, the ones, in fact, that form the most stars. The contributions of these highest mass haloes ($>10^{10}\,\mathrm{M}_\odot$) to the photon budget in these studies is therefore partially absent. This could have a further, profound effect, on the rate LyC photons are released from the lowest mass galaxies, the ones that are reionized and suppressed by external sources, as well as dramatically alter the geometry of ionized regions throughout the EOR. Moreover, when reionization is simulated in too small a box, the duration of reionization is too small compared with that found in a volume large enough to capture the globally averaged history. In turn, this can affect the relative importance of haloes of different mass as their relative abundances evolve with redshift (Iliev et al. 2006, 2014), alongside their importance to reionization.

To overcome these limitations, we will address the photon budget of galaxies during the EOR using the CoDa II simulation (Ocvirk et al. 2020), produced with the RAMSES-CUDATON code (Ocvirk et al. 2016), which couples RAMSES (Teyssier 2002), the code for baryonic hydrodynamics and dark matter N-body dynamics, to ATON (Aubert & Teyssier 2008), the code for radiative transfer of ionizing radiation and non-equilibrium ionization rate equations. The CoDa II simulation ran from $z = 150$ to $z = 5.8$ in a comoving cubic box 94.533 Mpc on a side, with a high-enough mass resolution to form every galaxy in that volume with halo mass above $10^8\,\mathrm{M}_\odot$. This is sufficient to satisfy the requirement for large enough volume to simulate the EOR and its inhomogeneity, with a statistically meaningful halo mass function over the full mass range that may contribute significantly to reionization. Further featuring the fully coupled radiation-hydrodynamics (including radiation transport at the full speed of light) necessary to study the release of ionizing starlight into the IGM by galaxies during the EOR and its transport between galaxies involving highly supersonic ionization fronts. This combination of very large volume with complete sampling of the galactic sources within it represents a necessary compromise. The focus on large scales comes at a cost; we do not attempt to achieve the higher resolution inside galaxies that some other recent simulations do (e.g.Trebitsch et al. 2017; Katz et al. 2018, 2019; Rosdahl et al. 2018; Kimm et al. 2019; Trebitsch et al. 2020), which may affect some of the internal halo physics that are important for our problem, such as the escape of ionizing photons. One of the goals of this study, however, is to demonstrate that, despite our relatively coarser spatial resolution internal to individual galaxies in CoDa II , the 'global' halo quantities relevant for describing the radiative properties of high-redshift galaxies, such as their escape fraction and total escape luminosity in ionizing photons, are meaningful and well captured, thereby validating the CoDa II-like approach, and paving the way towards even larger numerical simulations of the EOR with it in the future.

In this paper, we first, in Section 2, outline our numerical approach and computations. Then, in Section 3, we lay out our escape fraction results. Then, in Section 4, we present the ionizing galactic photon budget. Finally, in Section 5, we summarize our findings and propose some directions in which to take our subsequent efforts.

2 METHODS

2.1 Cosmic Dawn simulations

CoDa I and CoDa II are the largest coupled radiation hydrodynamics cosmological grid-based simulations aimed at studying the EoR. In the simulation code RAMSES-CUDATON, the RAMSES hydrodynamics+N-body code (Teyssier 2002) and the ATON radiative transfer code (Levermore 1984; Aubert & Teyssier 2008) are coupled, forming a hybrid code : hydrodynamics, gravitation, star formation, and supernova (SN) feedback are managed by the central processing units (CPUs), while the more computationally intensive mono-group[2] radiative transfer, Hydrogen photochemistry and cooling are managed by the graphics processing units (GPUs). The resulting acceleration allows us to perform simulations using the full speed of light, thereby circumventing possible artefacts arising from the use of a reduced speed of light (see Gnedin 2016, Deparis et al. 2019, Ocvirk et al. 2019, and Wu et al. 2019 for details on the impact of reduced or variable speed of light approximations).

We do not account for chemical enrichment, nor dust. The simulation focuses on the stellar component's ability to drive cosmic reionization, and therefore active galactic nuclei (AGN) formation, feedback, and ionizing emissivity are not taken into account.

The large box size (94.44 cMpc box) and relatively high resolution for a simulation of cosmic reionization (4096^3 dark matter particles and 4096^3 cells) yield a comoving cell size of 23.06 ckpc (i.e. 3.3 kpc physical at $z = 6$) and a dark matter (stellar) particle mass of $4.07 \times 10^5\,\mathrm{M}_\odot$ ($11\,732\,\mathrm{M}_\odot$). With these specifications, CoDa II can represent the large-scale spatial variance in the reionization process whilst self-consistently dealing with the physics of the haloes that interests us (i.e. those within $10^8\,\mathrm{M}_\odot \lesssim M_\mathrm{halo} \lesssim 10^{12}\,\mathrm{M}_\odot$), and providing us a huge sample of galactic haloes: there are around 13 million dark matter haloes at the end of the EoR in CoDa II.

CoDa II is compatible with a number of observational constraints related to the EoR, most notably the reionization history of the Universe, the Thomson optical depth measured from the cosmic microwave background, and the UV luminosity function of galaxies, as shown in Ocvirk et al. (2020). For further information relating to the code's design, set-up, and runs, we refer the reader to Ocvirk et al. (2020).

2.2 Halo detection and boundaries

Dark matter haloes (haloes throughout the text) are detected using a Friends-of-Friends algorithm described in Roy, Bouillot & Rasera (2014), which produces a catalogue of haloes with their positions

---

[2] Effective photon energy 20.28 eV.

and masses M_{halo}. We can define a halo's virial radius, based on it's mass M_{halo}, as r_{200}, as in Ocvirk et al. (2016, 2020):

$$M_{\mathrm{halo}} = \frac{4}{3}\pi r_{200}^3 \times 200\, \bar{\rho}_{\mathrm{DM}}\,, \tag{1}$$

where $\bar{\rho}_{\mathrm{DM}}$ is the average cosmic dark matter density.

As in Ocvirk et al. (2016, 2020), we assume that one galaxy resides in each dark matter halo, and that the r_{200} limit of the halo is the limit of the galaxy. This assumption is valid in the vast majority of cases, in which the star-forming region of each halo has one clear stellar mass peak within r_{200}.

2.3 Halo escape fraction: ray-tracing, sub-grid, and net

Using CoDa II's gas density and ionization fields, we can compute the optical depths encountered by photons emitted in the halo along paths, from their injection to r_{200}. For a given halo, and for a given halo cell, we use the *healpix* PYTHON module *healpy* to sample the r_{200} sphere with 768 evenly distributed end points (we pick this number so as to adequately resolve our largest haloes. Increasing the number of rays by a factor of 2 only yields a difference of the order of 10^{-4} in $f_{\mathrm{esc}}^{\mathrm{ray}}$ for the most error susceptible computation). We then compute the optical depth along the path from the source cell centre to each end point as

$$\tau_{\mathrm{path}} = \int_{\mathrm{path}} \sigma_{\mathrm{E}} \times \rho_{\mathrm{H\,I}} \times \mathrm{dl}\,, \tag{2}$$

where $\sigma_E = 2.493 \times 10^{-22}\mathrm{m}^2$ is the effective Hydrogen ionization cross-section of the photon group considered here in CoDa II, $\rho_{\mathrm{H\,I}}$ is the neutral Hydrogen physical density of the cell, and dl is an element of length.

The fraction of photons reaching r_{200} from a cell, Tr_{200}, is then obtained as the average of the transmissions along all 768 paths connecting the cell to the r_{200} sphere: $\mathrm{Tr}_{200} = \langle \exp(-\tau_{\mathrm{path}})\rangle_{\mathrm{paths}}$ where the bracket denotes the average over all 768 paths. Fig. C1 is provided as an explanatory figure in Appendix C.

Finally, the halo escape fraction $f_{\mathrm{esc}}^{\mathrm{ray}}$ is obtained as the SFR-weighted average of the transmissions of all the emitting cells of the halo, i.e. for a halo containing N cells indexed by the integer i:

$$f_{\mathrm{esc}}^{\mathrm{ray}} = \frac{\sum_{\mathrm{i}} \mathrm{SFR_i Tr_{200,i}}}{\sum_{\mathrm{i}} \mathrm{SFR_i}}. \tag{3}$$

The SFR of a cell is obtained as the stellar mass formed in the last 10 Myr, $M_{stars}(< 10\,\mathrm{Myr})$ divided by a 10 Myr duration:

$$\mathrm{SFR} = \mathrm{M_{stars}}(< 10\,\mathrm{Myr})/10^7\ \mathrm{M_\odot\,yr^{-1}}, \tag{4}$$

In the rest of the paper, 'star forming' (halo or cell) always means that some stellar mass has been formed in the last 10 Myr, i.e. $M_{\mathrm{stars}}(< 10\,\mathrm{Myr}) > 0$. The 'ray' superscript is used to clarify at all times that $f_{\mathrm{esc}}^{\mathrm{ray}}$ is obtained via ray-tracing. We will refrain from discussing the interstellar medium (ISM) or circum-galactic medium (CGM) nature of the absorbing material. Since these are difficult to separate in CoDa II, we prefer the more general term 'halo escape fraction' for $f_{\mathrm{esc}}^{\mathrm{ray}}$.

Furthermore, since CoDa II does not resolve the ISM of our galaxies, CoDa II also uses a sub-grid escape fraction $f_{\mathrm{esc}}^{\mathrm{sub}} = 0.42$ (chosen To obtain an EoR ending around z = 6), which accounts for the photons lost to the star's birth cloud and the ISM, i.e. only 0.42 of the ionizing photons produced by a stellar particle is deposited in the cell containing it. We can now also define the net halo escape fraction $f_{\mathrm{esc}}^{\mathrm{net}}$ as

$$f_{\mathrm{esc}}^{\mathrm{net}} = f_{\mathrm{esc}}^{\mathrm{sub}} \times f_{\mathrm{esc}}^{\mathrm{ray}}\,. \tag{5}$$

This is the fraction of a halo's stellar population photon production that manages to reach r_{200}. Since $f_{\mathrm{esc}}^{\mathrm{sub}}$ is constant by construction in CoDa II, we focus mostly in the rest of the paper on the determination and behaviour of $f_{\mathrm{esc}}^{\mathrm{ray}}$.

2.4 Escape luminosity

To obtain the instantaneous amount of produced ionizing photons within a given halo $\dot{\eta}_\star$, we sum the ionizing photon production of all emitting star particles (i.e. younger than 10 Myr) within r_{200}, i.e.

$$\dot{\eta}_\star = \mathrm{M_{stars}}(< 10\,\mathrm{Myr}) \times 4.32 \times 10^{46}\,\mathrm{ph\,s^{-1}}\,, \tag{6}$$

where the factor 4.32×10^{46} is the stellar ionizing emissivity in $\mathrm{ph\,s^{-1}\ M_\odot{}^{-13}}$ (taken from table 1 of Ocvirk et al. 2020). We then define the escape luminosity ($\mathrm{L_{esc}}$) of a halo as the product of the intrinsic luminosity and the net halo escape fraction. This gives the number of ionizing photons that reach r_{200} per second:

$$L_{\mathrm{esc}} = \dot{\eta}_\star \times f_{\mathrm{esc}}^{\mathrm{sub}} \times f_{\mathrm{esc}}^{\mathrm{ray}}\,\mathrm{ph\,s^{-1}}, \tag{7}$$

Note that, with these definitions, $\mathrm{L_{esc}}$ is, as expected, proportional to SFR and $f_{\mathrm{esc}}^{\mathrm{ray}}$. If we interpret r_{200} as the limit between the halo and the IGM, $\mathrm{L_{esc}}$ is effectively an estimate of the halo luminosity exiting the halo and entering the IGM. Similarly, we can define the cell escape luminosity, by considering $\dot{\eta}_\star$ as the photon production within that cell and the average transmission Tr_{200} of the paths from that cell to the r_{200} sphere of its host halo.

3 RESULTS: HALO ESCAPE FRACTIONS

3.1 Halo escape fraction as a function of mass and redshift

Fig. 1 shows the distribution of halo escape fraction $f_{\mathrm{esc}}^{\mathrm{ray}}$ as a function of halo mass for four redshifts : $z = 6, 7, 8, 10$.

For masses $\gtrsim 10^{9.5}\,\mathrm{M_\odot}$, the average $f_{\mathrm{esc}}^{\mathrm{ray}}$ decreases, reaching values of <10 per cent for $M_{\mathrm{halo}} > 10^{11}\,\mathrm{M_\odot}$. These smaller values are observed despite massive haloes appearing to heat their r_{200} sphere surroundings with SN activity, bringing their neutral fraction down to 10^{-6} and below. Indeed, massive haloes tend to feature dense, neutral, and opaque cores that trap a large fraction of their ionizing photon production as seen in Fig. 2. A more detailed investigation of the internal properties of these haloes is performed in Section 3.3.

Halo escape fractions are generally higher at low masses, saturating at 1 for all redshifts. The elongated vertical feature of the distribution of haloes at $M_{\mathrm{halo}} \approx 9 \times 10^8\,\mathrm{M_\odot}$, with f_{esc} values lower than 10^{-1}, and at $z = 10, 8, 7$ (top left, top right, bottom left-hand panels of Fig. 1) is populated by haloes in which a star has formed recently, and in which the haloes' expanding H II regions have not yet fully reached the halo boundary at distance r_{200}. Fig. 3 shows the neutral fraction in a plane containing such a halo, illustrating this case.

Comparing the panels of Fig. 1 shows that the extended vertical distribution attributed to haloes in which stars have recently formed progressively disappears between $z = 10$ and $z = 6$, owing to the ionization of the material of the lower mass haloes by the combination of local UV production, SNe energy injection, and outside radiation affecting these haloes, thereby rendering them more and more UV transparent within r_{200}.

There is an intrinsically high scatter in the escape fractions. It is due to the wide range spanned by the properties of the halo population

---

[3] This value is computed from the number of ionizing photons per stellar baryon produced by a binary stellar population of metallicity Z = 0.001 using the BPASS (Eldridge et al. 2017) models, and with a Kroupa IMF (Kroupa 2001).

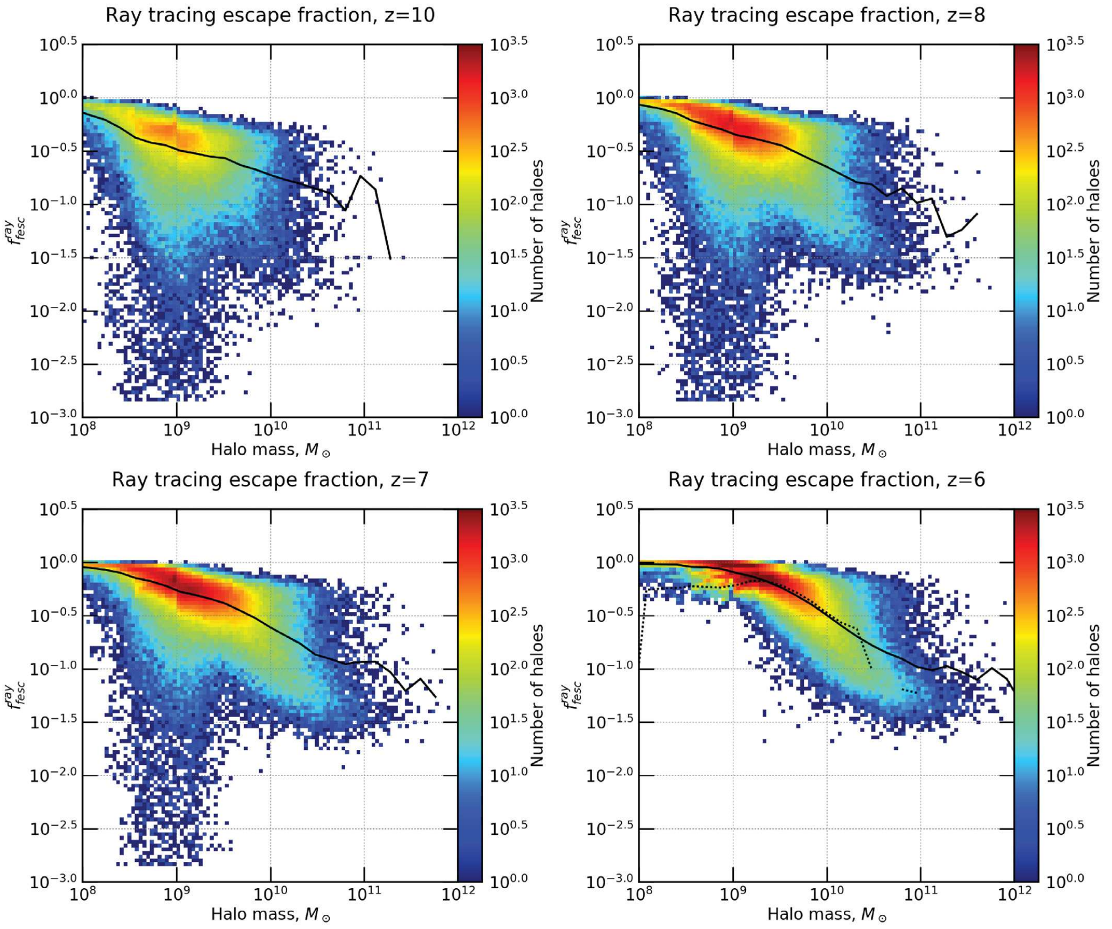


Figure 1. Distributions of halo escape fraction $f_{\rm esc}^{\rm ray}$ of star-forming haloes as a function of $M_{\rm halo}$ at epochs $z = 10, 8, 7, 6$. The full lines indicate the averages of the distributions, i.e. for the star-forming haloes, whereas the dotted line indicates averages computed for the whole population, for the $z = 6$ epoch only. The net halo escape fraction can be obtained by multiplying by the sub-grid escape fraction $f_{\rm esc}^{\rm net} = 0.42\ f_{\rm esc}^{\rm ray}$.

such as the maximum central density, the gas density profile, and the individual accretion history of the halo. These different properties are well sampled because of CoDa II very large size and therefore abundant halo population.

Close examination of the density maps reveals discontinuities around $5 \times 10^8\ {\rm M}_\odot$ and $10^9\ {\rm M}_\odot$ in the highly populated red/orange areas. These are due to resolution effects. In the case of our less massive haloes, the number of cells that represent them can be small therefore increases in r_{200} can affect the resulting $f_{\rm esc}^{\rm ray}$.

Fig. 4 shows the average escape fractions as a function of mass for five epochs, $z = 6$, 7, 8, 10, 14.9. Again, we see high $f_{\rm esc}^{\rm ray}$ values for low-mass haloes, and a negative mass trend from $\approx 10^9\ {\rm M}_\odot$ onwards. There is a clear evolution with redshift for haloes $M_{\rm halo} \lesssim 10^{10}\ {\rm M}_\odot$: the $f_{\rm esc}^{\rm ray}$ average increases with decreasing redshift, reaching $f_{\rm esc}^{\rm ray} \approx 1$ at $z = 6$. Indeed, for a fixed halo mass, higher redshift haloes are denser than their lower redshift counterparts, leading to lower escape fractions. The average behaviour of $f_{\rm esc}^{\rm ray}$ for haloes $\gtrsim 10^{11}\ {\rm M}_\odot$ is unclear as the number of such objects is small.

3.2 Comparison with the literature

The most prominent feature of Fig. 1 is a decrease of escape fraction with mass, in agreement with the literature investigating the escape fraction of ionizing radiation in high-redshift galaxies from numerical simulations (Razoumov & Sommer-Larsen 2010; Yajima et al. 2011; Kimm & Cen 2014; Wise et al. 2014; Paardekooper et al. 2015; Katz et al. 2018), although the slope and extent of the decrease may vary, as can be expected given the range in resolution and modelling of these studies. For dark matter haloes of $10^{11}\ {\rm M}_\odot$, the simulations of Yoo et al. (2020) yield an escape fraction of 7 per cent, also in rather good agreement with our results.

The other striking feature of our results is the evolution of halo escape fraction with time, in particular at masses below $10^{10}\ {\rm M}_\odot$, showing that in CoDa II, such star-forming haloes tend to be more opaque at higher redshifts. For a given halo mass, haloes tend to be denser at higher redshift, and are therefore more likely to yield higher optical depths. Similar evolution, with escape fractions decreasing at higher redshifts, is seen in Kimm & Cen (2014). Razoumov & Sommer-Larsen (2010) report an opposite evolution, perhaps due to their modelling of dust (in their study, dust UV optical depth scales linearly with cell density and metallicity, the latter can be expected to increase on average over time, increasing CGM absorption). This may explain their different result, since dust is not accounted for in our work or in Kimm & Cen (2014).

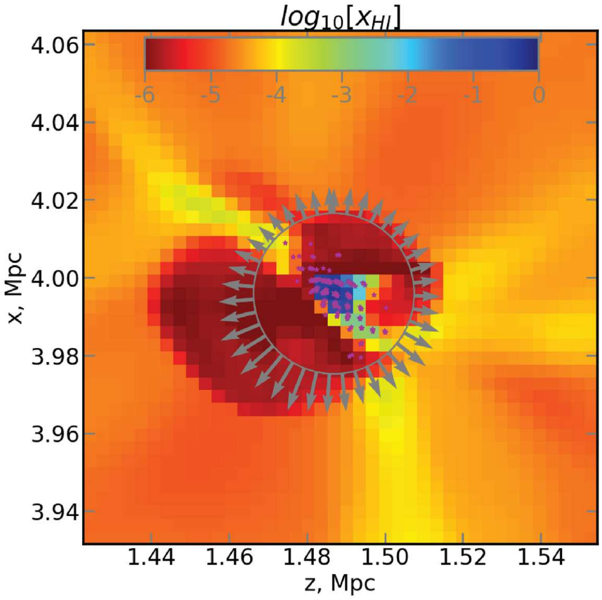


Figure 2. $X_{\rm H\,I}$ (colour) centred in a plane containing an example high-mass halo ($\approx$5.10^{10} M$_\odot$) at $z = 6$. The pink symbols denote active (younger than 10 Myr) stellar particles. Notice the central, neutral region (several cells are close to fully neutral) that exhibits high star formation. The surrounding ionized region is superheated ($T > >10^5$ K) and is generated by supernova explosions and accretion shocks. The circle represents r_{200}. The gas in the halo centre appears to be connected to several filament-like structures (with $10^{-4.5} \gtrsim X_{\rm H\,I} \gtrsim 10^{-3}$).

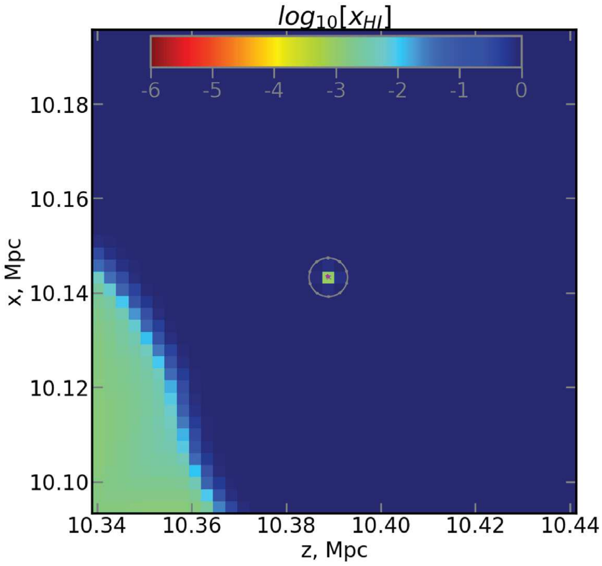


Figure 3. $X_{\rm H\,I}$ map (coded by colour) centred on a plane containing an example halo that has just formed a stellar particle at $z = 8$. Ionizing radiation has not yet reached the surfaces where its flux is sampled. Notice the $X_{\rm H\,I}$ gradient generated by an incoming UV front in the lower left. The symbols are as in Fig .2.

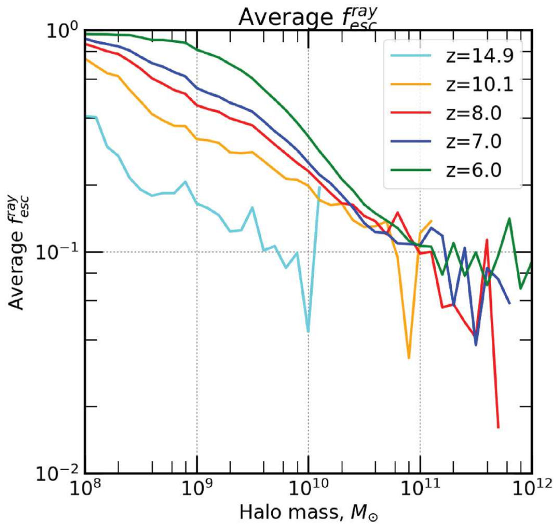


Figure 4. Evolution of the average halo escape fraction $f_{\rm esc}^{\rm ray}$ of star-forming haloes as a function of $M_{\rm halo}$ and redshift. The net halo escape fraction is obtained by multiplying by the sub-grid escape fraction $f_{\rm esc}^{\rm net} = 0.42\ f_{\rm esc}^{\rm ray}$.

3.3 What drives the decrease of escape fractions with mass?

The main trend in all escape fraction plots is a decrease with increasing mass. Since halo escape fraction is determined by density and neutral hydrogen fraction, we want to investigate these particular properties. We show the distributions of the gas properties of star-forming cells in Fig. 5, and their contribution to the SFR and escape luminosity of their halo mass bins.

The top left-hand panel of Fig. 5 shows the distribution of cell transmissions weighted by their SFR for three representative mass bins, normalized by the total SFR in each mass bin (so that the integral of each histogram is 1). The star-forming cells of the lowest mass bin are concentrated around high transmission values, as observed in individual haloes and yielding the low-mass plateaus of $f_{\rm esc}^{\rm ray}$ values close to 1.

There is a stark difference with the intermediate-mass bin, where star-forming cells present a much wider spread of transmission values. The distribution is dominated by cells with intermediate opacity, and peaks at a transmission of $\sim$0.18. For the most massive mass bin, however, the dispersion in transmissions is even higher. Between a few times 0.1 and 10^{-2} the distribution is almost flat, and it extends to extremely opaque cells with transmissions as low as 10^{-6} and below. The high value of the distribution at 10^{-6} indicates that there are close to 40 per cent of star-forming cells with a transmission lower or equal to 10^{-6}. There are very few star-forming cells with a transmission of 1. This implies that most of the absorption of UV photons within our most massive haloes occurs within the cells containing the sources. We remind the reader that this is the distribution of the transmission of halo *cells*, i.e. one value per cell, and not the distribution of the halo-averaged (which would yield one value per halo). This is why the distributions extend much lower than in Fig. 1, which show only halo-averaged values. Please note that the peak at the low transmission end is caused by binning all values $<10^{-6}$ together.

We recompute this histogram, weighing this time by the cells' escape luminosity L$_{\rm esc}$. The result is shown in the top right-hand panel of Fig. 5. This allows us to quantify the contribution of cells to the total escape luminosity of haloes. Unsurprisingly, low-mass

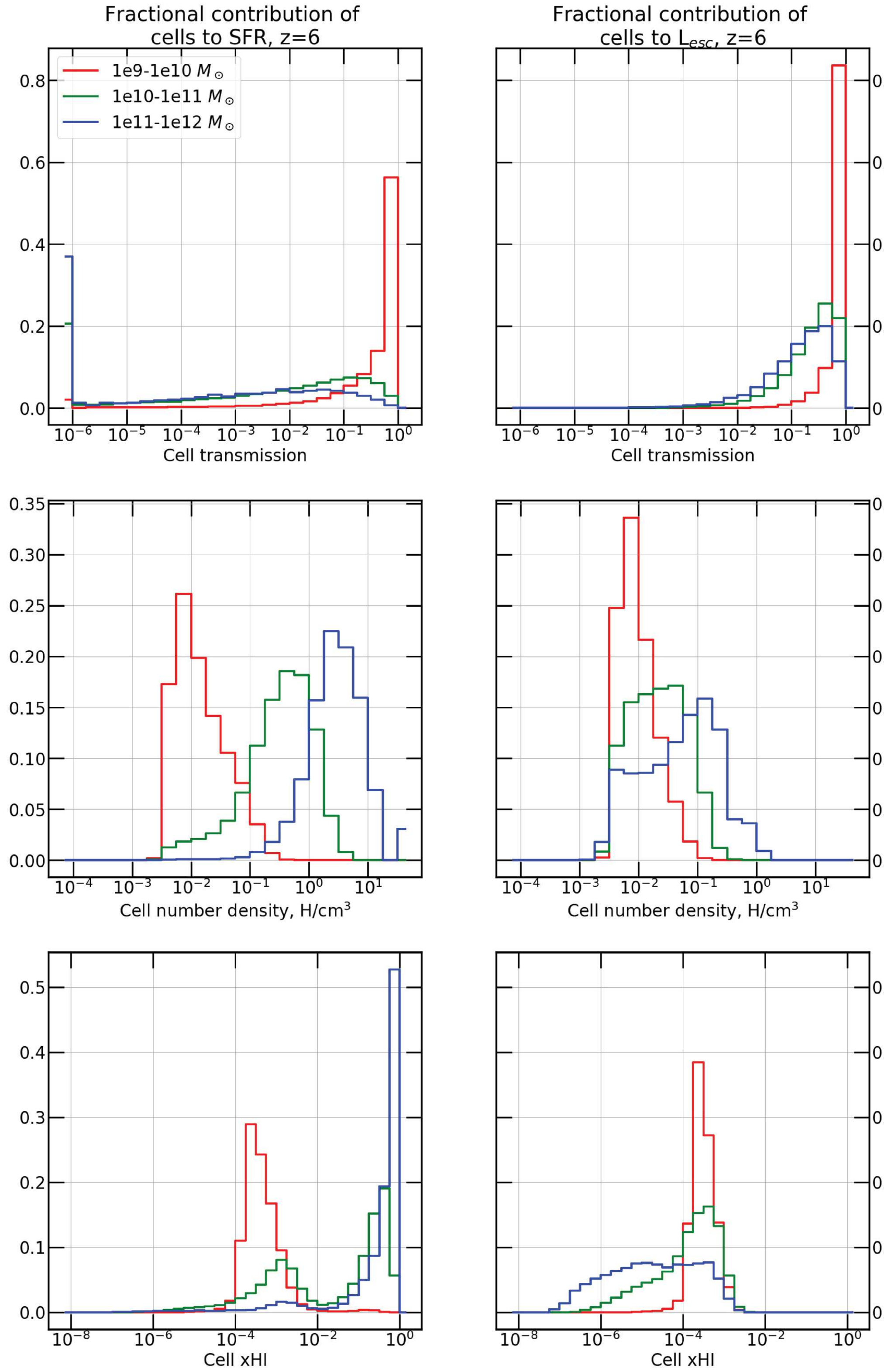


Figure 5. Distribution of star-forming halo cell properties: transmissions (top row), physical gas densities in H cm^{-3}(middle row), neutral fractions (bottom row), for three different mass bins of haloes. The histograms show the contribution of halo cells to (*left:*) total SFR of the mass bin considered, and to (*right:*) the total escape luminosity of the mass bin considered.

haloes' escape luminosities originate from high transmission cells. For intermediate- and high-mass haloes, the situation is slightly more contrasted: while most of the escaping photons originate from high transmission cells, a small fraction (10–15 per cent) are actually produced by moderately opaque regions with transmission below 0.1.

To gain deeper insight into the properties of UV-bright and UV-dark cells, we further examine their physical properties.

The middle left-hand panel of Fig. 5 shows the distribution of cells' gas density for our three representative mass bins, weighted by their SFR, and normalized by the total SFR in each mass bin. The mode of the distribution shifts to higher densities with increasing halo mass, and the density of star-forming cells in the high mass bin extends up to a few 10 H cm^{-3}. This is in stark contrast with the right-hand panel of Fig. 5, which shows the same distribution, but this time weighted by the cells' escape luminosity $\mathrm{L_{esc}}$. The mode of the distributions for the highest and intermediate mass bins are located at much lower densities, and extend no further than about 1 H cm^{-3} therefore showing that in CoDa II, all stars forming at a density larger than this do not contribute to the escape luminosity of their host halo because their cell transmission is too low.

The bottom left-hand panel of Fig. 5 shows the distribution of cells' neutral hydrogen fractions for our three representative mass bins, weighted by their SFR, and normalized by the total SFR in each mass bin. For the lowest mass bin, the neutral fraction distribution has one strong peak at just over 10^{-4}, which explains their high transmission. The value of $\approx 10^{-4}$ reflects the ionization equilibrium for a cell of CoDa II resolution at $z = 6$, with an overdensity of $\delta \approx 50$ that contains a single emitting stellar particle. In the case of the $10^{10-11}\,\mathrm{M}_\odot$ mass bin, haloes' cells are split more or less evenly between ionized and neutral, with two peaks, one centred around 10^{-3} and the other one around a few times 10^{-1} in neutral fractions. A possible origin for the binary nature of the distribution may be the rapidity with which dense star-forming cells ionize and recombine. In this context, cells would jump very quickly from one peak to the other, depending on their star formation activity or lack thereof, and spend very little time between the peaks. The star-forming cells of the highest mass bin are neutral in majority. Though the distribution presents a small ionized peak as well, the high neutral fraction peak is hugely dominant (about 90 per cent of the SFR takes place there) when compared to the ionized peak.

The bottom right-hand panel of Fig. 5 shows the same distribution, but this time weighted by the cells' escape luminosity $\mathrm{L_{esc}}$. It shows the contribution of the cells in terms of escaping photons. For low-mass haloes, the uni-modal distribution is almost unchanged with respect to the left-hand panel. However, for the two higher mass bins, this different weighing has a strong impact on the relative modes of the distribution: indeed, the high neutral fraction mode of the distribution has completely vanished, so that only the high ionization mode remains ($X_{\mathrm{H\,I}} = 10^{-3}$ and below), showing that for these high halo masses, escaping ionizing photons originate predominantly from strongly ionized regions. Another striking aspect of the bottom right-hand panel Fig. 5 is the existence of a tail of the distribution, extending to cells with very high ionization (neutral fractions as as low as 10^{-4}–10^{-7}) and therefore completely transparent. This tail is absent for the low-mass bin. By examining maps of the physical properties of haloes, such as Fig. 2, we found that such high ionization within high-mass haloes is typical of shocked regions. These shocks can either be accretion shocks as seen in Ocvirk, Pichon & Teyssier (2008), Ocvirk et al. (2016), or shocks due to SNe explosions, or a combination of both. Indeed, several studies have shown SN feedback to play an important role in the escape of ionizing photons (Kimm & Cen 2014; Trebitsch et al. 2017), and we interpret this very high ionization tail in our distribution as another manifestation of this effect. Examining Fig. 2 in the light of these results also allows us to locate the regions contributing to the escape luminosity of the halo. The cells of the low ionization peak, at $X_{\mathrm{H\,I}} \geq 0.1$, belong to the opaque neutral central core of the halo, from which no ionizing photons escape. The high ionization mode at $X_{\mathrm{H\,I}} \sim 10^{-3}$ and below, consists of cells located along the accreting gas filaments, up to the virial radius. Finally, the star-forming cells in the very strong ionization tail of the distribution, at $X_{\mathrm{H\,I}} < 10^{-5}$–$10^{-6}$ are located in regions heated by SN feedback.

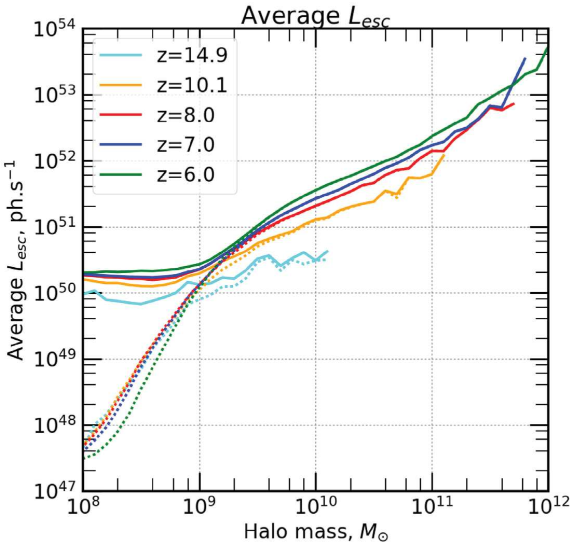


Figure 6. Evolution of the average halo escape luminosity $\mathrm{L_{esc}}$ of star-forming haloes (the solid line) and all haloes (the dotted line) as a function of M_{halo} and redshift.

Although the shock-heated tail is not seen in the low-mass bins, it shows up when increasing spatial resolution, as demonstrated in Appendix A. However, it does not result in an increase in f_{esc}, because neutral cells also become more frequent within these haloes, and the net result of increasing resolution is a decrease of f_{esc}.

3.4 Escape luminosities

Using our determinations of halo escape fractions, we can now compute the halo escape luminosities as in equation (7). We show the resulting $\mathrm{L_{esc}}$ as a function of halo mass and their evolution with redshift in Fig. 6. The escape luminosity increases with halo mass, despite the decrease of f_{esc}. Indeed, at $z = 6$, for instance, f_{esc} decreases roughly as $M_{\mathrm{halo}}^{-0.34} \sim M_{\mathrm{halo}}^{-1/3}$, as shown in Appendix B, whereas halo SFR increases as $M_{\mathrm{halo}}^{5/3}$, and therefore the product $\mathrm{L_{esc}} \propto M_{\mathrm{halo}}^{4/3}$ increases with halo mass. At the low-mass end, $\mathrm{L_{esc}}$ flattens as we reach the quantization limit of star formation in the simulation: between 10^{8} and $10^{9}\,\mathrm{M}_\odot$, star-forming haloes host only one emitting star particle of the same elementary mass $M_\star = 11732\,\mathrm{M}_\odot$. The evolution of $\mathrm{L_{esc}}$ with redshift reflects that of the escape fraction f_{esc}. The impact of star formation suppression by radiation feedback is not readily seen in the solid lines of the figure because they represent the average $\mathrm{L_{esc}}$ of star-forming haloes only. However, the dotted lines shows the average $\mathrm{L_{esc}}$ for all haloes, where the suppression-driven decrease of SFR with redshift leads to lower $\mathrm{L_{esc}}$ at low redshifts below $10^{9}\,\mathrm{M}_\odot$.

4 RESULTS: PHOTON BUDGET

We now turn to investigating the contribution of haloes of different masses and different M_{AB1600} to cosmic reionization in CoDa II.

4.1 Photon budget versus mass

We sort the CoDa II dark matter halo population into 40 logarithmic mass bins between 10^8 and 10^{12} $\mathrm{M}_{\odot}$. We define the total escape luminosity of a given mass bin as the sum of the escape luminosity of all haloes within that mass bin. This quantity depends on the total number of haloes in the CoDa II volume, and we wish to conduct our study using a quantity independent of simulation box size, to ease comparison with future semi-analytical models and simulations. Hence, we further define the escape emissivity, as the total escape luminosity of a given mass bin divided by the simulation volume.

The left-hand panel of Fig. 7 shows this escape emissivity as a function of halo mass, at five epochs (the full lines). This represents the cosmic ionizing photon budget for the CoDa II simulation, i.e. the distribution of the contributions of each mass bin to the total rate of photons reaching the IGM and driving reionization, for a 1 cMpc^3 volume.

The contribution to the cosmic escape emissivity culminates around $\sim 10^{9.5}$ $\mathrm{M}_{\odot}$ between $z = 6$ and $z = 8$ and decreases at lower and higher masses. This leads to a photon budget that is dominated by haloes of a few times 10^8 up to a few times 10^{10} $\mathrm{M}_{\odot}$ between these redshifts.

To quantify the impact of the smaller escape fraction of massive haloes on the photon budget, we also show in the left-hand panel of Fig. 7 the intrinsic photon *production* budget, i.e. proportional to the SFR of haloes (the dotted line), compared to the escaping photon budget. At $z = 6$, for instance, intrinsic photon production peaks at $10^{10.5}$ $\mathrm{M}_{\odot}$. This is the result of two opposite trends: SFR increases with increasing halo mass, whereas the number abundance of haloes decreases, as dictated by the halo mass function. This competition yields $10^{10.5}$ $\mathrm{M}_{\odot}$ haloes as the foremost contributors to the total cosmic star formation at $z = 6$. However, their contribution to the escaping photon budget is strongly affected by their low escape fractions, to the point that, even though they dominate all other haloes in terms of SFR, they are outshined in total escape emissivity by the 10^9–10^{10} $\mathrm{M}_{\odot}$ halo population.

To further compare the contribution of various mass bins, the right-hand panel of Fig. 7 shows the cumulative version of the photon budget. It allows us to directly read from the plot that at $z = 6$, for instance, haloes within $2 \times 10^9\,\mathrm{M}_{\odot} \lesssim M_{\mathrm{halo}} \lesssim 10^{10}\,\mathrm{M}_{\odot}$ produce around 50 per cent of all ionizing photons. The shape of the cumulative photon budget is rather similar at all redshifts.

However, the cumulative distributions shift towards lower masses as redshift increases, by about half a decade between $z = 6$ and $z = 10$. This is also seen as a shift of the peak of the distribution (left-hand panel) between $z = 6$ and $z = 8$. This shift reflects the build-up and evolution of the halo mass function towards more abundant and more massive halo populations. At all redshifts but the highest, the cumulative distribution has its 10 per cent and 90 per cent levels separated by about 1.5 decades in mass, meaning this range is responsible for 80 per cent of the ionizing photon budget. In CoDa II, the ionized volume fraction goes from 20 per cent to 100 per cent between $z = 6$ and 8, and it is $\sim$50 per cent at $z = 7$. We therefore retain $z = 7$ as the epoch most representative of 'ongoing' reionization and detail the photon budget for this redshift: we read from the cumulative distribution that the photon budget at this epoch is dominated by galaxies with dark matter masses of 6×10^8–$3 \times 10^{10}\,\mathrm{M}_{\odot}$, which produce 80 per cent of the ionizing photons reaching the IGM. They are therefore the main drivers of cosmic reionization.

4.2 Photon budget by M_{AB1600}

The luminosity function of CoDa II haloes was shown to be in good agreement with observations in Ocvirk et al. (2020). Here, we use the halo magnitudes to recast our photon budget analysis into a M_{AB1600} scale instead of the halo mass scale.

Fig. 8 shows the total ionizing photon contribution as a function of M_{AB1600} (left), as well as the equivalent cumulative distribution (right). The shape of the distribution and its temporal evolution are similar to that obtained as a function of mass (Fig. 7). This is a direct consequence of the SFR–halo mass relation shown in Ocvirk et al. (2020), which drives an M_{AB1600}–halo mass relation.

Combined, these figures illustrate that the main contributors to reionization lie within a magnitude range of around -12 to about -19. Which is in broad agreement with the semi-analytical results of Liu et al. (2016) for $z > 7$. More specifically, the 80 per cent escape luminosity range (reading the 10–90 per cent levels of the cumulative escape luminosity distribution function) is $M_{\mathrm{AB1600}} = [-13, -19]$ at $z = 6$, and $[-12, -17]$ at $z = 8$. This shift with redshift is due to the build-up of the galaxy mass function.

This suggests that very deep surveys in cluster fields, such as Bouwens et al. (2017) and Atek et al. (2018), if reliable down to $M_{\mathrm{AB1600}} = -13$, are indeed starting to see the bulk (>80 per cent) of the population driving cosmic reionization at $z = 6$. However, at $z = 6$, cosmic reionization is already finished in CoDa II, and in the middle of reionization, i.e. at $z = 7$, the galaxies seen by Bouwens et al. (2017) are brighter than $M_{\mathrm{AB1600}} = -17$, and Fig. 8 shows that these can only account for $\sim$20 per cent of the ionizing luminosity. To see the bulk of galaxies driving reionization when it is in full spin, surveys would need to achieve a similar $M_{\mathrm{AB1600}} = -12$ depth at $z = 7$.

Finally, early reionization at $z = 15$ is driven by galaxies of a very narrow range of halo masses, 10^8–$10^9\,\mathrm{M}_{\odot}$, corresponding to magnitudes fainter than $M_{\mathrm{AB1600}} = -14$, out of reach of current and future planned observatories at these redshifts.

4.3 Comparison with the literature

Katz et al. (2018, 2019) use tracers and RAMSES-RT RHD simulations to study the contribution of haloes to the cosmic ionizing luminosity. Although their technique differs in many respects from ours (we do not use such tracers, they use variable speed of light, AMR), their results are in rather good agreement with ours. At $z = 8$, they find that 70 per cent of the ionizing luminosity is produced by haloes of $10^9\,\mathrm{M}_{\odot} \lesssim M_{\mathrm{halo}} \lesssim 10^{10}\,\mathrm{M}_{\odot}$ (from fig. 6 of Katz et al. 2018). We can read directly from Fig. 7 that at the same redshift, this mass range is responsible for 60 per cent of the ionizing photons in CoDa II. However, at $z = 6$, Katz et al. (2018) find that high-mass haloes of $\gtrsim 10^{10}\,\mathrm{M}_{\odot}$ produce the majority of ionizing photons (60 per cent), while we find that they account for only 40 per cent of the ionizing luminosity, i.e. the largest contribution originates from haloes less massive than $10^{10}\,\mathrm{M}_{\odot}$ in CoDa II.

At higher redshifts, though ($z > 12$), the Katz et al. (2018) values fluctuate too much for a meaningful comparison. Also,

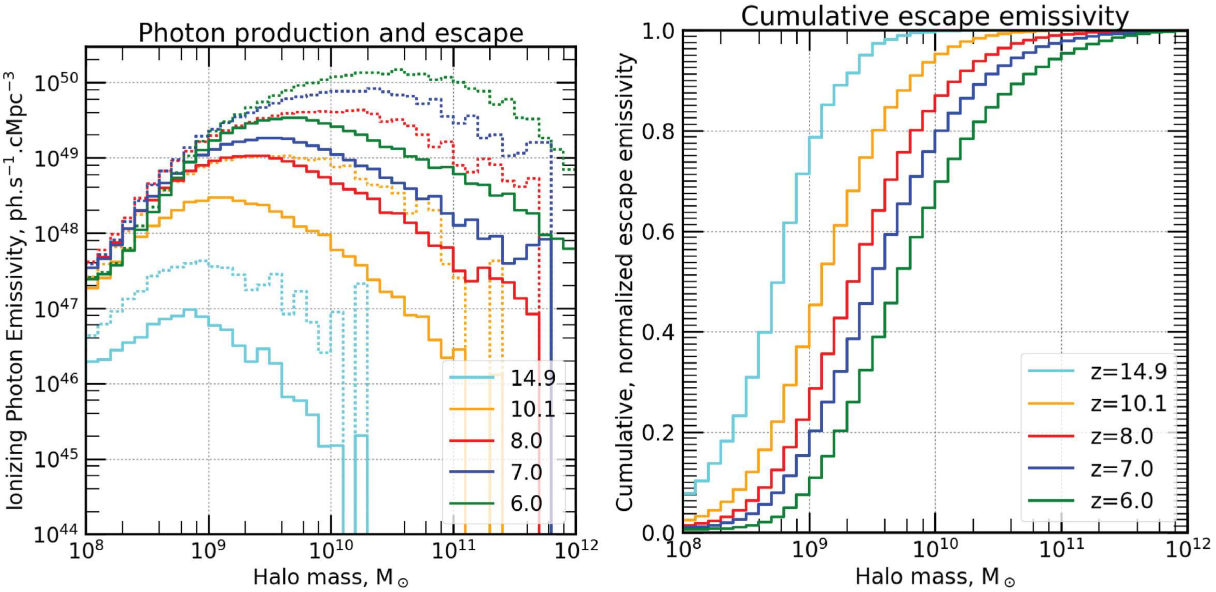


Figure 7. Left : escape emissivity as a function of halo mass for several CoDa II epochs. The displayed emissivity is the sum of all haloes' escape luminosities in the mass bin considered at that redshift, divided by the simulation volume. There are 10 mass bins per mass decade. The full lines correspond to the escape emissivity, whereas the dashed lines correspond to the intrinsic (unabsorbed) emissivities (i.e. derived directly from the star formation rate). Right : Cumulative version of left-hand panel.

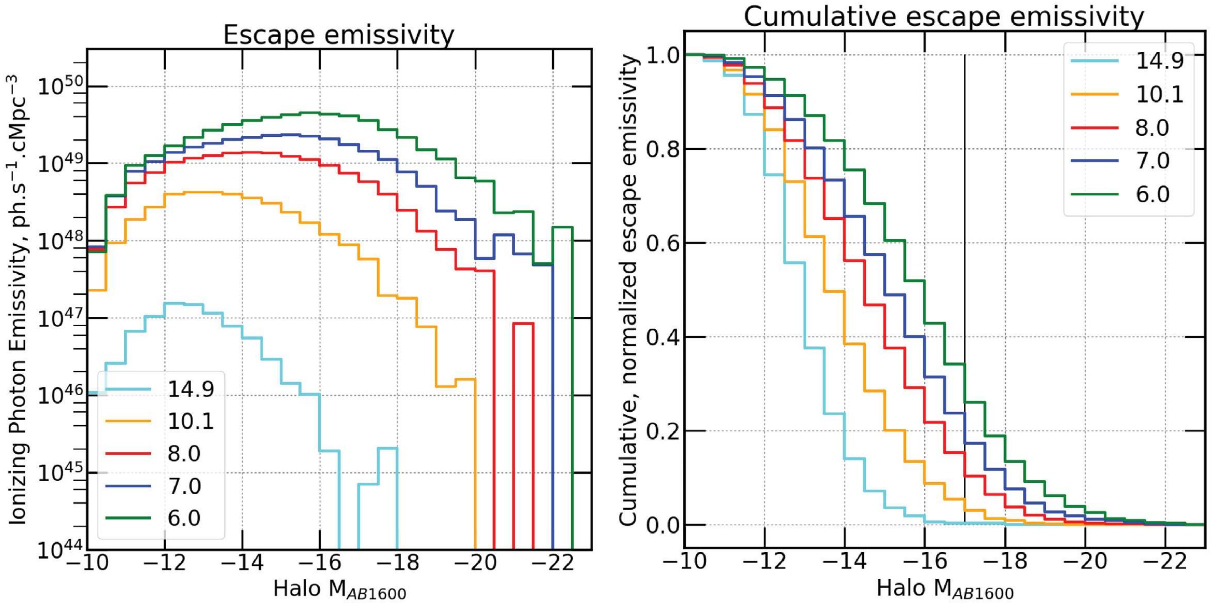


Figure 8. Left : Total escape emissivity as a function of magnitude for several CoDa II epochs. The displayed emissivity is the sum of all haloes' escape luminosities in the $M_{\rm AB1600}$ bin considered, divided by the simulation volume. There are four bins per mag. Right : Cumulative version of the total escape emissivity plot on the left. The vertical bar marks the $M_{\rm AB1600} = -17$ limit.

because of their smaller box size, their sample is devoid of haloes more massive than $10^{11}\,\rm M_\odot$ even at $z = 6$, unlike in CoDa II, where such massive haloes are present already at $z = 10$. However, their low number density and escape fractions prevent them from contributing significantly to cosmic reionization: their total escape luminosity is less than 10 per cent at all redshifts, which is why our results are in fair agreement with Katz et al. (2018), even though they do not include these high-mass haloes.

Yajima et al. (2011) find that haloes below $M_{\rm halo} \lesssim 10^{10}\,\rm M_\odot$ contribute about ≈75 per cent of the ionizing luminosity at $z = 6$ (summing the two lowest mass bins of the $z = 6$ panel of their fig. 12).

Again, this is in reasonable agreement with our findings, although at this redshift the largest contribution (45 per cent) comes from haloes in the mass range 10^9–$10^{9.5}$ $\mathrm{M}_\odot$, while the contribution of this mass range is about two times smaller in CoDa II. This discrepancy could be due to the lack of radiative feedback on the SFR of low-mass haloes in (Yajima et al. 2011): indeed, their study performs RT as post-processing, whereas radiation and hydrodynamics are fully coupled in CoDa II, which mitigates star formation in low-mass haloes, as shown in Ocvirk et al. (2020), Dawoodbhoy et al. (2018) and similarly in Wu et al. (2019), Ma et al. (2018), and therefore intrinsically reduces their contribution to cosmic reionization.

A possible caveat in our work is the assumption we make in Section 2.2 that each halo contains one galaxy. This could potentially result in the blending of galaxies whose real equivalents would be distinguishable in observations as two separate star-forming objects. A visual inspection of a sample of our haloes reveals that this preferentially occurs in some of the most massive haloes ($M_{\mathrm{halo}} > 10^{11}$ $\mathrm{M}_\odot$), at a rate of less than 20 per cent. Since this mass range only contributes to a few per cent of the photon budget, tentatively correcting for this effect would impact the photon budget only at the percent level, and would not change our conclusions.

Finally, we address the possible dependence of our results on spatial resolution by performing a follow up higher resolution simulation.

We show in appendix A that indeed, increasing spatial (mass) resolution in RAMSES-CUDATON by a factor of 4 (64) may yield lower halo escape fractions by a factor of 2, while retaining a similar slope. Such a global re-scaling over all masses, without changing the slope of the f_{esc}–halo mass relation, should not dramatically affect our results on the photon budget and the predominant halo mass scale driving reionization because it does not alter the *relative* contributions between halo mass bins.

5 CONCLUSIONS

We use the CoDa II fully coupled RHD simulation of the EoR to study the photon budget of galaxies during the EoR. To do so, we start out by investigating the escape fractions of CoDa II galaxies. We find that the halo escape fraction (i.e. the fraction of ionizing photons produced by the halo's stars reaching the virial radius) is a decreasing function of halo mass.

To gain insight into the evolution of the $f_{\mathrm{esc}}^{\mathrm{ray}}$ with halo mass, we examine the properties of the halo cells as a function of their SFRs and escape luminosity. We find that for intermediate- and high-mass haloes, the neutral fractions of star-forming cells exhibit a strongly bi-modal distribution, with a neutral mode and an ionized mode at $X_{\mathrm{H\,I}} \sim 10^{-3}$. The neutral mode is completely opaque, meaning that escaping ionizing photons originate from the star-forming, ionized regions of the haloes. The halo escape fractions we obtained closely reflect the distribution of the star-forming cells between the neutral opaque mode and the ionized, transparent mode. For instance, CoDa II high-mass haloes (10^{11} $\mathrm{M}_\odot$) have an average halo escape fraction of ~10 per cent because 90 per cent of their young stars reside in central, dense, fully opaque regions, while the remaining 10 per cent of their young stars reside in transparent regions allowing their photons to escape.

Moreover, we find a slow evolution of the halo escape fraction with redshift: haloes of a given mass are more opaque at higher redshift. This is due to the fact that for a fixed mass, haloes at higher redshifts tend to be more concentrated than their low-redshift counterparts.

In Appendix B, we provide a functional form fit to our average halo escape fraction results, so as to allow its use in semi-analytical models of the EoR such as 21cmFAST (Mesinger, Furlanetto & Cen 2011) or Fialkov et al. (2013).

We then use the halo escape fractions of our haloes to investigate the contributions of galaxies of various masses to the total ionizing emissivity during the EoR. We show that CoDa II galaxies within 3×10^{10} $\mathrm{M}_\odot \gtrsim M_{\mathrm{halo}} \gtrsim 6 \times 10^8$ $\mathrm{M}_\odot$ produce about 80 per cent of all the ionizing photons reaching the IGM at $z = 7$, which is the middle of reionization in CoDa II ($x_{\mathrm{H\,I}} = 50$ per cent). They can therefore be considered as the main drivers of cosmic reionization, although, at $z = 6$ (8), the mass range accounting for 80 per cent of the photon budget is slightly more (less) massive, by 0.25 dex.

The foremost mass range reionizing the Universe emerges as the result of a competition between the different processes exposed throughout this paper, and can be summarized as follows: the numerous low-mass haloes are too inefficient at forming stars to contribute significantly despite their high halo escape fractions, whereas the high-mass haloes are too few and have escape fractions that are too low to contribute significantly, despite their high SFR.

As a consequence, the low-mass end (below 5×10^9 $\mathrm{M}_\odot$) and the high-mass end (above 5×10^{10} $\mathrm{M}_\odot$) contribute, respectively, only less than 10 per cent each to the total ionizing photon budget between $z = 8$ and $z = 6$.

Our results are in reasonable agreement with the (not exhaustive) literature reviewed, despite a number of differences in numerical treatment, and assumptions on stellar populations and their feedback, which explain the deviations from our results.

Ideally, we would like to follow up on our study by pushing future CoDa II-like simulations to higher spatial resolutions, possibly using AMR to provide a better description of the ISM, and its processes and if possible rely less on a sub-grid escape fraction, as well as improved physics such as chemical enrichment, AGN mechanical and radiative feedbacks, and their possible contribution to reionization. The continued growth of supercomputers, because of hybrid nodes mixing many-core CPUs and GPUs, may allow us to get there in the near future, provided we overcome a number of technical hurdles related to code architecture and optimization.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the judicious comments of the anonymous referee, as well as those of Max Gronke. This work has used PYTHON, and the following packages : PYTHON : MATPLOTLIB (Hunter 2007), NUMPY (Van Der Walt, Colbert & Varoquaux 2011), SCIPY (Virtanen et al. 2020), HEALPIX (Górski et al. 2005). PO acknowledges support from the French ANR funded project ORAGE (ANR-14-CE33-0016). ND and DA acknowledge funding from the French ANR for project ANR-12-JS05-0001 (EMMA). ITI was supported by the Science and Technology Facilities Council (grant numbers ST/F002858/1 and ST/I000976/1) and the Southeast Physics Network (SEPnet). JS acknowledges support from "l'Oréal-UNESCO 'Pour les femmes et la Science'" and the 'Centre National d'études spatiales (CNES)' postdoctoral fellowship programs. KA was supported by NRF (grant number NRF-2016R1D1A1B04935414). GY acknowledges financial support by the MINECO/FEDER under project grant AYA2015-63810-P and MICIU/FEDER under project grant number PGC2018-094975-C21. PRS was supported in part by U.S. NSF grant number AST-1009799, NASA grant number NNX11AE09G, NASA/JPL grant number

RSA Nos. 1492788 and 1515294, and supercomputer resources from NSF XSEDE grant number TG-AST090005 and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. The CoDa II simulation was performed at Oak Ridge National Laboratory/Oak Ridge Leadership Computing Facility on the Titan supercomputer (INCITE2016 award AST031). Processing was performed on the Eos and Rhea clusters. Auxiliary simulations were performed at pôle HPC de l'Université de Strasbourg (méso-centre). The simulations used for the resolution study were performed on CSCS/Piz Daint (Swiss National Supercomputing Centre), as part of the 'SALT: Shining a light through the dark ages' PRACE allocation obtained via the 16th call for PRACE Project Access (project id pr37). A series of test simulations for the initial conditions of CoDa II were performed at LRZ Munich within the project pr74no. This work has used v2.1 of the Binary Population and Spectral Synthesis (BPASS) models as last described in Eldridge et al. (2017).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. See Ocvirk et al. (2020) for specifics about the availability of CoDa II.

REFERENCES

Ahn K., Iliev I. T., Shapiro P. R., Mellema G., Koda J., Mao Y., 2012, ApJ, 756, L16
Anderson L., Governato F., Karcher M., Quinn T., Wadsley J., 2017, MNRAS, 468, 4077
Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, MNRAS, 479, 5184
Aubert D., Teyssier R., 2008, MNRAS, 387, 295
Aubert D., Deparis N., Ocvirk P., 2015, MNRAS, 454, 1012
Aubert D. et al., 2018, ApJ, 856, L22
Barkana R., Loeb A., 2007, Rep. Prog. Phys., 70, 627
Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, ApJ, 843, 129
Dawoodbhoy T. et al., 2018, MNRAS, 480, 1740
Dayal P., Ferrara A., 2018, Phys. Rev., 780, 1
Dayal P. et al., 2020, MNRAS, 495, 3065
Deparis N., Aubert D., Ocvirk P., Chardin J., Lewis J., 2019, A&A, 622, A142
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Aust., 34, e058
Ferrara A., Loeb A., 2013, MNRAS, 431, 2826
Fialkov A., Barkana R., Visbal E., Tseliakhovich D., Hirata C. M., 2013, MNRAS, 432, 2909
Gnedin N. Y., 2016, ApJ, 833, 66
Górski K. M., Hivon E., Banday A. J., Wand elt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Haardt F., Salvaterra R., 2015, A&A, 575, L16
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625
Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725
Izotov Y. I., Orlitová I., Schaerer D., Thuan T. X., Verhamme A., Guseva N. G., Worseck G., 2016, Nature, 529, 178
Katz H., Kimm T., Haehnelt M., Sijacki D., Rosdahl J., Blaizot J., 2018, MNRAS, 478, 4986
Katz H., Kimm T., Haehnelt M. G., Sijacki D., Rosdahl J., Blaizot J., 2019, MNRAS, 483, 1029
Kimm T., Cen R., 2014, ApJ, 788, 121
Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 466, 4826
Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, MNRAS, 486, 2215
Kroupa P., 2001, MNRAS, 322, 231
Legrand L. et al., 2019, MNRAS, 486, 5468
Levermore C. D., 1984, J. Quant. Spectrosc. Radiat. Transfer, 31, 149
Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 462, 235
Ma X., Kasen D., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Kereš D., Murray N., 2015, MNRAS, 453, 960
Ma X. et al., 2018, MNRAS, 478, 1694
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Ocvirk P. et al., 2016, MNRAS, 463, 1462
Ocvirk P., Aubert D., Chardin J., Deparis N., Lewis J., 2019, A&A, 626, A77
Ocvirk P. et al., 2020, MNRAS
Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544
Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, MNRAS, 485, 47
Razoumov A. O., Sommer-Larsen J., 2010, ApJ, 710, 1239
Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
Rosdahl J., et al., 2018, MNRAS, 479, 994
Roy F., Bouillot V. R., Rasera Y., 2014, A&A, 564, A13
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Sorce J. G. et al., 2016, MNRAS, 455, 2078
Teyssier R., 2002, A&A, 385, 337
Trebitsch M., Blaizot J., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 470, 224
Trebitsch M., Volonteri M., Dubois Y., Madau P., 2018, MNRAS, 478, 5607
Trebitsch M. et al., 2020, preprint (arXiv:2002.04045)
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230
Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560
Wu X., Kannan R., Marinacci F., Vogelsberger M., Hernquist L., 2019, MNRAS, 488, 419
Yajima H., Choi J.-H., Nagamine K., 2011, MNRAS, 412, 411
Yoo T., Kimm T., Rosdahl J., 2020, preprint (arXiv:2001.05508)

APPENDIX A: ESCAPE FRACTION RESOLUTION STUDY

We have shown that the main features of our $f_{\rm esc}^{\rm ray}$ values are that they decrease with increasing halo mass above a certain mass scale, with haloes of $\lesssim 10^9\,{\rm M}_\odot$ having high $f_{\rm esc}^{\rm ray}$ values, and increase slightly with the progress of the EoR. Since this first trend is driven by the presence of dense, neutral, and UV opaque cells in haloes, one may expect that increasing the number of resolution elements (and thus within the same volume and same total mass, increasing the maximum possible density) could affect the exact relation between $f_{\rm esc}^{\rm ray}$ and halo mass, as well as the mass scale at which the break between trends occurs.

To investigate this, and to test the robustness of our previously presented results, we proceed to study the resolution convergence of $f_{\rm esc}^{\rm ray}$. We ran a series of high-resolution simulations using RAMSES-CUDATON: we used a 4 cMpc.h^{-1} sided box, with 1024^3 resolution elements (i.e. four times the spatial resolution of CoDa II, i.e. 64 times better mass resolution). We will call it high-res from now on. This set-up yields a comoving cell size of 5.76 ckpc (0.82 kpc physical at $z = 6$), and a dark matter particle mass of $6.4 \times 10^3\,{\rm M}_\odot$.

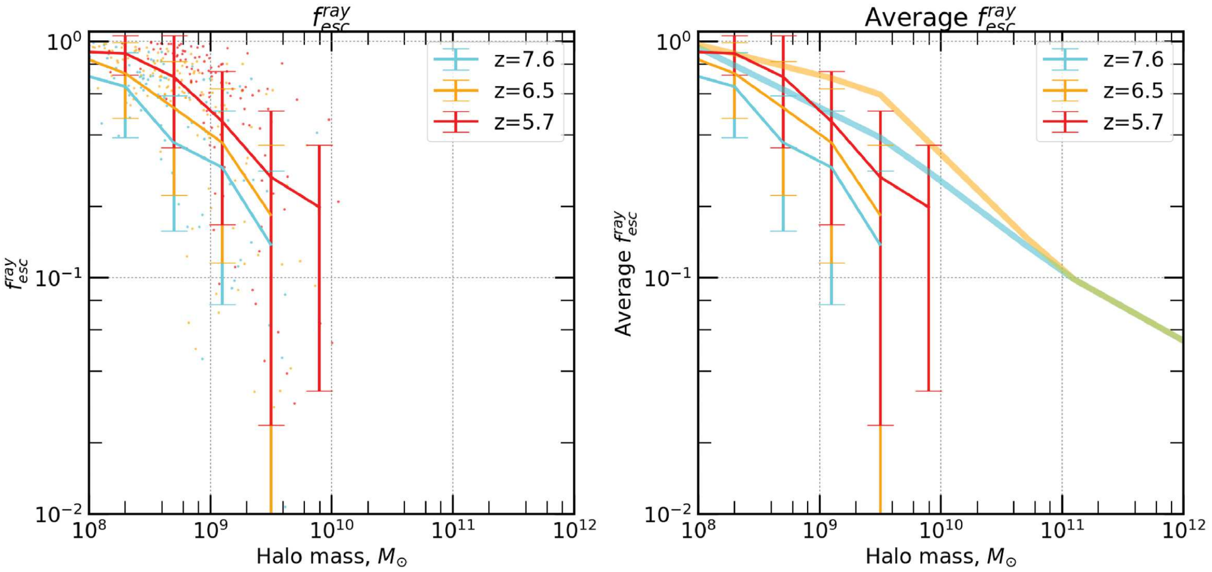


Figure A1. Left : Scatter and mass bin average of $f^{\rm ray}_{\rm esc}$ of star-forming haloes over the full range of masses in the high-res simulation box | Right : average $f^{\rm ray}_{\rm esc}$ of star-forming haloes for the high-res box, compared with our fitting formula (see equation B1) for CoDa II $f^{\rm ray}_{\rm esc}$ (at redshifts that are appropriate, $10 > z > 6$). Here, we plot within the mass range of CoDa II haloes.

All other parameters are kept exactly as in CoDa II, including the minimum stellar mass particle of 11 732 $\mathrm{M}_{\odot}$. The initial conditions are necessarily different from CoDa II because the high-res run uses a smaller box. As a consequence, it contains fewer star-forming haloes of a given mass than CoDa II, making $f^{\rm ray}_{\rm esc}$ measurement noisier. The latter must therefore be carefully compared with oru former results.

We performed our previous $f^{\rm ray}_{\rm esc}$ measurements on the high-res simulation to test the impact of numerical resolution on $f^{\rm ray}_{\rm esc}$ in CoDa II.

The left-hand panel of Fig. A1 shows $f^{\rm ray}_{\rm esc}$ as a function of halo mass for the high-res simulation. This high-resolution case also presents high $f^{\rm ray}_{\rm esc}$ values for low-mass haloes, as well as decreasing $f_{\rm esc}$ with halo mass. Moreover, $f^{\rm ray}_{\rm esc}$ increases on average with time for low-mass haloes, as in CoDa II. The right-hand panel presents a direct comparison of these averages with the fitting formula for CoDa II at the same redshifts where appropriate ($10 \gtrsim z \gtrsim 6$, the domain of validity for our fitting formula given in Fig. B1).

However, as anticipated, there are differences. In the high-res box, the average $f^{\rm ray}_{\rm esc}$ is lower for all masses than in CoDa II, and the slope with mass is slightly more pronounced. However, there is a large scatter around the average of the high-res data set. Indeed, there are only a few hundred star-forming haloes of all masses in the high-res box at $z = 5.7$. Reassuringly, the difference between the two boxes is akin to a global re-normalization of the average $f^{\rm ray}_{\rm esc}$. Therefore, while numerical resolution may change the *absolute* $\mathrm{L}_{\rm esc}$ of a given halo mass bin, it is not likely to change the *relative* balance between mass bins in the photon budget, which is our main result.

To explain the smaller $f^{\rm ray}_{\rm esc}$ values measured in high-res, we proceed as in 3.3, and examine the properties of the gas of star-forming cells. Fig. A2 shows the distribution of the neutral fractions of star-forming cells in high-res, weighted by their SFR ($\mathrm{L}_{\rm esc}$) in the left-hand right)-hand panel, at $z = 5.7$. In high-res, the neutral fractions of star-forming cells in haloes of $10^{10}\,\mathrm{M}_{\odot} > M_{\rm halo} > 10^{9}\,\mathrm{M}_{\odot}$ are grouped into two peaks: an ionized mode, centred around $X_{\rm H\,I} \sim 10^{-4}$, and a neutral/quasi-neutral mode at $X_{\rm H\,I} \sim 0.5$. The $\mathrm{L}_{\rm esc}$ weighted distribution shows that only the cells with $X_{\rm H\,I} \leq 10^{-3}$ contribute significantly to the final $\mathrm{L}_{\rm esc}$ of the haloes in high-res. The cells belonging to the high neutral fraction account for ≈50 per cent of star formation, and do not contribute to haloes' $\mathrm{L}_{\rm esc}$.

Fig. A2 also shows the CoDa II distributions, allowing us to gauge directly the impact of increased resolution. The first striking difference between the two simulations is the shape of the SFR-weighted distribution. In CoDa II, most of the star formation in the $10^{10}\,\mathrm{M}_{\odot} > M_{\rm halo} > 10^{9}\,\mathrm{M}_{\odot}$ haloes happens in cells with neutral fractions $<10^{-3}$, and there is no equivalent to the high neutral fraction peak seen in high-res. However, this feature is present in the distributions of cell xHI in higher mass haloes in CoDa II, which are better resolved. This is a convincing hint that the increase of resolution between CoDa II and high-res allows higher density cells to exist within haloes of the same mass, in turn allowing for higher recombination rates, higher neutral fractions, and finally lower LyC transmissions and escape fractions.

There is, however, a second aspect worthy of discussion. In CoDa II, the star-forming cells of haloes of $10^{10}\,\mathrm{M}_{\odot} > M_{\rm halo} > 10^{9}\,\mathrm{M}_{\odot}$ do not exhibit neutral fractions smaller than 10^{-4}, whereas in high-res the distributions of neutral fraction stretch all the way to 10^{-5}. Higher mass haloes in CoDa II have similarly low neutral fractions that can only occur when the gas is heated by SN. This suggests that although the higher resolution of high-res allows for higher gas densities that can decrease the escape fraction, it also permits the existence of lower gas density cells that are more susceptible to the heating from SN, plausibly increasing the $f^{\rm ray}_{\rm esc}$ of haloes by allowing higher transmission cells to exist. The fact that the escape fractions of

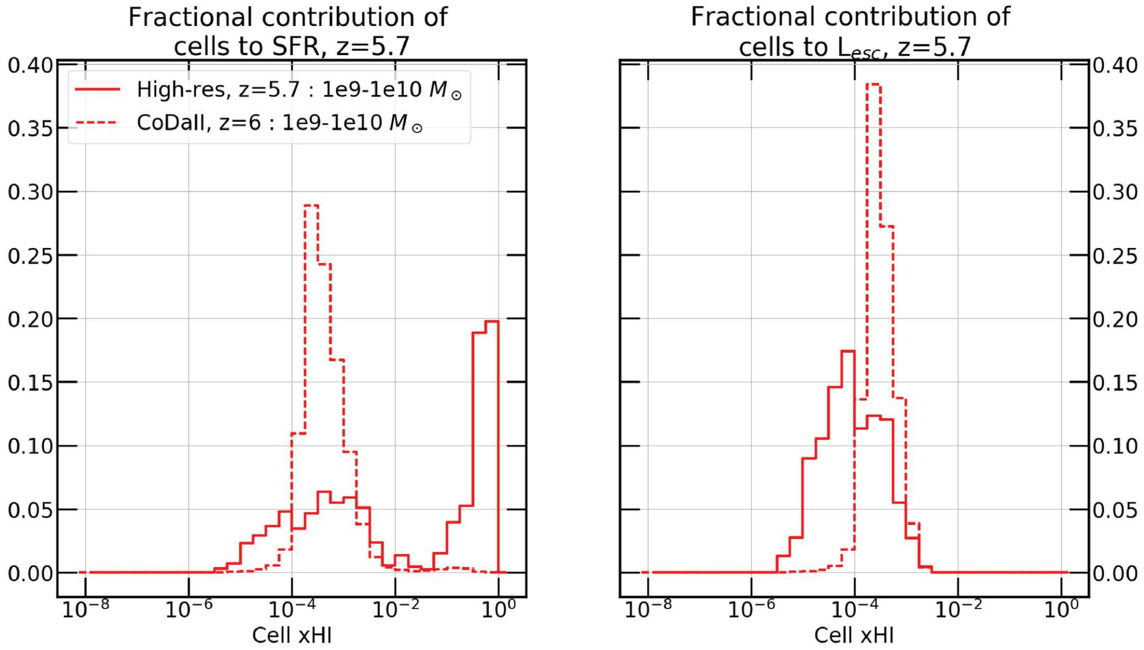


Figure A2. Distribution of star-forming halo cell neutral fractions for haloes of masses $10^{10}\,\mathrm{M}_\odot > M_\mathrm{halo} > 10^9\,\mathrm{M}_\odot$. The full lines show the curves for the high-res simulation, whereas the dashed lines show the curves for the CoDaII simulation. The histograms show the contribution of halo cells to (*left:*) total SFR of the mass bin considered, and to (*right:*) the total escape luminosity of the mass bin considered.

the high-res box are lower suggests that this transmission boost does not compensate the opacity boost occurring at higher resolution, and the latter dominates.

APPENDIX B: ESCAPE FRACTION FITS

In this section, we propose simple functional forms for the both the average of $f_\mathrm{esc}^\mathrm{ray}$ and the SFR-weighted average of $f_\mathrm{esc}^\mathrm{ray}$ as a function of halo mass and of redshift. These could be useful as a $f_\mathrm{esc}^\mathrm{ray}$ model for projects in which the determination of $f_\mathrm{esc}^\mathrm{ray}$ is impossible or difficult: either when working with simulated data that was not produced with fully coupled radiation-hydrodynamics simulation codes, performing radiative transfer in post-processing, or with simulations with lower spatial resolution than CoDa II, and also for semi-analytical models of the EoR (Mesinger et al. 2011; Fialkov et al. 2013).

In all cases, we caution the reader that the functions presented here represent the halo escape fraction, i.e. to obtain the net halo escape fraction of haloes, $f_\mathrm{esc}^\mathrm{net}$, one needs to multiply $f_\mathrm{esc}^\mathrm{ray}$ by $f_\mathrm{esc}^\mathrm{sub}$ (0.42), as show in equation (5).

We provide the reader with PYTHON functions for plotting the fits that follow here (github).

B1 Average escape fraction

Based on the trends presented by the average values of $f_\mathrm{esc}^\mathrm{ray}$ as a function of mass and of redshift between $z = 6$ and $z = 10$ shown in Fig. 4, we opted for a double power law of halo mass, with a knee separating both laws, that shifts with redshift. Due to the noisy appearance of the data for $\mathrm{M_{halo}} > 10^{11}\,\mathrm{M}_\odot$, the time behaviour of $f_\mathrm{esc}^\mathrm{ray}$ within this mass range is unclear, which is why we chose to simply model the evolution of $f_\mathrm{esc}^\mathrm{ray}$ with mass as the extension of the $f_\mathrm{esc}^\mathrm{ray}(M_\mathrm{halo}, z = 10)$ fit, to produce indicative values despite the high amount of noise in the average curves for these masses.

We propose the following functional form (equation B1):

$$f_\mathrm{esc}^\mathrm{ray} = \begin{cases} \mathrm{f_{esc}^{ray,10.1\leq z\leq 14.9}(M_{halo}, z)}; & \text{if } \mathrm{M_{halo} > M_{join}}, \\ \mathrm{f_{esc}^{knee}(z) f^{knee}(M_{halo}, z)}; & \text{if } \mathrm{M_{join} \geq M_{halo} \geq M_{knee}(z)}, \\ \mathrm{f_{esc}^{max}(z) f^{max}(M_{halo}, z)}; & \text{if } \mathrm{M_{halo} < M_{knee}(z)}, \end{cases} \tag{B1}$$

$$\text{with} = \begin{cases} \mathrm{z_6} & = \left(\frac{1+z}{1+6}\right), \\ \mathrm{f_{esc}^{max}(z)} & = \min\left(\mathrm{f_{esc}^{0,max} z_6^{\beta}},\ 1.0\right), \\ \mathrm{f_{esc}^{knee}(z)} & = \max\left(\min\left(\mathrm{f_{esc}^{0,knee} z_6^{\gamma}},\ 1.0\right), \mathrm{f_{min}^{knee}}\right), \\ \Delta^\mathrm{max}(z) & = \frac{\log_{10}(\mathrm{f_{esc}^{knee}(z)}) - \log_{10}(\mathrm{f_{esc}^{max}(z)})}{\log_{10}(\mathrm{M_{knee}(z)}) - \log_{10}(10^8)}, \\ \Delta^\mathrm{knee}(z) & = \frac{\log_{10}(\mathrm{f_{esc}^{knee}(z)}) - \log_{10}(\mathrm{f_{esc}^{join}})}{\log_{10}(\mathrm{M_{knee}(z)}) - \log_{10}(\mathrm{M_{join}})}, \\ \mathrm{M_{knee}(z)} & = 2.5 \times 10^9 z_6^{\zeta}\,\mathrm{M}_\odot, \\ \mathrm{f^{max}(M_{halo}, z)} & = \left(\frac{M_\mathrm{halo}}{\mathrm{M_{knee}(z)}}\right)^{\Delta^\mathrm{max}(z)}, \\ \mathrm{f^{knee}(M_{halo}, z)} & = \left(\frac{M_\mathrm{halo}}{\mathrm{M_{knee}(z)}}\right)^{\Delta^\mathrm{knee}(z)}, \end{cases} \tag{B2}$$

where $f_\mathrm{esc}^\mathrm{max}(z)$ is the maximum value below M_knee, it is $f_\mathrm{esc}^{0,\mathrm{max}}$ at $z = 6$, and evolves as z^β.

$f_{\rm esc}^{\rm knee}(z)$ is where both power laws join at $M_{\rm knee}(z)$, it is $f_{\rm esc}^{0,\rm knee}$ at $z = 6$, and evolves as z^{γ}. It cannot be lower than $f_{\rm min}^{\rm knee}$.

$\Delta^{\rm max}(z)$ is the slope of the power law above $M_{\rm knee}(z)$, it is defined so as to reach $f_{\rm esc}^{\rm max}(z)$ when $\mathrm{M} = 10^8\,\mathrm{M}_{\odot}$.

$\Delta^{\rm knee}(z)$ is the slope of the power law below $M_{\rm knee}(z)$, it is defined so as to reach $f_{\rm esc}^{\rm knee}(z)$ when $\mathrm{M} = M_{\rm knee}(z)$, and $f_{\rm esc}^{\rm join}$ when $\mathrm{M} = M_{\rm join}$.

$M_{\rm knee}(\mathrm{z})$ gives the mass where the power laws join, it is $2.5 \times 10^9\,\mathrm{M}_{\odot}$ at $z = 6$, and evolves with z as z^{ζ}.

$f^{\rm max}(M_{\rm halo},\, z)$ accounts for the mass slope for $M_{\rm halo} < \mathrm{M}_{\rm knee}(\mathrm{z})$ given by $\Delta(z)$.

$f^{\rm knee}(M_{\rm halo},\, z)$ accounts for the mass slope for $M_{\rm halo} \geq \mathrm{M}_{\rm knee}(\mathrm{z})$ given by δ.

The adjusted values are summarized in equation (B3) next:

$$\begin{cases} \mathrm{f}_{\rm esc}^{0,\rm max} & = \quad 0.98\,, \\ \mathrm{f}_{\rm esc}^{0,\rm knee} & = \quad 0.8\,, \\ \mathrm{M}_{\rm join} & = 10^{11}\,\mathrm{M}_{\odot}\,, \\ \mathrm{f}_{\rm esc}^{\rm join} & = \quad 0.1\,, \\ \mathrm{f}_{\rm min}^{\rm knee} & = \quad 0.23\,, \\ \beta & = \quad -0.5\,, \\ \gamma & = \quad -3.5\,, \\ \zeta & = \quad -2.0\,. \end{cases} \tag{B3}$$

The average curves corresponding to $10 \leq z \leq 14.9$ behave somewhat differently. They appear to be simple power laws, offset in escape fraction by ~0.25 dex. Hence, instead of using the previously discussed formulas for this redshift range, we tentatively provide the following fit (equation B4) for $f_{\rm esc}^{\rm ray}$ for $10 \leq z \leq 14.9$.

$$\mathrm{f}_{\rm esc}^{\rm ray,10.1\leq z\leq 14.9}(\mathrm{M}_{\rm halo},\, \mathrm{z}) = \min\left(\mathrm{f}_{\rm esc,0}^{10\leq z\leq 14.9} \times \left(\frac{\mathrm{M}_{\rm halo}}{10^8\,\mathrm{M}_{\odot}}\right)^{\Delta^{10\leq z\leq 14.9}} \times \left(\frac{10.}{z}\right)^{\gamma},\, 1\right), \tag{B4}$$

$$\text{with} = \begin{cases} \mathrm{f}_{\rm esc,0}^{10.1\leq z\leq 14.9} & = 0.77, \\ \Delta^{10.1\leq z\leq 14.9} & = \dfrac{\log_{10}(0.77) - \log_{10}\left(\mathrm{f}_{\rm esc}^{\rm join}\right)}{\log_{10}(10^8\,\mathrm{M}_{\odot}) - \log_{10}(\mathrm{M}_{\rm join})}, \\ \gamma & = -1.64. \end{cases} \tag{B5}$$

Once again, as we aim to represent the averages and not fit the data, we adjust these values by hand.

The aforementioned fits are presented in Fig. B1 (the full, thick lines), where they are also compared to the mass bin averages of $f_{\rm esc}^{\rm ray}$ (the full, thin lines).

B2 SFR-weighted average escape fraction

We also provide fits for the SFR-weighted average values of $f_{\rm esc}^{\rm ray}$ as a function of mass and of redshift between $z = 6$ and $z = 14.9$. As can be seen in Fig. B2, the SFR-weighted average of $f_{\rm esc}^{\rm ray}$ has a slightly different behaviour. Indeed, in this case, the average curves do not join above a certain mass. To fit them, we opted for a slightly different double power law of halo mass, with a knee separating both laws. We propose the following functional form (equation B6):

$$f_{\rm esc} = \begin{cases} \mathrm{f}_{\rm esc}^{\rm knee}(\mathrm{z})\mathrm{f}^{\rm knee}(\mathrm{M}_{\rm halo},\, \mathrm{z});\ \text{if } \mathrm{M}_{\rm halo} \geq \mathrm{M}_{\rm knee}(\mathrm{z}), \\ \mathrm{f}_{\rm esc}^{\rm max}(\mathrm{z})\mathrm{f}^{\rm max}(\mathrm{M}_{\rm halo},\, \mathrm{z});\ \text{if } \mathrm{M}_{\rm halo} < \mathrm{M}_{\rm knee}(\mathrm{z})\,, \end{cases} \tag{B6}$$

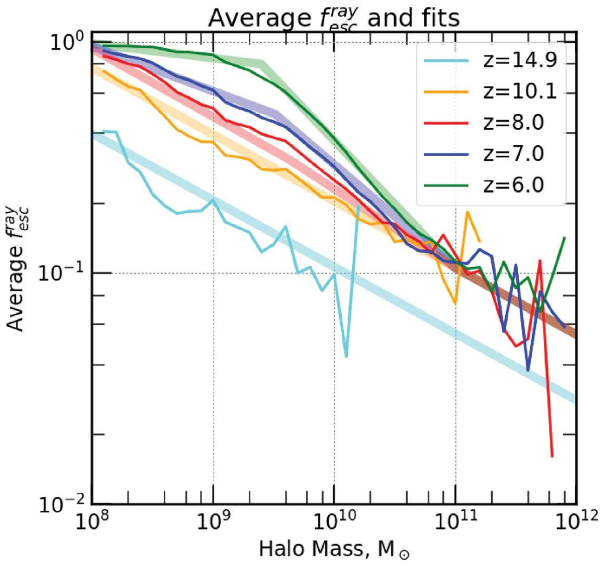


Figure B1. Average escape fractions (thin lines) of star-forming haloes and corresponding fits (thick lines) at z = 6, 7, 8, 10.1, 14.9.

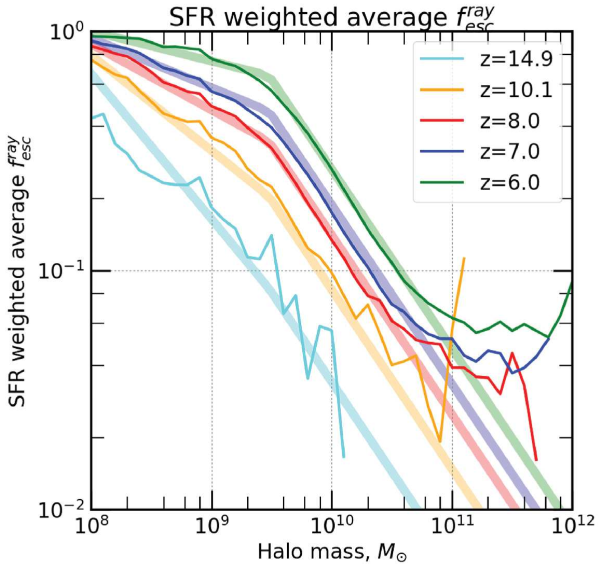


Figure B2. SFR-weighted average escape fractions (the thin lines) of star-forming haloes and corresponding fits (the thick lines) at $z = 6$, 7, 8, 10.1, 14.9.

$$\text{with} = \begin{cases} \mathrm{z}_6 & = \left(\frac{1+z}{1+6}\right), \\ \mathrm{f}_{\rm esc}^{\rm max}(\mathrm{z}) & = \min\left(\mathrm{f}_{\rm esc}^{0,\rm max}\mathrm{z}_6^{\beta},\ 1.0\right), \\ \mathrm{f}_{\rm esc}^{\rm knee}(\mathrm{z}) & = \min\left(\mathrm{f}_{\rm esc}^{0,\rm knee}\mathrm{z}_6^{\gamma},\ 1.0\right), \\ \Delta(z) & = \frac{\log_{10}(\mathrm{f}_{\rm esc}^{\rm knee}(\mathrm{z}))-\log_{10}(\mathrm{f}_{\rm esc}^{\rm max}(\mathrm{z}))}{\log_{10}(\mathrm{M}_{\rm knee}(\mathrm{z}))-\log_{10}(10^8)}, \\ \mathrm{M}_{\rm knee}(\mathrm{z}) & = 3 \times 10^9\,\mathrm{M}_{\odot}, \\ \mathrm{f}^{\rm max}(\mathrm{M}_{\rm halo},\, \mathrm{z}) & = \left(\frac{M_{\rm halo}}{\mathrm{M}_{\rm knee}(\mathrm{z})}\right)^{\Delta(z)}, \\ \mathrm{f}^{\rm knee}(\mathrm{M}_{\rm halo},\, \mathrm{z}) & = \left(\frac{M_{\rm halo}}{\mathrm{M}_{\rm knee}(\mathrm{z})}\right)^{\delta}, \end{cases} \tag{B7}$$

where $f_{\rm esc}^{\rm max}(z)$ is the maximum value below $M_{\rm knee}$, it is $f_{\rm esc}^{0,\rm max}$ at $z = 6$, and evolves as z^{β}.

$f_{\rm esc}^{\rm knee}(z)$ is where both power laws join at $M_{\rm knee}(z)$, it is $f_{\rm esc}^{0,\rm knee}$ at $z = 6$, and evolves as z^{γ}.

$\Delta(z)$ is the slope of the power law below $M_{\rm knee}(z)$, it is defined so as to reach $f_{\rm esc}^{\rm max}(z)$ when $M_{\rm halo} = 10^{8}\,{\rm M}_{\odot}$.

$M_{\rm knee}$(z) gives the mass where the power laws join, it is $3\times10^{9}\,{\rm M}_{\odot}$ at $z = 6$, and evolves with z as z^{ζ}.

$f^{\rm max}(M_{\rm halo}, z)$ accounts for the mass slope for $M_{\rm halo} < {\rm M}_{\rm knee}({\rm z})$ given by $\Delta(z)$.

$f^{\rm knee}(M_{\rm halo}, z)$ accounts for the mass slope for $M_{\rm halo} \geq {\rm M}_{\rm knee}({\rm z})$ given by δ.

Since we want to reproduce the behaviour with mass and with redshift of the SFR-weighted averages of $f_{\rm esc}^{\rm ray}$, and not model the full distribution of points, we again adopt the simple approach of adjusting the fit by hand (as opposed to computing the fit of the model to the data sample).

The adjusted values are summarized in equation (B8) next:

$$\begin{cases} {\rm f}_{\rm esc}^{0,\rm max} & = \quad 1.0\,, \\ {\rm f}_{\rm esc}^{0,\rm knee} & = \quad 0.65\,, \\ \beta & = \ -0.5\,, \\ \gamma & = \ -2.5\,, \\ \delta & = -0.75\,, \end{cases} \tag{B8}$$

The aforementioned fits are presented in Fig. B2 (the full, thick lines), where they are also compared to the SFR-weighted average $f_{\rm esc}^{\rm ray}$ measure in CoDa II (the full, thin lines).

APPENDIX C: COMPUTING ESCAPE FRACTION

Fig. C1 shows a simplified, explanatory drawing of the computation process for $f_{\rm esc}^{\rm ray}$ for an individual star-forming halo cell.

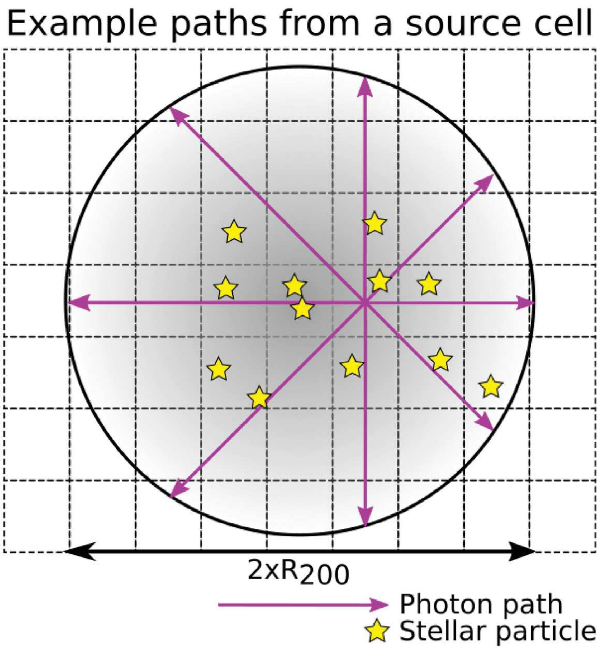


Figure C1. Explanatory drawing of the computation of $f_{\rm esc}^{\rm ray}$. For each cell containing an emitting particle, we trace 768 rays from its centre to a sphere of radius r_{200} centred on the halo centre given by the fof halo finder. By averaging this result for every emitting cell and weighting by SFR, we obtain the halo value $f_{\rm esc}^{\rm ray}$ at r_{200}.

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

1.2.6 The limits of the definition of galaxies in CoDa II

As reported in Ocvirk et al. [2020], dark matter haloes are detected using the friends of friends algorithm pFOF [Roy et al., 2014] to create a catalogue of halo masses and positions. In this work, we assume that every dark matter halo contains one galaxy, which is true in most cases. An oft used convention to define galactic limits is to equate the virial radius of a galaxy to its r_{200}, or the radius that contains a density two hundred times greater than the average dark matter density (Computed using Eq. 1.1).

$$\mathrm{M}_{\mathrm{halo}} = \frac{4}{3}\pi\, \mathrm{r}_{200}^{3} \times 200\, \bar{\rho}_{\mathrm{DM}}\,, \tag{1.1}$$

This simple approach is not without its drawbacks (especially at high redshift) , as not all galaxies are fully-virialised within r_{200} , nor are they all perfectly captured by a spherical volume (and are often triaxial for example as shown in Ocvirk et al. [2020]). This being said, throughout the rest of the manuscript we shall consider the limits of galaxies (and the beginning of the IGM) to be at r_{200} around every dark matter halo containing a galaxy.

Over-linking can be an issue with our use of pFoF. Where there are two nearby dark matter haloes, pFoF may link the particles of both of them, which rather unfortunately means that the centre of this dual halo ends up between the potential wells of the halos, and thus far from the maximal density of baryons. The result is a massive dark matter halo whose r_{200} contains none or little of the baryonic matter of both galaxies. Fig. 1.12 gives an example of this for a detected halo of $1.59 \times 10^{10}\,\mathrm{M}_{\odot}$. In cases such as these, the result of $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ is difficult to interpret. It tends to be higher than a typical $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ value from the same mass range, as the paths sampled to compute $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ are in the IGM or the outer limits of the CGM which is optically thin for LyC photons in massive haloes. The worst case scenario, wherein no active stellar particles are in r_{200}, is easy to detect. This appears to only occur to less than one percent of haloes with $\mathrm{M}_{\mathrm{halo}} > 5 \times 10^{9}\,\mathrm{M}_{\odot}$. However, it is difficult to estimate the total number of problematic cases. One solution is to examine the spatial distribution of the associated dark matter particles. In the case of two separate linked structures, one can expect an elongated shape of dark matter. Although this is the case, it is non-trivial to separate high largest axis over-linked detected haloes from actual elongated haloes. Ultimately, it would be desirable to use a halo finding algorithm that uses the dark matter particle velocities as well as their positions. Indeed, using full phase space information could eliminate some of the cases where two near or merging haloes are interpreted as a single massive halo.

1.2.7 On the computation of the escape fraction

Computation by integrating the photon flux

In Lewis et al. [2020], we used the state of the Hydrogen gas to compute $\mathrm{f}_{\mathrm{esc}}$ in CoDa II galaxies. However, since RAMSES-CUDATON tracks the flux of ionising photons in each cell, the most straightforward and intuitive approach to computing the escape fractions of dark matter haloes in CoDa II, is to directly measure the photons that escape r_{200} by integrating the photon fluxes on shells of radii

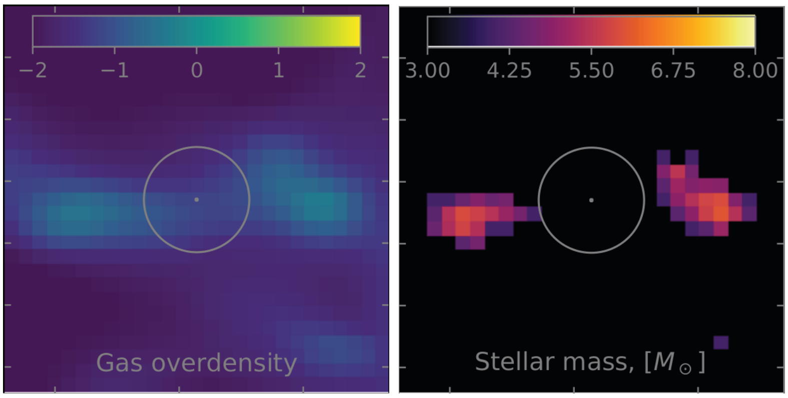


Figure 1.12: Two maps of a case of overlinking by pFoF (integrated over $2\times$ r_{200} of the detected halo along the Z direction). Each cell is 3.3 ckpc wide. *Left:* Map of the H gas density in units of over-density. *Right:* Map of the density of stellar particles in $M_\odot$ per cell.

r_{200} as in Eq. 1.2.

$$L_{esc} = \sum_i \tilde{F}_i.\tilde{S}_i \,, \tag{1.2}$$

Where L_{esc} is the rate of escaping photons or escape luminosity, $\tilde{F}_i$ is the flux vector of the cell containing the surface element i, and $\tilde{S}_i$ is the outwards direction surface vector of the surface element i.

In practice we use the python healpy module to create a discrete partition of r_{200} radius shells into N_{surf} surfaces of equal surface. We do not perform an interpolation of the flux values, and each surface is associated with the flux of the cell that contains it. To obtain the escape fraction, we must then divide the obtained rate of escaping photons by the production rate of photons within r_{200}. We compute this second number by summing the masses of the stellar particles young enough[9] to emit ionising radiation and associated with a halo, taking care to weight each mass by the fraction of its lifetime that it spent over 10 Myr if it is older (This step makes sure that we match the output of the code, by guaranteeing that each stellar particle emits a fixed amount of photons per mass over its real life time). Finally, we can convert this stellar mass into a luminosity by multiplying by the stellar emissivity used in CoDa II and the sub-grid escape fraction (0.42 in CoDa II).

$$L_{intr} = \sum_{\text{stars} \in r_{200},\, \text{ages}<10+\Delta T} \Big(M_\star \times \min(1.0, \frac{10.0-(\text{Age}-\Delta T)}{\Delta T})\Big) \times \epsilon \times f^{sub}_{esc} \,, \tag{1.3}$$

Where L_{intr} is the intrinsic luminosity, $M_\star$ is the stellar mass in $M_\odot$, ϵ the stellar emissivity or the number of photons produced per second per solar mass of stellar mass ($4.3\times10^{46}\,\mathrm{ph}\cdot\mathrm{s}^{-1}\cdot M_\odot^{-1}$ in CoDa II), ΔT is the length of the last time step in Myr, Age is the current age of the stellar particle in Myr, and 10.0 Myr is the time spent emitting ionising photons for every stellar particle, and f^{sub}_{esc} is the sub-grid escape fraction (0.42 in CoDa II). One then has that $f^{flux}_{esc} = L_{esc}/L_{intr}$.

We pick N_{surf}=768 equal healpy surfaces onto which we perform our integrations. Doubling this

---

[9] younger than 10+ΔT Myr, where ΔT is the length of the last time step

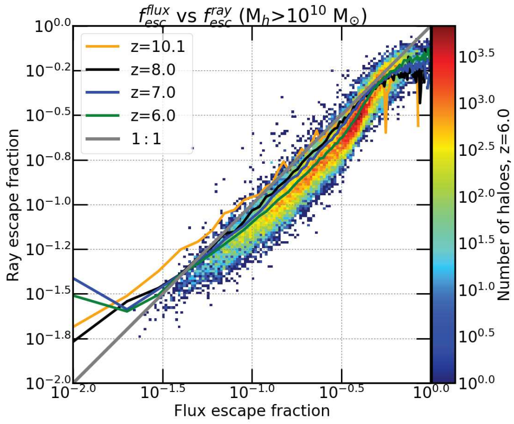


Figure 1.13: A comparison between flux based escape fractions and ray tracing based escape fractions at several redshift and for $M_{halo}>10^{10}\,M_\odot$ haloes.

number only results in changes at the 10^{-4} level for the most massive halo, that requires the best angular resolution. This precision is satisfactory as most of our f_{esc} results are within 10^{-2} to 1 (At high redshifts, low mass haloes can have much lower f_{esc} from 10^{-2} to 0, but they represent a small fraction of the total population).

Why we pick the ray tracing method

Fig. 1.13 compares the results obtained by integrating halo fluxes with the results yielded by our ray tracing method for $M_{halo}>10^{10}\,M_\odot$ haloes, that are the best resolved. Overall there is a good agreement between the two methods, that appears to degrade for low f_{esc} ($<10^{-1.5}$) but the significance of these data is poor due to their small population. A lesser break from the agreement occurs for the highest f_{esc} values $>10^{-0.2}$, for which the ray tracing method values plateau at around $10^{-0.1}$(around 0.8), whereas the flux based values increase till 1. This plateau increases slightly with time, possibly following box ionisation, reducing the severity of the disagreement, but never reaching 1. Elsewhere, the agreement deteriorates as the mid-range ray tracing values slightly decrease with time. At lower masses, the agreement between the two methods significantly deteriorates, as the flux integration method yields results that increase sharply with halo radius (and thus mass) at constant L_{intr} until $r_{200}\approx 1.5$ (mass $\approx 2\times 10^9\,M_\odot$). This stems from the implementation of the M1 closure used in ATON over such a small (< 3 cell sided cubes) number of cells. The ray tracing method has no such particularity.

The M1 closure method for computing solutions to the radiative transfer equations has many benefits, but comes with a specific disadvantage that is troublesome when computing f_{esc}. Indeed, in the M1 method RT, simulated photon fluxes no longer cross one another, but instead collide like a fluid. In the case of two (or more) close or closing galaxies, this collision of their outward fluxes can affect the measurement of L_{esc} via flux integration. As shown illustratively in Fig. 1.14, if the point where the fluxes collide is near to one of the integration surfaces, then the result of the integration can change, as the colliding photons tend to emerge from their point of collision in a perpendicular direction to their original propagation direction and therefore somewhat more tangentially (depending on where the opposing fluxes meet in relation to the integration surfaces) to the surface of the r_{200}

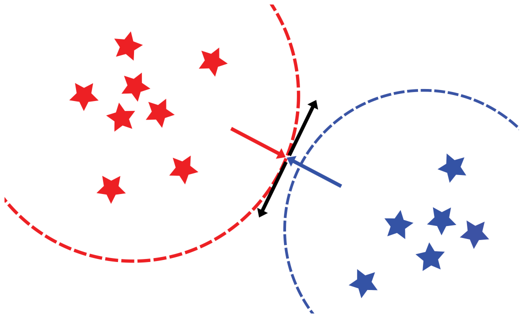


Figure 1.14: Drawing illustrating the collisional behaviour of photons when using the M1 method, and the issues it can cause when integrating the fluxes of galaxies on spheres. Stars represent stellar particles, arrows radiative fluxes, and the dashed lines r_{200}. In this particular case where two star forming galaxies are near one another (possibly merging) the point where the fluxes collide can be near or on the r_{200} of one of the galaxies. In this case, the flux seen exiting the red galaxy is lower than one might otherwise expect as the collisional behaviour of the M1 photons renders the flux at r_{200} tangential or close to tangential to the red galaxy's integration surface.

shell. This problem is known to the community, and ray tracing methods can be used to side step it entirely as in Trebitsch et al. [2020a].

A further drawback of the flux based method comes from geometrical considerations, as illustrated by Fig. 1.15: for a halo of a given SFR and mass, the r_{200} sphere may intersect cells containing flux from stellar particles that are outside r_{200}, leading to higher (lower) f_{esc} values when a particle is counted and its integrated flux is greater than L_{intr} (when a particle is counted and its integrated flux is null or lower than L_{intr}). For small mass haloes, that have few stellar particles this can affect the f_{esc} dramatically, sometimes giving completely non-physical >1 values.

For all of the reasons above, we make the decision to use the ray tracing method for computing L_{esc} and escape fractions, as it provides a better result with good precision across the full sample of haloes.

1.2.8 SFR weighted average escape fractions

Although the fit of the SFR weighted average is given in Lewis et al. [2020], we made no comment as to its shape or redshift evolution. Fig. 1.16 shows the SFR weighted average f_{esc}^{ray} as a function of halo mass for z=14.9,10.1,8,7,6. The SFR weighted average curves decrease with increasing halo mass just as the number weighted ones do in Lewis et al. [2020]. However, weighting by SFR drastically changes the shape of the curves beyond $10^{10}\,M_\odot$: above a few $10^{9}\,M_\odot$ the SFR weighted averages decrease with M_{halo} at a much faster rate than the number weighted curves, yielding lower values. For instance, at z=6, the SFR weighted average of f_{esc}^{ray} is ≈0.06 for haloes of $10^{11}\,M_\odot$, versus ≈ 0.1 for the number weighted average at the same mass. The higher the redshift, the greater this effect. Indeed, whereas the number weighted average f_{esc}^{ray} of the highest mass haloes did not vary with redshift, the SFR weighted average f_{esc}^{ray} increases with decreasing redshift at a rate similar to the time evolution of other masses. These differences between the number weighted case and the SFR weighted case suggest

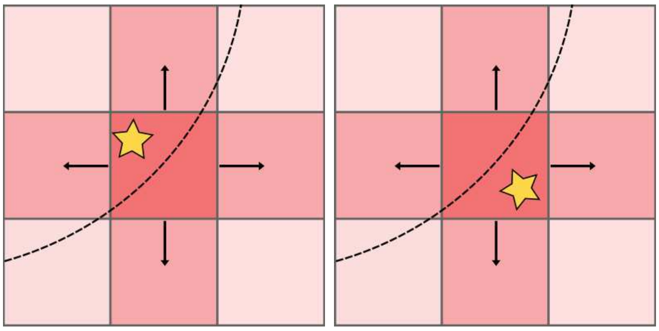


Figure 1.15: Two schematic examples of the possible problems arising when using the flux integration method. The red squares represent the source field in 2D; the higher their photon density, the redder their colour. Yellow stars represent stellar particles. The black arrows represent the flux exiting the central source in 2D. Finally the black line represents r_{200}. In the left panel, f_{esc} will be lower than expected, as L_{intr} will include the stellar particle, and since a fraction of the source's flux is not captured by our r_{200} sphere surfaces. In the right panel, we would obtain a higher f_{esc} measurement as L_{intr} does not include the stellar particle, and a fraction of its flux will leave the r_{200} sphere and increase L_{esc}.

some correlation between f_{esc}^{ray} and SFR, which we explore in Sec. Scatter of f_{esc}values around the mass averages.

A commonly computed statistic in Reionization simulations is the global average f_{esc} weighted by star formation rate. Fig. 1.17 shows this average between z=14.9 and z=6 along side the equivalent curves from Kimm and Cen [2014]; Katz et al. [2018]. The displayed curves are similarly flat and close to ≈ 0.1. Our average is below the two others (≈ 0.08), but remains close. The similarity of our results is somewhat astonishing as the two other simulations in our comparison have much better spatial resolutions. The flatness of the SFR box weighted average of f_{esc}^{ray} is relatively surprising, as we have shown that over time the f_{esc}^{ray} of low mass haloes increases substantially. At the same time, the progressive build up of high mass haloes which dominate the SFR and have low f_{esc}^{ray} must pull down the SFR weighted average in the opposite direction. By reducing the number of low mass star forming haloes, star formation suppression also acts to reduce the average SFR weighted $f_{esc}^{ray}$by "removing" low mass high f_{esc}^{ray} haloes from the average.

1.2.9 Scatter of f_{esc}values around the mass averages

For a given mass, we find a large range of f_{esc}^{ray} values, indicating that there are other parameters that play a role in setting f_{esc}^{ray} . The left panel of Fig. 1.18, shows the average f_{esc}^{ray} of haloes with only one stellar particle as a function of its age, for several redshifts. For all redshifts, the average f_{esc}^{ray} of single stellar particle haloes increases gradually with stellar age. This is a very clear sign that the ionisation of a low mass halo by a single stellar particle that forms within it is not instantaneous. For redshifts of 14.9 to 7.0, f_{esc}^{ray} roughly doubles before 1 Myr of age, after which the increase is much more gradual, eventually reaching a maximum towards ≈ 6 Myr. As redshift decreases, the average f_{esc}^{ray} increases substantially for all particle ages : at z=14.9 the maximum average f_{esc}^{ray} is around 0.25, whereas at

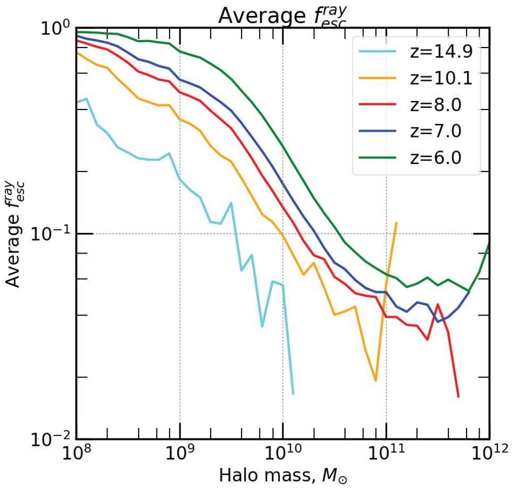


Figure 1.16: SFR weighted average $f_{\rm esc}^{\rm ray}$ as function of mass for z=14.9,10.1,8,7,6.

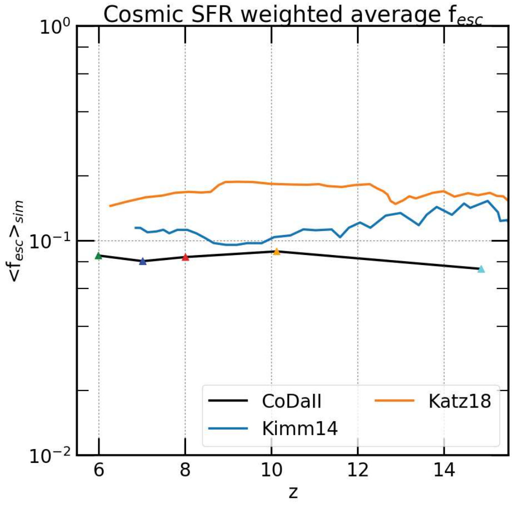


Figure 1.17: Simulation SFR weighted average $f_{\rm esc}$ in CoDa II as a function of redshift (black lines). Also plotted are the same curves from Kimm and Cen [2014]; Katz et al. [2018].

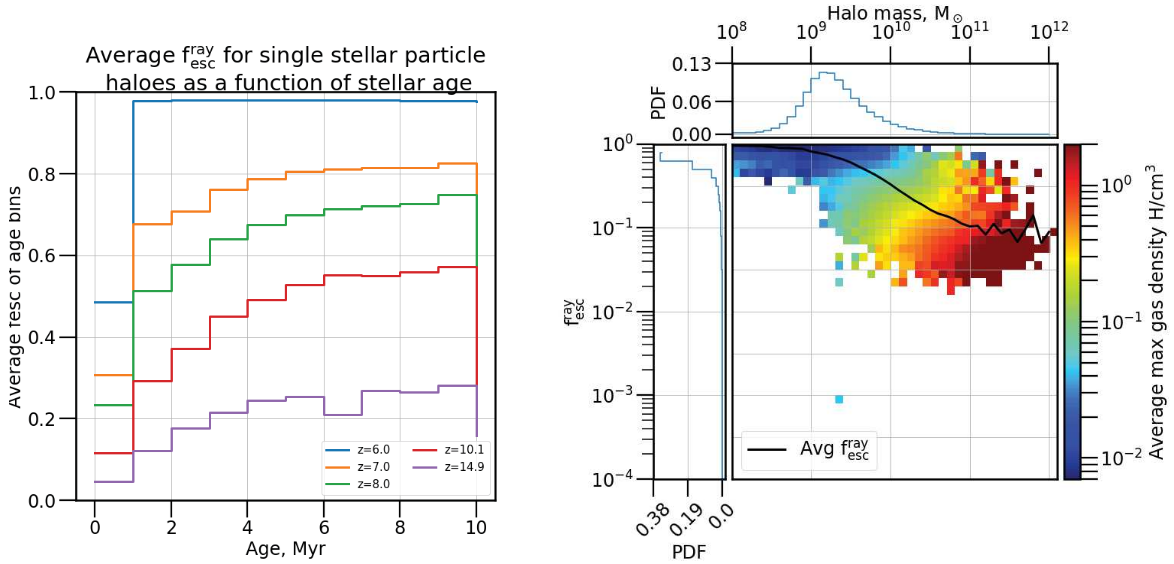


Figure 1.18: *Left*: Average single stellar particle halo f_{esc}^{ray} as a function of stellar age. *Right*: Average halo max gas density as a function of halo mass and escape fraction at z=6. The black lines represents the number weighted average f_{esc}^{ray} for every mass bin.

z=6 it is 1.0. By z=6, the initial sharp increase yields an average f_{esc}^{ray} of roughly 1.0. It could be that this increase is due to cosmological expansion of simulation grid in two ways: Firstly, once a star forms, and since optical depth scales as $\rho_{gas}dx \propto a^{-2}$, the optical depths encountered by photons are smaller, and the measured f_{esc}^{ray} is larger. Secondly, because the density criterion for star formation is co-moving then as time goes on, stars can form in cells with lower and lower physical gas densities where the opacity to ionising photons is lower. The increase between z=7 and z=6 is greater than the increase between the other successive redshifts. Coincidentally, it is between these two redshifts that the overlap of ionised regions in CoDa II finishes. It is possible, then, that this surge in average f_{esc}^{ray} in single stellar particle haloes is due to an increase in the UVBG, or to the UVBG reaching more of these haloes.

For more massive objects that can have a multitude of stellar particles, one very plausible origin for this scatter is variation in the neutral gas mass and the neutral gas density distribution. It may be that for a given halo mass some galaxies have denser more opaque centres (optical depth scales with the gas density, as does the recombination rate, further augmenting the optical depth). Such differences in the gaseous morphology could arise due to environmental effects : indeed, some galaxies may form in more gas poor(rich) environments leading to lower(greater) accumulations of gas over the galaxy's lifetime. Further, some regions may experience feedback events intense enough to heat the IGM gas sufficiently to significantly hinder the accretion of gas. At the same time, galaxies forming in denser regions may be more likely to have massive neighbours that might disrupt their gas content through tidal effects or close encounters, or more directly when major mergers occur. The right panel of Fig. 1.18 shows the average of every bin halo's maximum gas density for several bins of halo mass and f_{esc}^{ray} at z=6. The black line indicates the number weighted average f_{esc}^{ray} as a function of mass that was found in Lewis et al. [2020]. For haloes more massive than 10^9 $M_\odot$, haloes with higher than average f_{esc}^{ray} have lower maximum gas densities on average. For haloes with under-average f_{esc}^{ray}, the opposite is true. This indicates that the scatter in f_{esc}^{ray} values for a given halo mass is at least partially driven by differences in the internal distribution of gas within haloes. For instance, for halo masses of $10^{10}\,M_\odot$, f_{esc}^{ray} varies between 1.0 and 3×10^{-2}.

1.2.10 Time integrated ionising photon budget

The photon budget in Fig. 1.11 only gives an instantaneous view of the drivers of Reionization. Determining the overall dominant contributors to Reionization requires integrating the instantaneous contribution over the duration of the EoR. As we only performed our analysis of the ionising photon budget at certain redshifts (z=14.9, 10.1, 8, 7, 6), we perform a simple linear interpolation between these time steps in order to determine the contribution of haloes of various masses as a function of time, and the total time integrated ionising photon budget during Reionization. Fig. 1.19 shows the results of the interpolation. The left panel shows the total emissivity of ionising photons, emitted by several mass ranges of haloes at each chosen time step. It serves as an interesting complement to the photon budget in Fig. 8 of Lewis et al. [2020]. At z=14.9, haloes of 10^8 to 10^9 $M_\odot$ contribute about 0.1 dex more than those of 10^9 to 10^{10} $M_\odot$. At the same time, haloes with masses in 10^{10} to 10^{11} $M_\odot$ contribute close to 2 dex less. Haloes more massive than $10^{11}\,M_\odot$ appear between z=14.9 and z=10.1 (probably later rather than sooner, however without intermediate snapshots it is impossible to determine). The higher the mass of the bin the greater the slope with redshift is. By z=10.1, the 10^9 to 10^{10} $M_\odot$ haloes have become the dominant contributors, surpassing the 10^9 to 10^{10} $M_\odot$, 10^{10} to 10^{11} $M_\odot$, and 10^{11} to 10^{12} $M_\odot$ bin by 0.3, 0.5 and 2.4 dex respectively. After z=10.1, the slopes of all of the bins decrease, however none so much as the slope of the 10^8 to 10^9 $M_\odot$ bin, which even becomes negative after z=7. Decreases in these curves result either from a decrease in f_{esc}^{ray} or a decrease in SFR within the mass bin. Since we have shown in Lewis et al. [2020], that the average CoDa II f_{esc}^{ray} increases on average with time for all but the highest masses, the dramatic change in the slope of emissivity for the least massive bin seems to point to a decrease in SFR, and is another sign of star formation suppression ongoing during Reionization. At z=6, 10^8 to 10^9 $M_\odot$ is the least emissive bin, having been narrowly overtaken by the 10^{11} to 10^{12} $M_\odot$ bin, which has the highest slope at this point. The dominant contributor at z=6 is the 10^9 to 10^{10} $M_\odot$ bin, which is also the overall dominant contributor with $2.5 \times 10^{50}\,\mathrm{ph \cdot s^{-1} \cdot Mpc^{-3}}$ (this is expected since it contains the masses which are the modes of the galactic ionising photon budget between z=10.1 and z=6). The second and third most emissive bins are the 10^{10} to 10^{11} $M_\odot$ and 10^8 to 10^9 $M_\odot$ bins, that have emissivities of $\approx 2 \times 10^{50}\,\mathrm{ph \cdot s^{-1} \cdot Mpc^{-3}}$ and $4 \times 10^{49}\,\mathrm{ph \cdot s^{-1} \cdot Mpc^{-3}}$, respectively at z=6.

The right panel shows the total density of ionising photons emitted between z=14.9 and z=6 as a function of halo mass. The slightly skewed bell shape of the curve is highly reminiscent of the galactic ionising photon budget from z=8 to z=6. Its width and mode are very similar too, the latter being situated at around $3 \times 10^9\,M_\odot$. The maximum value is just above $2 \times 10^{64}\,\mathrm{ph \cdot Mpc^{-3}}$, and the curve falls to around $6 \times 10^{62}\,\mathrm{ph \cdot Mpc^{-3}}$ at $10^8\,M_\odot$, and at $8 \times 10^{11}\,M_\odot$. This relative decrease of almost two dex is also similar to that that is shown in Fig. 8 of [Lewis et al., 2020]. Overall, the total amount of ionising photons that escape galaxies in CoDa II during Reionization is dominated by those that

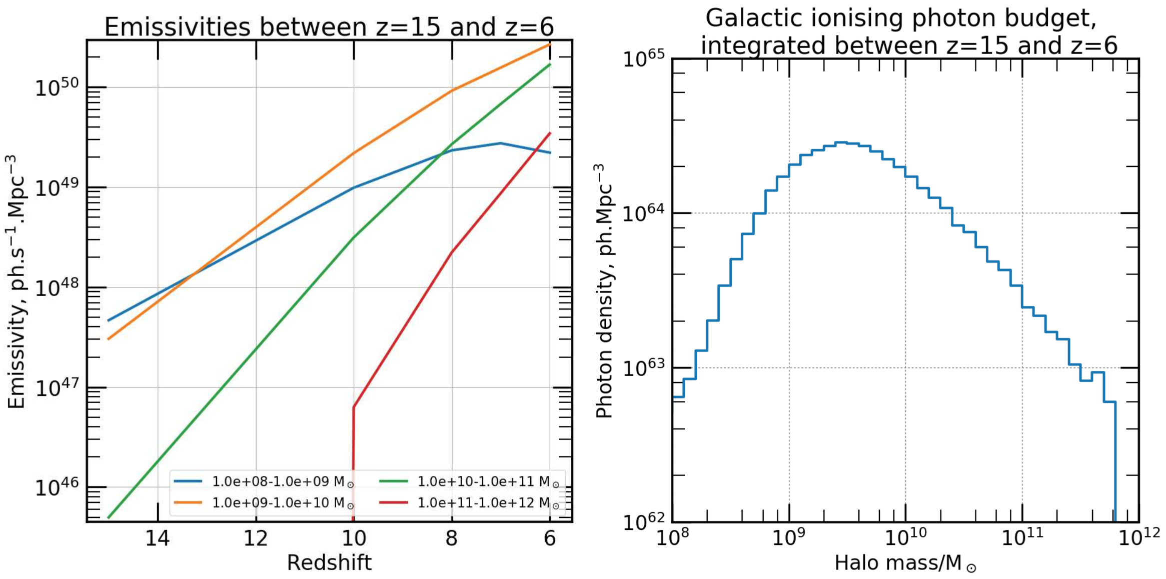


Figure 1.19: *Left:* Total galactic ionising photon budget for several mass ranges for the studied epochs *Right:* Total time integrated galactic ionising photon budget between z=14.9 and z=6.

escaped towards the end of Reionization, that were mostly produced within galaxies of 3×10^8 to 6×10^{10} M$_\odot$.

1.3 Résumé des résultats

Cosmic Dawn II Dans ce chapitre j'ai présenté la simulation Cosmic Dawn II [Ocvirk et al., 2020], une simulation de la Réionisation à très grand échelle (94.44 cMpc par côté) avec une grande résolution spatiale pour une telle simulation (4096 éléments de résolution par côté ou 4096^3 éléments au total). Sa grande taille lui permet de représenter un ensemble statistique de galaxies pendant la Réionisation, évitant des biais dû à la variance cosmique. En même temps, sa résolution permet l'étude des galaxies qui forment des étoiles grâce au refroidissement dû au rayonnement d'atomes d'hydrogène (soit les galaxies dont la masse en matière noire est $\gtrsim 10^8\,\mathrm{M}_\odot$, à cette époque). De plus, dans CoDa II le rayonnement et le gaz sont couplés permettant une description réaliste de la contre-réaction du rayonnement stellaire ionisant avec le gaz, très importante pour représenter la suppression de la formation stellaire à l'époque de la Réionisation. Dû au coût numérique considérable du transfer radiatif (la condition de courant impose des pas de temps très petits à cause de la vitesse de la lumière, très grande), on retrouve souvent des vitesses de la lumière réduites ou variables dans la littérature [Comme dans Rosdahl et al., 2018; Katz et al., 2018; Trebitsch et al., 2020]. De telles approximations peuvent avoir des répercussions sur la vitesse des fronts ionisants [Deparis et al., 2019], et sur l'ionisation moyenne du gaz d'Hydrogène [Ocvirk et al., 2019]. CoDa II permet de contourner ce problème grâce aux gains de performance obtenus avec l'algorithme hybride du code RAMSES-CUDATON, répartissant les calculs de transfert du rayonnement et de chimie de photo-ionisation sur GPUs, et les calculs d'hydrodynamique et de dynamique de la matière noire sur CPUs Ocvirk et al. [2016, 2020]). Finalement, la simulation CoDa II reproduit avec succès certaines des contraintes principales sur la Réionisation (notamment, la fonction de luminosité ultra-violette des galaxies). C'est l'ensemble de ces qualités qui a motivé le choix de CoDa II pour étudier le budget des photons ionisants de galaxies pendant la Réionisation.

Production de photons ionisants des galaxies Dans CoDa II, les galaxies produisant le plus de photons pendant la Réionisation sont entre 2×10^9 et $10^{11}\,\mathrm{M_\odot}$. Par exemple, à z=7 quand la Réionisation avance très rapidement, les galaxies de masse $\approx 10^{10}\,\mathrm{M_\odot}$ produisent plus de 10 fois plus de photons ionisants que les galaxies $\mathrm{M_{halo}}<6 \times 10^8\,\mathrm{M_\odot}$ et $\mathrm{M_{halo}}>2 \times 10^{11}\,\mathrm{M_\odot}$. Les premiers sont trop peu massifs et forment des étoiles trop peu efficacement, tandis que les seconds sont trop rares malgré leur grandes masses et leur formation stellaire efficace. Au fur et au mesure que la Réionisation avance, cet intervalle en masse se décale vers les grandes masses, avant tout à cause de l'apparition de galaxies de plus en plus massives, mais aussi à cause de la suppression de formation stellaire dans les galaxies peu massives.

Fraction d'échappement des galaxies Pour aboutir au budget de photons ionisants des galaxies il faut prendre en compte l'absorption des photons ionisants par le gaz d'Hydrogène neutre contenu dans les galaxies. Pour ce faire, j'ai calculé les fractions d'échappement des galaxies dans CoDa II. La fraction d'échappement d'une galaxie, ou $\mathrm{f_{esc}}$ est la fraction de lumière ionisante produite qui parvient à s'échapper d'une galaxie vers le milieu intergalactique[10]. Pour calculer $\mathrm{f_{esc}}$ j'ai utilisé l'état du gaz pour calculer des épaisseurs optiques le long de rayons tirés entre les particules stellaires et $\mathrm{r_{200}}$.

J'ai trouvé que les galaxies de faibles masses avaient des grandes valeurs de $\mathrm{f_{esc}}$, proches de 1 à $z = 6$ alors que le $\mathrm{f_{esc}}$ des galaxies les plus massives était autour de 0.1 (signifiant que ≈ 90 % de la lumière produite est absorbée avant d'atteindre le milieu intergalactique). Cette tendance est aussi présente dans d'autres travaux tels que Razoumov and Sommer-Larsen [2010]; Wise et al. [2014]; Kimm and Cen [2014]; Paardekooper et al. [2015]; Katz et al. [2018] . J'ai montré que dans CoDa II cette décroissance moyenne de $\mathrm{f_{esc}}$ en fonction de la masse trouvait son origine dans l'ionisation des sites de formation stellaire. En effet, dans les galaxies plus massives, le gaz est plus dense et il parvient plus facilement à recombiner (en dépit de la plus grande formation stellaire qui s'y trouve), réduisant l'ionisation du gaz et donc sa transparence aux photons ionisants.

Dans les galaxies de moins de $\lesssim 10^{11}\,\mathrm{M_\odot}$, $\mathrm{f_{esc}}$ augmente au cours de la Réionisation. En étudiant cette évolution dans les galaxies à une seule particule stellaire, j'ai proposé comme explication l'effet combiné de l'expansion des cellules et de l'utilisation d'un critère en densité physique (et non comobile), donnant lieu à des galaxies avec la même luminosité ionisante mais des cellules de gaz de moins en moins dense et plus transparente aux photons ionisants, et ainsi aboutissant à des valeurs de $\mathrm{f_{esc}}$ qui croissent en moyenne avec le temps.

Finalement, j'ai montré qu'à masses égales, le $\mathrm{f_{esc}}$ d'une galaxie massive était inversement corrélé au maximum de densité de gaz de la galaxie. Ce résultat met en évidence l'importance de la distribution de gaz dans le résultat de $\mathrm{f_{esc}}$ d'une galaxie, suggérant un effet de l'environnement de formation et de l'histoire d'accrétion de gaz des galaxies sur leur $\mathrm{f_{esc}}$.

---

[10] Pour être cohérent avec les autres résultats CoDa II, j'ai suivi la convention de prendre $\mathrm{r_{200}}$, ou le rayon englobant une densité de matière noire 200 fois la moyenne, comme la limite des galaxies.

Budget galactique des photons ionisants En combinant la production de photons ionisants des galaxies et leur f_{esc} j'ai montré qu'une classe de masse et de luminosité intermédiaire ($6 \times 10^8\,M_\odot < M_{halo} < 3 \times 10^{10}\,M_\odot$, ou $-19 < M_{AB1600} < -13$) domine le budget de photons ionisants à z=7. Les faibles f_{esc} des galaxies massives ($\gtrsim 3 \times 10^{10}\,M_\odot$) réduit leur contribution au budget des photons ionisants, qui favorise d'avantage les galaxies peu massives et peu brillantes que ce que la production de photons ionisants laissait imaginer. Ce résultat est similaire à celui mis en évidence dans Katz et al. [2018, 2019]. Poussé par l'évolution de la production de photons ionisants, l'intervalle en masse dominant le budget de photons ionisant augmente en masse d'environ 0.5 dex au cours de la Réionisation.

BIBLIOGRAPHY

Atek H., et al.. The extreme faint end of the UV luminosity function at z $\sim$ 6 through gravitational telescopes: a comprehensive assessment of strong lensing uncertainties. *Mon Not R Astron Soc*, 479(4):5184–5195, 2018.

Aubert D. and Teyssier R. A radiative transfer scheme for cosmological reionization based on a local Eddington tensor. *Mon Not R Astron Soc*, 387(1):295–307, 2008.

Aubert D., et al.. The Inhomogeneous Reionization Times of Present-day Galaxies. *The Astrophysical Journal Letters*, 856:L22, 2018.

Bouwens R.J., et al.. Reionization After Planck: The Derived Growth of the Cosmic Ionizing Emissivity Now Matches the Growth of the Galaxy UV Luminosity Density. *The Astrophysical Journal*, 811(2):140, 2015.

Dawoodbhoy T., et al.. Suppression of star formation in low-mass galaxies caused by the reionization of their local neighbourhood. *Monthly Notices of the Royal Astronomical Society*, 480(2):1740, 2018.

Deparis N., et al.. Impact of the reduced speed of light approximation on ionization front velocities in cosmological simulations of the epoch of reionization. *Astronomy and Astrophysics*, 622:A142, 2019.

Fan X., et al.. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132:117–136, 2006.

Finkelstein S.L., et al.. The Evolution of the Galaxy Rest-Frame Ultraviolet Luminosity Function Over the First Two Billion Years. *ApJ*, 810(1):71, 2015.

Gnedin N.Y. On the Proper Use of the Reduced Speed of Light Approximation. *The Astrophysical Journal*, 833:66, 2016.

Iliev I.T., et al.. Simulating cosmic reionization: how large a volume is large enough? *Mon Not R Astron Soc*, 439(1):725–743, 2014.

Katz H., et al.. A Census of the LyC Photons that Form the UV Background During Reionization. *arXiv:1802.01586 [astro-ph]*, 2018. ArXiv: 1802.01586.

Katz H., et al.. Tracing the sources of reionization in cosmological radiation hydrodynamics simulations. *Monthly Notices of the Royal Astronomical Society*, 483:1029–1041, 2019.

Kimm T. and Cen R. Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars. *The Astrophysical Journal*, 788:121, 2014.

Lewis J.S.W., et al.. Galactic ionising photon budget during the Epoch of Reionisation in the Cosmic Dawn II simulation. *Monthly Notices of the Royal Astronomical Society*, 2020.

Ocvirk P., et al.. HIGH-RESOLUTION SIMULATIONS OF THE REIONIZATION OF AN ISOLATED MILKY WAY-M31 GALAXY PAIR. *ApJ*, 777(1):51, 2013.

Ocvirk P., et al.. Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe. *Monthly Notices of the Royal Astronomical Society*, 463(2):1462–1485, 2016.

Ocvirk P., et al.. Impact of the reduced speed of light approximation on the post-overlap neutral hydrogen fraction in numerical simulations of the epoch of reionization. *Astronomy and Astrophysics*, 626:A77, 2019.

Ocvirk P., et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 2020.

Paardekooper J.P., et al.. The First Billion Years project: the escape fraction of ionizing photons in the epoch of reionization. *Mon Not R Astron Soc*, 451(3):2544–2563, 2015.

Planck Collaboration, et al.. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, 1807:arXiv:1807.06209, 2018.

Razoumov A.O. and Sommer-Larsen J. Ionizing Radiation from z = 4–10 Galaxies. *ApJ*, 710(2):1239, 2010.

Rosdahl J., et al.. The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Monthly Notices of the Royal Astronomical Society*, 479:994–1016, 2018.

Roy F., et al.. pFoF: a highly scalable halo-finder for large cosmological data sets. *A&A*, 564:A13, 2014.

Sorce J.G., et al.. Cosmicflows Constrained Local UniversE Simulations. *Monthly Notices of the Royal Astronomical Society*, 455(2):2078, 2016.

Teyssier R. Cosmological hydrodynamics with adaptive mesh refinement - A new high resolution code called RAMSES. *A&A*, 385(1):337–364, 2002.

Trebitsch M., et al.. Modelling a bright z = 6 galaxy at the faint end of the AGN luminosity function. *arXiv e-prints*, 2004:arXiv:2004.03611, 2020a.

Trebitsch M., et al.. The Obelisk simulation: galaxies contribute more than AGN to HI reionization of protoclusters. *arXiv e-prints*, 2002:arXiv:2002.04045, 2020b.

Wise J.H., et al.. The Birth of a Galaxy - III. Propelling reionisation with the faintest galaxies. *Monthly Notices of the Royal Astronomical Society*, 442(3):2560–2579, 2014.

Towards Cosmic Dawn III : Dust, metals, and cooling in RAMSES-CUDATON

2

2.1 Beyond CoDa II and missing physics

In Chp. 1, we presented CoDa II, a fully coupled RHD cosmological simulation of Reionization. Fig. 2.1 shows both the average neutral fraction (left) and the average J21 (right) of the simulation as a function of redshift. Although CoDa II satisfies constraints on the average neutral fraction during Reionization, after the overlap of reionising regions (z≈5.7), the average neutral fraction is too low, and there are too many ionising photons in the CoDa II box. Indeed, CoDa II's average neutral fraction is far below one of the most commonly used constraints used in Reionization simulations after overlap (Lyα forest derived average neutral fractions from Fan et al. [2006]). In Chp. 3, we shall investigate how a different calibration could help with these issues. And in particular, how tweaking the sub grid recipe for star formation might improve our agreement with observations. In this chapter, we focus on additions to the simulation code, adding in missing physics that could contribute to more realistic Reionization.

Evolving stellar populations One possible solution to this issue lies in the emissivities of stellar particles. In CoDa II, stellar particles were attributed a global value of ionising emissivity, that was constant over their lifetimes, and completely extinguished after 10 Myr. This is quite different from real stellar populations, whose ionising emissivities vary substantially based on their age and metallicity. Of particular interest when considering the above problematic excess of photons and ionisations, is that the ionising emissivities of both older and more metal rich stellar particles, that are more abundant towards the end of Reionization, are smaller. This could lead to substantial differences in the speed with which the IGM reionises, the post-overlap average neutral fraction, as well as in the galactic ionising photon budget, as it could alter the relative importance of galaxies depending on the ages and metallicities of their populations.

Cooling processes Star formation is a cornerstone of RHD simulations of Reionization. The SFR, as it is typically implemented in such simulations (like in CoDa II), is a power law of gas density. It therefore follows that it is of paramount importance that the simulated gas densities be realistic. As it turns out, maybe the most determinant factor in the density of an astrophysical plasma is the effectiveness of cooling mechanisms within it. Thus, it is crucial for RHD simulation projects to strive to best represent the important gas cooling processes. In CoDa I and II, two relevant sets of cooling processes were left out : Helium, and metal cooling. Helium is produced by BBN and is available straight after the Big Bang, and metallic species are present in star forming gas within the stellar generations that follow the first stars. In spite of their omission from the CoDa project so far, Helium

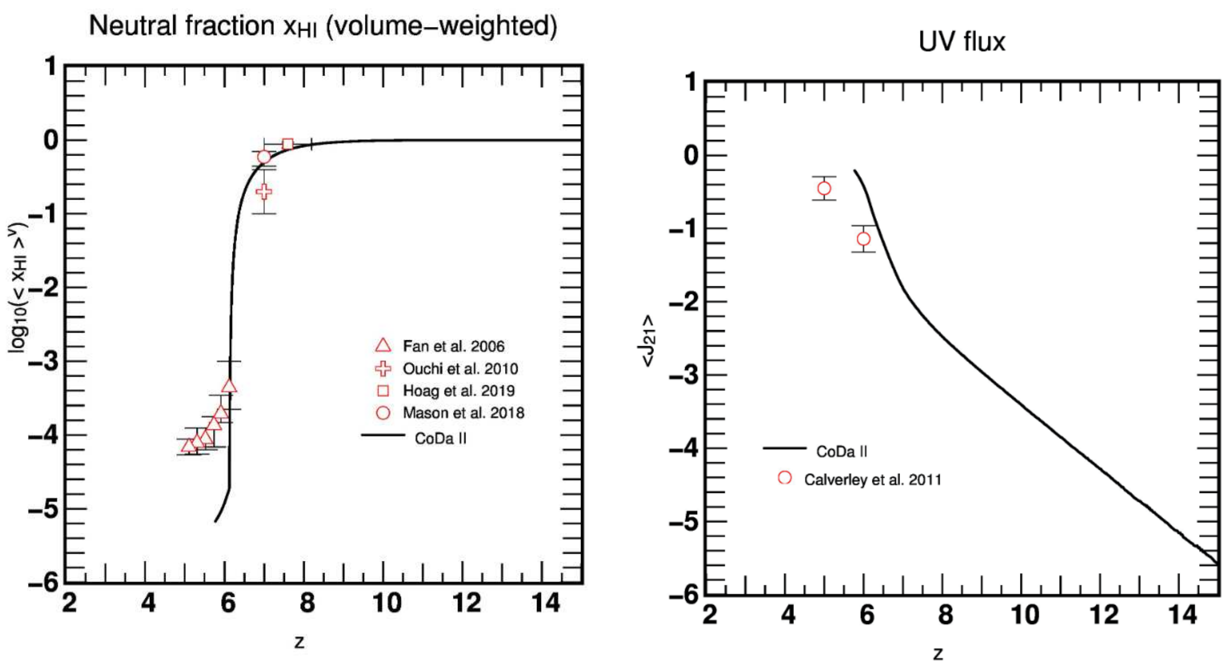


Figure 2.1: *Left*: Average neutral fraction in CoDa II versus redshift with constraints from Fan et al. [2006]; Ouchi et al. [2010]; Mason et al. [2018]; Hoag et al. [2019]. *Right*: Average J_{21} in CoDa II as a function of redshift, with constraints from Calverley et al. [2011].

cooling can dominate the cooling in hot gas (10^5 K, see Fig. 3 and Barkana and Loeb [2001]), whereas metal cooling is thought to allow star formation in mini-haloes[1] that are credible contributors to Reionization [Wise et al., 2014], and allows very hot ionised metal rich gas to cool. Adding to the cooling in our simulations could have drastic effects on our depiction of the EoR. Indeed, since the cooling can affect both the gas density and temperature, it can alter the balance between ionisation, feedback, and star formation.

Dust It is becoming increasingly apparent that some of the high-redshift (z>5) massive star forming galaxies contain large masses of dust [for instance : Burgarella et al., 2020; Laporte et al., 2017; Béthermin et al., 2015]. If this is the case, then there are wide reaching implications. Firstly, for the interpretation of the observations of high redshift galaxies. For example, extinction due to dust could significantly alter the slope of the UVLF. Indeed, Wu et al. [2020] show a substantial change in the slope of the UVLF in their simulations when accounting for dust extinction. Secondly, in simulations, large column densities of dust could reduce the LyC luminosity of galaxies, reducing the contribution of the most dust rich and possibly brightest galaxies to Reionization, potentially altering the geometry of Reionization, or the speed with which it progresses. At the same time, this could also lead to a reduction of external radiative feedback in and around the most massive dusty galaxies. Finally, dust grains can be heated by collisions with ions in hot gas, re-emitting the collision energy in the far-infrared and cooling the gas [Ostriker and Silk, 1973].

---

[1] with $<10^8\,M_\odot$ in which atomic cooling is ineffective, but metal cooling could still allow for star formation.

Throughout this chapter, we describe the implementation of the features related above in RAMSES-CUDATON, their effects on simulations, and the calibration of our implementation of dust.

2.2 Standard RAMSES chemical enrichment

Although they were deactivated in CoDa II and CoDa I, RAMSES-CUDATON possesses the chemical enrichment tracking features of RAMSES. In this body of work, we used the same parameters for chemical enrichment in RAMSES as in the SPHINX simulations Rosdahl et al. [2018], which are :

- η_{SN}=0.2, the fraction of stellar particle mass ejected as gas of supernovas
- yield=0.075, the fraction of metals in the ejected gas of supernovas

This parametrisation is compatible with stellar population models such as the BPASS V2.2.1 [Eldridge and Stanway, 2020] binary model, with a Kroupa initial mass function (IMF)d that we use systematically herein.

2.3 Evolving stellar emissivities

In CoDa II, the stellar particle emissivity was determined using a BPASS binary stellar population model Eldridge et al. [2017], assuming metallicities of Z=10^{-3} for all particles and no evolution with age. These emissivities corresponded to the ionising emission that should occur over the 10 Myr life time of the CoDa II stellar particles. Since we now track metallicity in gas and stars (in order to compute dust densities[2] and metal cooling[3]), we can allow the emissivities of particles to vary with it according to the latest version of BPASS (V2.2.1 binary model, with a Kroupa IMF, and stellar masses $\leq 100\,\mathrm{M}_{\odot}$). At the same time, we account for the change in stellar particles' emissivities as they age. To accomplish this, I implemented new routines in the radiation step of RAMSES, that read in and interpolate BPASS tables of emissivity as a function of age and metallicity. Instead of assigning a one value fits all emissivity to each stellar particle, during the construction of the source field, each particle receives an emissivity based on the interpolation of a table chosen in the RAMSES namelist configuration file.

Fig. 2.2 showcases the ionising emissivity of a stellar population as a function of its absolute metallicity and of its age, as we implemented it into our code. The emissivities decrease both with metallicity and with age, however they are much more sensitive to the age of the stellar population : for all metallicities, ageing from 10 Myr to 25 Myr reduces the population's emissivity by at least a factor of 100, and up to just under factor of 10^3. Changes in metallicity can only cause variations of up to a factor of 10 in the emissivity. Populations with the highest emissivities are those with ages of just under 10 Myr, and low metallicities (Z<10^{-4}). One apparent feature is that the older the population, the more its emissivity decreases with metallicity. Over the first 10 Myr, the emissivities vary little except for the particles with the highest metallicities (Z> $10^{-2.5}$). The main differences with the CoDa

---

[2] Sec. Dust model

[3] Sec. Implementing metal line cooling

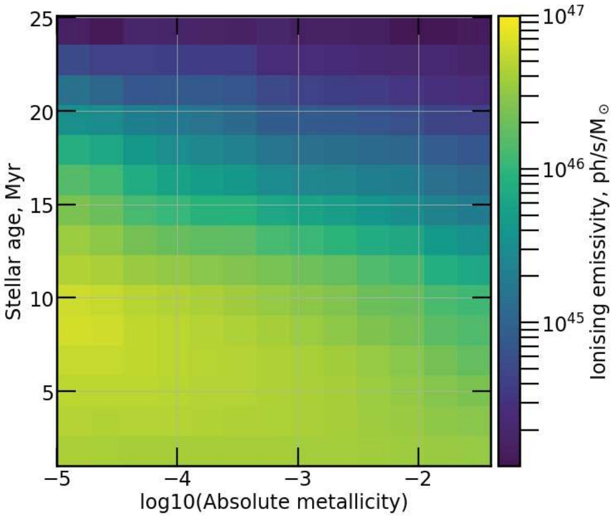


Figure 2.2: Map of the stellar particle emissivity as a function of absolute particle metallicity and particle age.

I and CoDa II models are therefore the extended lifetime and continued emission beyond 10 Myr. In particular, since the CoDa emissivities were calibrated to emit a lifetime of ionising photons within 10 Myr, our new model has on average lower emissivities. For instance : in CoDa II, the ionising stellar emissivities were fixed at $\approx 4.32 \times 10^{46}\,\mathrm{ph \cdot s^{-1} \cdot M_{\odot}^{-1}}$, whereas in our new model only young metal poor populations have stellar emissivities $\geq 10^{46}\,\mathrm{ph \cdot s^{-1} \cdot M_{\odot}^{-1}}$. Further, BPASS V 2.2.1 is slightly less emissive than BPASS V 2.1 as used in CoDa II.

Fig. 2.3 shows the average xHI as a function of redshift in a test simulation with a L4 N256 box[6]. Since these were test runs, no particular care has been taken in calibrating them so that Reionization is realistic. The left panel shows that introducing evolving stellar emissivities with particle metallicity and age delays Reionization somewhat. For a given redshift, this renders the post overlap <xHI> with the new emissivities somewhat higher. In the right panel, we shift the curves so that they meet at $\langle xHI \rangle = 10^{-4}$ · This shows that the changes in emissivities has not had a significant impact on the trend of the average neutral fraction, but merely causes a shift in the curves (a delay in Reionization).

Fig. 2.4 shows the mass weighted average of stellar particle emissivities in the CoDa II case (thick black line), and in the case of our new model that evolves with age and metallicity (blue full curve). The CoDa II emissivities are greater than the new model curves between z=14 and z=4 at all times, though the relative difference is small. This explains the difference in Reionization times seen in Fig. 2.3. The dotted constant age curve remains more or less constant over the presented time period, at around $4 \times 10^{46}\,\mathrm{ph \cdot s^{-1} \cdot M_{\odot}^{-1}}$, remaining between twice and over 2 dex higher than the other curves, illustrating the dramatic impact of stellar age. The dotted curve does in fact evolve with time, and just barely decreases due to enrichment of the stellar population. The curves that take into account the ageing of the particles (full, dot-dashed, and dashed lines) are already lower than the dotted curve at z=14 : all three are approximately $10^{46}\,\mathrm{ph \cdot s^{-1} \cdot M_{\odot}^{-1}}$ at z=14. Between z=14 and z=4, they decrease similarly till around $2.5 \times 10^{44}\,\mathrm{ph \cdot s^{-1} \cdot M_{\odot}^{-1}}$, again illustrating the strong impact of the age

---

[6] Throughout the manuscript this notation refers to a simulated box of 256^3 resolution elements, and a volume of

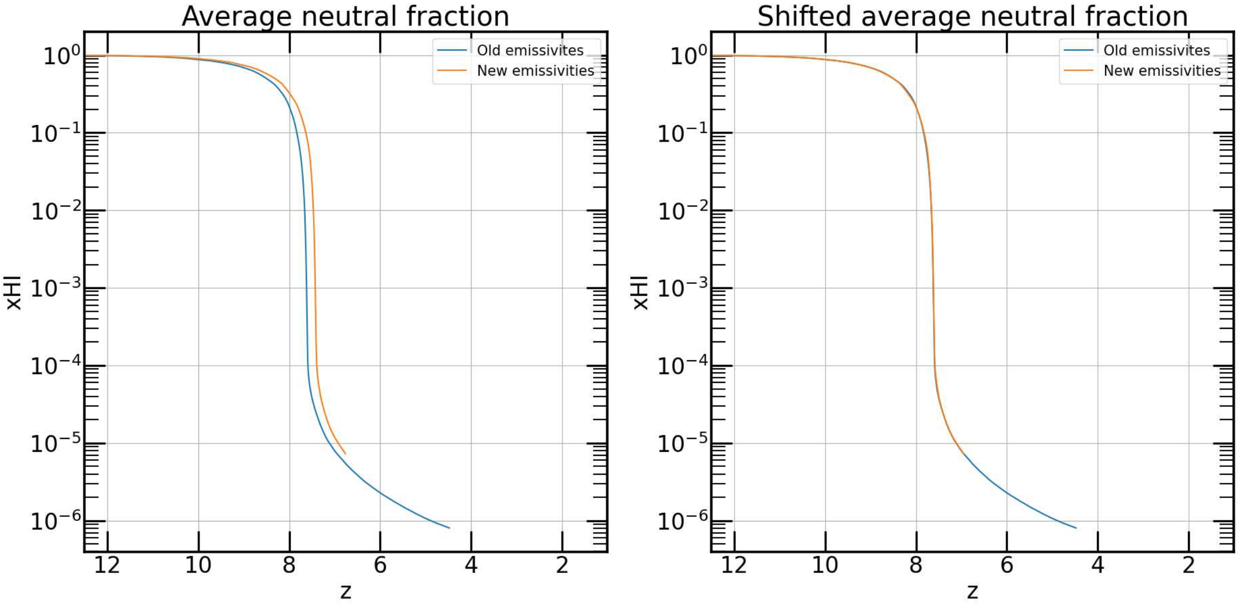


Figure 2.3: *Left* : Volume average neutral fraction for a CoDa II-like simulation with 256 cells per side with a corresponding side length of $4\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$[5] (blue) , and an equivalent resolution simulation with our new evolving stellar emissivities (orange). *Right* : The same curves, but the orange curve has been shifted in redshift so that the curves intersect where <xHI>=10^{-4} .

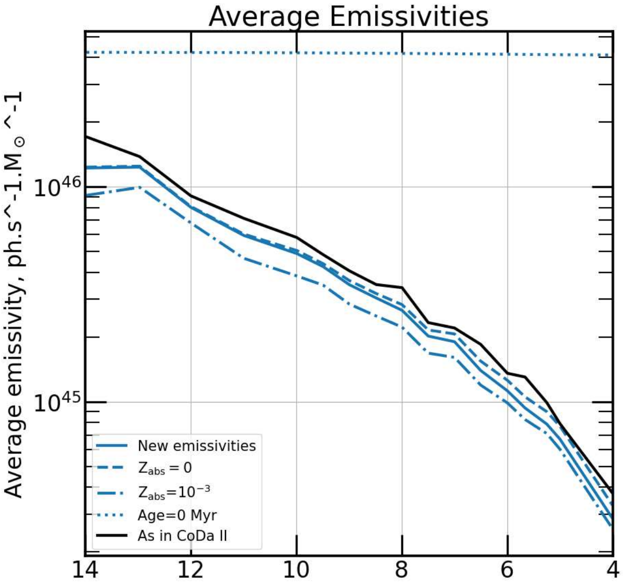


Figure 2.4: Mass weighted averages of the stellar particle emissivity of a L4 N256 simulation with the new emissivity model in blue lines. The blue full line shows the average when considering the actual particle ages and metallicities. The dotted curve is the case where we consider all stellar particles to have an age=0 Myr. For the dashed line we considered that all stellar particles were pristine (absolute metallicity Z=0). For the dot-dashed line we considered the particle absolute metallicities to be 10^{-3}.

of particles on their emissivities in the stellar population models. The full curve is greater than the dot-dashed curve at all times, and both are smaller than the dashed line. However as time progresses, and the stellar particles become more metal rich on average, the full curve pulls away from the dashed pristine (Z=0) curve and towards the metal enriched dot-dashed curve, almost intersecting it at z=4. Thus, showing the (slighter) evolution of stellar emissivity with metallicity.

2.4 Metal and Helium line cooling

In this section, I review my implementation of metal and Helium cooling in RAMSES-CUDATON, as well as some of its effects.

2.4.1 Implementing metal line cooling

RAMSES and RAMSES-RT [Rosdahl et al., 2013] already compute metal cooling rates on the fly using RAMSES' implementation of metals. Since tracking the various abundances and immense number of ionisation states is prohibitively costly, the approach they take is to interpolate a table of cooling rates based on cell temperature, metallicity, and redshift. The table in question was produced in CLOUDY [Ferland et al., 2017]. Below temperatures of 10^4K, they use cooling rates taken from Rosen and Bregman [1995]. However, since RAMSES-CUDATON relies on ATON for the gas thermochemistry, we cannot use the RAMSES metal cooling, as it is would introduce a second, different cooling step, as well as count Hydrogen cooling twice. I took the simple approach of porting[7] a version of the RAMSES metal cooling module to ATON in RAMSES-CUDATON. I therefore modified the interface between RAMSES and ATON to pass the RAMSES calculated metallicities to ATON during each RAMSES time step. Following the above method, interpolation based on temperature is used to find the total metal cooling rate (that does not include Helium nor Hydrogen contributions) for each cooling step. In my implementation, there is no redshift dependence (ie : an indirect dependence on a global UV background). The computed metal cooling rates are then added to ATON's Hydrogen cooling rate. Care is taken to update ATON's sub-cycling time steps to reflect the possible increased cooling.

2.4.2 Implementing Helium line cooling

Between 10^4K and 5×10^5K, the cooling rate of a primordial abundance gas is dominated by the collisional excitation of HI and HeII [Katz et al., 1996]. However, in RAMSES-CUDATON, only the first process was accounted for in the gas thermochemistry of ATON. To remedy this, and since tracking the abundances and different ionisation levels of Helium is beyond the scope of our aims, I implemented the following very simple approach into ATON. The main assumption is to consider that Helium is as ionised as Hydrogen (i.e. for every cell) as in Eq. 2.1 :

$$\mathrm{x_{He}} = \frac{\mathrm{n_{HeII} + n_{HeIII}}}{\mathrm{n_{HeI} + n_{HeII} + n_{HeIII}}} = \mathrm{x_H} = \frac{\mathrm{n_{HII}}}{\mathrm{n_{HI} + n_{HII}}}\,, \tag{2.1}$$

---

[7] I use the same CLOUDY table as in RAMSES and RAMSES-RT, described in Rosdahl et al. [2013, 2018] for instance

where $\mathrm{n_{HeII}}$ is the number density of Helium that has been ionised once, $\mathrm{n_{HeIII}}$ is the number density of Helium that has been ionised twice, $\mathrm{n_{HII}}$ is the number density of ionised Hydrogen, and $\mathrm{n_{HI}}$ is the number density of neutral Hydrogen.

If we then assume colisionnal ionisation equilibrium[8], we can then compute $\mathrm{x_{HeIII}}$ using the rates of HeII collisional recombination and collisional ionisation, and the rate of HeIII collisional recombination (Taken from Hui and Gnedin [1997]; Cen [1992]) as a function of cell temperature giving Eq. 2.2:

$$\mathrm{x_{HeIII}} = \frac{\beta_{\mathrm{HeII}}(\mathrm{T})\mathrm{x_{He}}}{\alpha_{\mathrm{HeIII}}(\mathrm{T}) + \beta_{\mathrm{HeII}}(\mathrm{T})}\,, \tag{2.2}$$

where $\mathrm{n_{HeIII}}$ is the fraction of HeIII ions, $\beta_{\mathrm{HeII}}(\mathrm{T})$ is the ionisation rate of $\mathrm{n_{HeII}}$, $\alpha_{\mathrm{HeIII}}(\mathrm{T})$ is the HeIII recombination rate, and T is the gas temperature in Kelvin.

From $\mathrm{x_{HeIII}}$ and $\mathrm{x_{He}}$ we then obtain $\mathrm{x_{HeII}}$, and $\mathrm{n_e}$ the density of electrons. Fig. 2.5 shows the relation between xHe, xHeII, and xHeIII for increasing temperatures. As temperature increases past 10^4 K, Hydrogen and Helium start to ionise via collisions. At around 4×10^4 K, the probability of collision is high enough to maintain all Helium atoms ionised once. Past this point, there start to be sufficient collisions to ionise HeII to HeIII. At temperatures above 3×10^5 K, the vast majority (>99%) of Helium atoms are maintained at double ionisation. Finally, assuming a BBN abundance of Helium ($\mathrm{Y} \approx 0.24$), we can then compute the cooling rates from excitation, ionisation, recombination, and Bremsstrahlung as function of cell temperature, density, and ionisation taken from Cen [1992]. When adding these cooling processes to ATON's cooling, care was taken to update ATON's sub-cycling time steps to reflect the possibly increased cooling.

2.4.3 New cooling rates

Fig. 2.6 breaks down the total cooling rate of a cell of gas of density $\mathrm{n_H} = 0.1\mathrm{H/cm^3}$ and metallicity $\mathrm{Z}=0.05\,\mathrm{Z_\odot}$ at collisional ionisation equilibrium. The left panel shows the contribution of the Hydrogen cooling processes and metal cooling (orange full line) to the total (blue full line) cooling and total Hydrogen (ATON, green full line) cooling. The total cooling rate increases very rapidly with temperature (and increasing ionisation) around 10^4 K, fast reaching a local maximum just under 2×10^4 K. At these temperatures, cooling is dominated by the collisional excitation of Hydrogen, whose contribution is more than one dex greater than the other cooling channels. As temperature increases past 2×10^4 K, the cooling rates associated with Hydrogen collisional ionisation and excitation decrease sharply as the ionisation fraction increases. At the same time, Compton cooling and Bremsstrahlung cooling rates increase steadily all the way past 10^7 K, where Bremsstrahlung is the dominant cooling mechanism. The term from Hydrogen recombination cooling increases very gradually until 2×10^5 K, after which it drops faster and faster with temperature. Finally, the metal cooling rate increases rapidly between 2×10^4 K and 10^5 K. Since the non-metallic cooling rates decrease after this point, metal

---

[8] Sutherland and Dopita [1993] show that assuming CIE can be a substantial approximation when compared with non equilibrium conditions (NEQ). At temperatures lower than 10^6 K, they show that some species do not reach collisional ionisation. In fact, the CIE cooling rates act as an upper limit to the NEQ case, which for some configurations of metallicity and temperature are up to two times lower than the CIE case. Rosdahl et al. [2013] show that in low density, hot plasma, equilibrium is not always reached, and that the difference between CIE and NEQ is highly dependant on the initial conditions of cells. This means that working with CIE probably leads to over-cooling, though the differences in our application are probably limited if it is mostly and issue in low density plasmas that do not form stars.

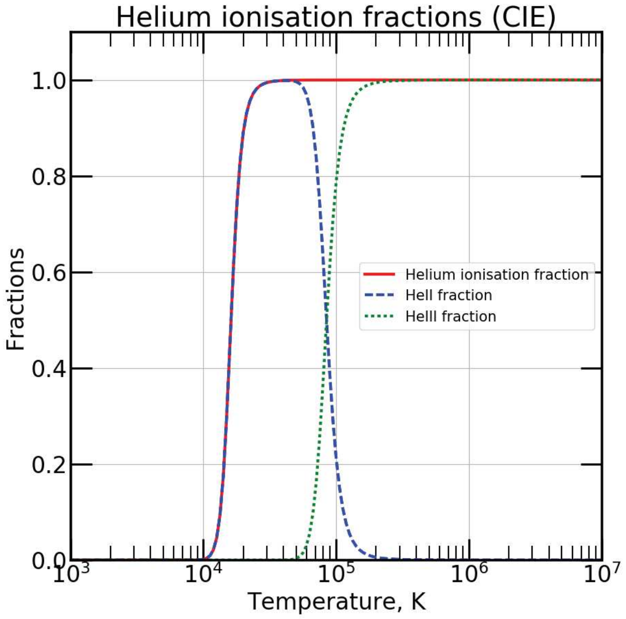


Figure 2.5: Assuming CIE across a range of temperatures and that xHe=xH, we compute the fractions of HeII and HeIII.

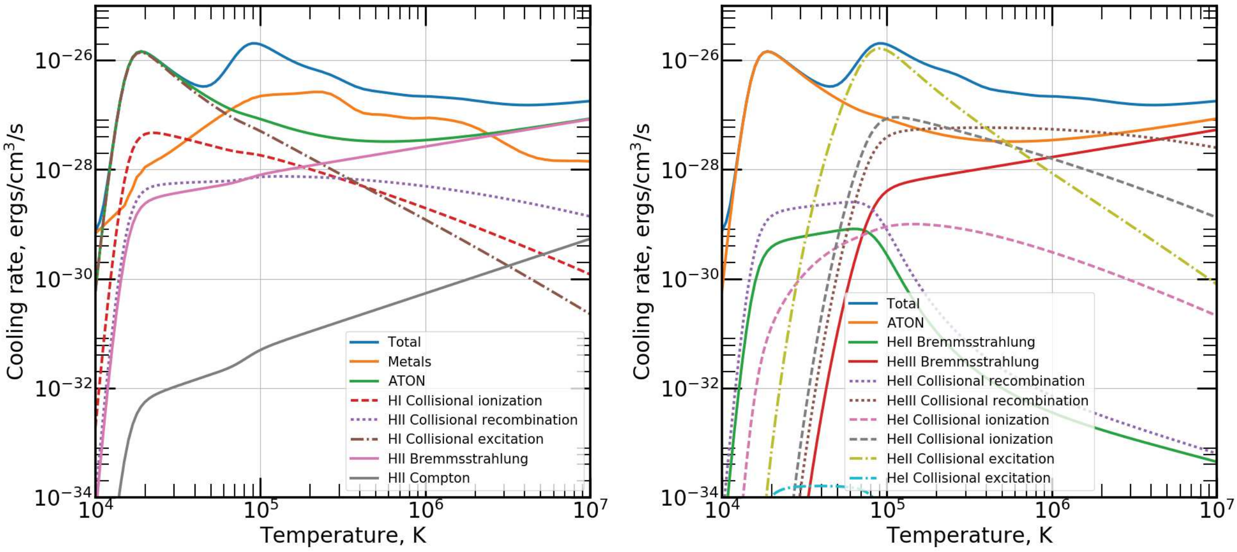


Figure 2.6: Cooling rates used in RAMSES-CUDATON as a function of temperature, assuming collisional ionisation equilibrium and a constant density of $n_H = 0.1 H/cm^3$, and varying metallicities. Full lines show the total cooling rate including Helium and metal cooling, whereas dashed lines only show the cooling rates dues to metal cooling.

cooling is the prominent cooling channel between 2×10^5 to 2×10^6 K, after which it is overtaken by Bremsstrahlung.

The right panel again shows the total (blue), and Hydrogen (ATON, orange) cooling rates, and allows us to compare them with the dominant Helium related cooling processes. All in all, the Helium cooling rates behave similarly to the Hydrogen ones, even rising to a maximum of similar height, albeit at 8×10^4 K instead of 2×10^4 K. At the lower end of the temperature range, HeII collisional recombination dominates the other Helium cooling rates until around 3×10^4 K, where HeII collisional excitation surpasses it and HeI collisional ionisation. The former process goes on to dominate all cooling channels at 9×10^4 K. Almost parallel to the HeII collisional excitation, HeII collisional recombination and HeII collisional ionisation also reach peaks at the same temperature and decline from thereonwards. At these temperatures, the fraction of HeII declines rapidly in favour of HeIII, causing a sharp drop off in the HeII Bremsstrahlung and HeII collisional recombination. The HeIII collisional recombination cooling rate remains almost constant until 10^6 K where it is the main Helium cooling rate, until it is overtaken by HeIII Bremsstrahlung. The cooling from HeI collisional excitation is several orders of magnitude smaller than the dominant Helium or Hydrogen cooling at any temperature.

Fig. 2.7 shows the new total cooling rate, as well as the metal cooling rate as a function of temperature, for various metallicities. The metal cooling rate increases linearly with metallicity, and reaches its maximum at temperatures between 7×10^4 to 3×10^5 K. Below 10^4 K, metal cooling rates are computed as in Rosen and Bregman [1995], and are the only source of cooling as we do not account for molecular H_2 cooling. However, this low temperature metal cooling is significantly lower than the peak from Hydrogen collisional ionisation, near 10^4 K, and only decreases towards lower temperatures. Below $Z=10^{-2}\,Z_\odot$, metal cooling is negligible when compared to the other implemented cooling channels. For metallicities greater than $10^{-1}\,Z_\odot$, the cooling rates from metals start to dominate the cooling from 2×10^5 to 10^6 K. Increasing the metallicity to $1\,Z_\odot$ yields cooling that is dominated by metals for all temperatures except in a narrow range around 2×10^4 K. Our metal cooling rates are very similar to those of Smith et al. [2008] (Fig. 1), who compute a grid of cooling coefficients using CLOUDY.

All in all, our new treatment of Helium cooling and metal cooling seems to have the potential to be impact the state of gas in simulations. Although potentially minor differences, we have shown in Lewis et al. [2020] that the state of the gas in star forming regions sets the f_{esc} of galaxies, therefore possibly affecting the ionising galactic photon budget. In particular, Helium cooling dominates cooling at ionisation equilibrium at temperatures of a few 10^5 K. The increased cooling rates in such gas could potentially boost star formation and feedback. Metal cooling is much more potent between temperatures of 2×10^5 to 2×10^6 K, though only in the most metal rich gas. In CoDa II, temperatures $>10^5$ K were typical of shocked gas (either by gravity or by supernova) around massive galaxies, that was transparent to LyC photons. Adding effective cooling channels at those temperatures could mean that these regions become denser and more opaque.

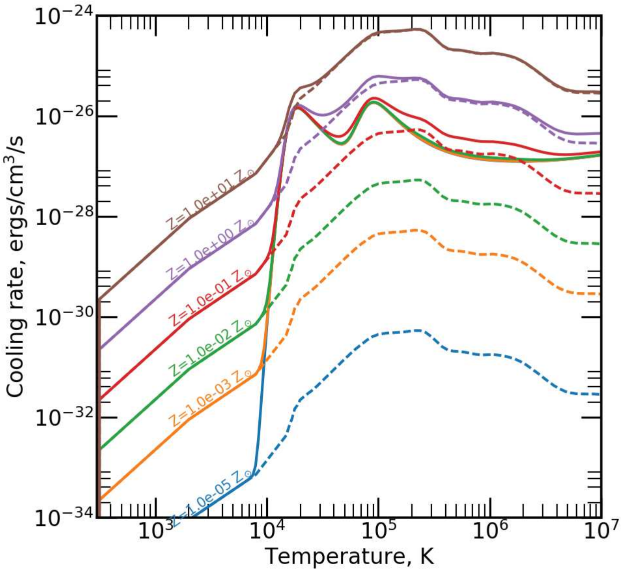


Figure 2.7: Cooling rates used in RAMSES-CUDATON as a function of temperature, assuming collisional ionisation equilibrium and a constant density of $n_H = 0.1 H/cm^3$, for various solar metallicities

2.4.4 Effects of He and Metal cooling on the temperature-density relation

Fig. 2.8 shows the temperature-density relation of simulated gas in several lower resolution (L16 N256) boxes where we have toggled our new cooling rates off and on. Starting with the top left panel, which shows the standard case of RAMSES-CUDATON with no added cooling, we see that most of the gas can be found scattered around a power law in temperature and density. Spanning roughly from $(T, \rho)=(4\times10^3\,K,\ 10^{-5}\,H\cdot cm^{-3})$ to $(T, \rho)=(7\times10^4\,K,\ 1\times10^{-3}\,H\cdot cm^{-3})$, and corresponding to $\rho \propto T^{1.61}$, this is close to RAMSES' gas equation of state (a mono-atomic perfect gas with an adiabatic index of 5/3). Deviations from the programmed equation of state are due to heating and cooling on time scales too small for the gas to react adiabatically. For instance, the very high temperature locus ($>10^5\,K$) in temperature-density space is due to shocks driven by infall or supernova explosions (as seen in Ocvirk et al. [2016] with the same code). The scatter around the EoS increases slightly with decreasing gas density.

The top right panel shows the same simulation set-up but with the additional Helium cooling that we have discussed. The most dramatic effect is upon the gas that is hotter than the EoS power law: the high temperature region above $10^5\,K$ has been emptied. The difference is particularly striking for the high density ($>10^{-2}\,H\cdot cm^{-3}$) cells, that appear to have cooled to under $10^5\,K$, where they form a ridge like number density structure. It is precisely around $10^5\,K$ that the collisional excitation of Helium is an effective coolant. This emptying affects all densities. Indeed, in lower density cells ($<10^{-3}\,H\cdot cm^{-3}$), the effect of He cooling is more moderate, however it has still had an appreciable effect as the high temperature low density region is less populated than without He cooling. In this simulation, the maximum encountered gas densities are slightly higher than in the "Vanilla" one (which was ran using default RAMSES-CUDATON cooling). By extending the temperature range at which cooling is effective, and at the same time increasing its effectiveness, we allow higher temperature

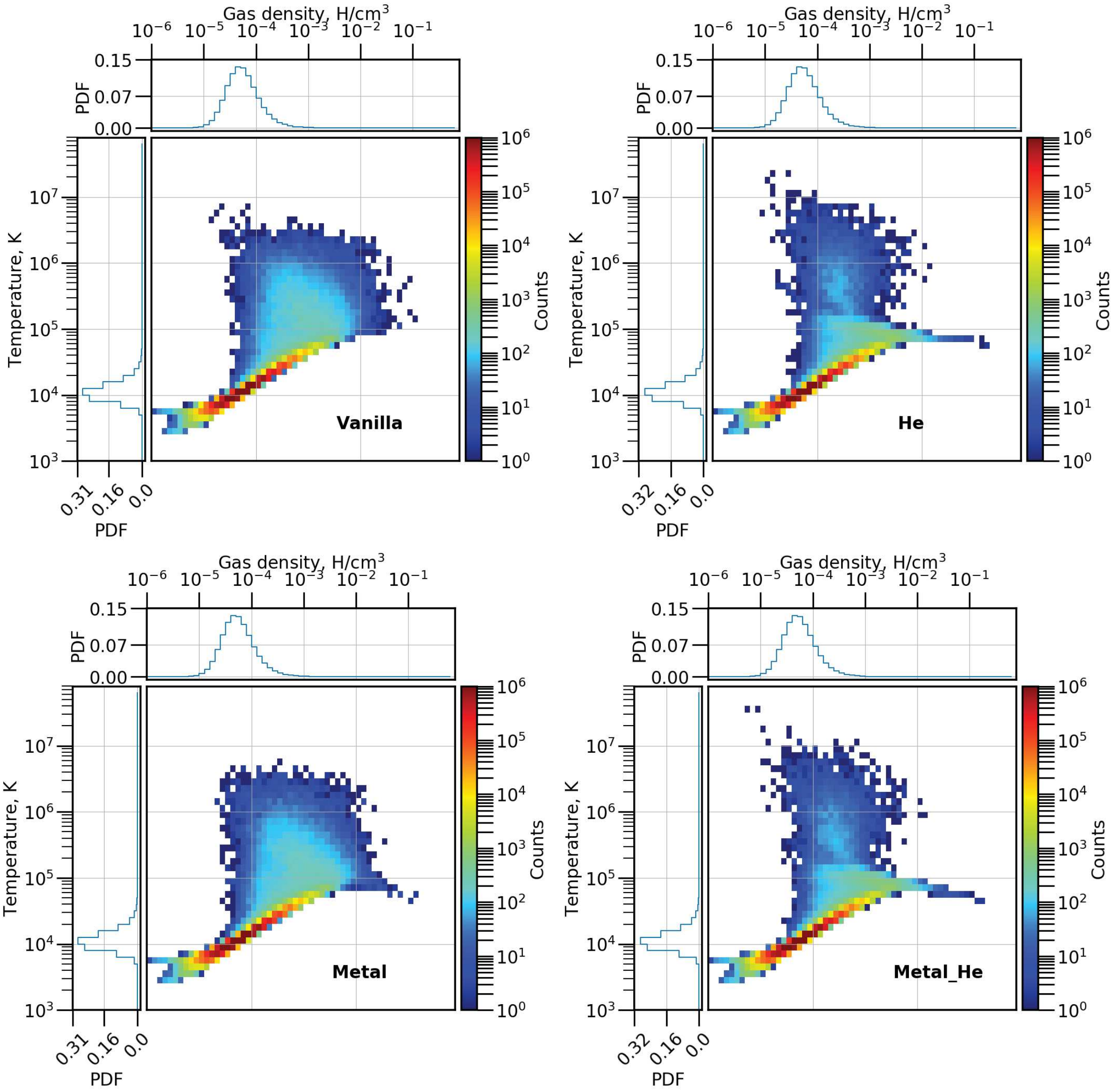


Figure 2.8: Temperature-density phase space of gas at z=7.5 in several star forming boxes ($16\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$), with different cooling channels active. In clockwise order starting from the top left : no extra cooling, He cooling, metal cooling, He and metal cooling

gas to cool more rapidly, thereby ensuring higher gas densities. Interestingly, the highest temperature ($>3 \times 10^6$ K) region of low density cells is more populated than in the Vanilla case. This could be the sign that our increased cooling in the higher density cells allows more star formation, in turn leading to more photons and supernova events, thereby producing hotter low density cells.

The bottom left panel presents the effect of our metal cooling on the gas temperature-density distribution. The effect of metal cooling on the state of our simulated case is much more moderate than that of Helium cooling. This can be understood simply, as although metal cooling can be effective for a wider range of temperature than Hydrogen or Helium cooling, the gas must be highly enriched (>0.05 $Z_\odot$) to be of comparable effectiveness. Further, metal cooling is most effective between 2×10^5 K and 10^6 K, and great majority of cells are cooler than this. However, there are differences with the Vanilla case :firstly, the hottest ($>3 \times 10^6$ K) and lowest density ($<10^{-4}\,\mathrm{H \cdot cm^{-3}}$) cells of the Vanilla simulation are cooler in the metal cooling simulation. Secondly, some of the densest cells ($>10^{-2}\,\mathrm{H \cdot cm^{-3}}$) of the Vanilla simulation appear to have cooled to below 10^5 K. In fact, in the metal cooling simulation (just like in the Helium cooling simulation), there is a high density tail of gas cells. However, in the He box, the tail is more densely populated and at the same temperature. Whereas in the metal box, the high gas density tail decreases somewhat in temperature with increasing density. It could be that the highest density regions have some of the highest metallicities, thereby rendering metal cooling more effective in higher density gas.

Finally, we examine the bottom right panel, that shows the case where both metal cooling and Helium cooling are active in a simulation. The resulting map of cell temperature and gas density is very similar to that of the He simulation. There is however one noteworthy difference : just as in the metal cooling simulation, the high gas density tail is cooler the higher the density.

2.4.5 Effects of new cooling on star formation

Fig. 2.9 shows the total star formation rate density in four small separate test boxes (L4 N256), namely : with ATON cooling (red), with ATON and metal cooling (black), with ATON and Helium cooling (silver), and with ATON, metal, and Helium cooling (pink). In all cases, the first stars appear in haloes near z=11. The SFRD curves increase with decreasing redshift until $z\approx 5.5$. At z=10, although the curves appear to fluctuate, a difference in the curves' trends as a function of their cooling appears to emerge causing the curves to split into two distinct groups from $z\approx 8$ and onwards : Both simulations with He cooling have higher SFRD than the two simulations without it. This is consistent with our analysis of Fig. 2.8, that showed the He cooling had the most dramatic effect on the highest densities and their temperature. By z=6, the He cooled curves' SFRDs are about twice as great than the non He cooled curves' SFRDs. Although the gap between the ATON only curve and the ATON+Metal cooling curve is very small ($\lesssim$ 0.1 dex), the latter is consistently higher than the ATON only curve from z=8 and onwards, showing the positive impact of metal cooling on star formation. The relative positions of the curves before z=8 are somewhat confusing. For instance, the ATON+Metal cooling box has the lowest SFRD between z=11 and z=8, when naively one would expect very similar values to the ATON only box until later times when the halo gas is on average more metal rich, and has reached temperatures for which metal cooling can have an impact. Similarly, until z=9, the ATON

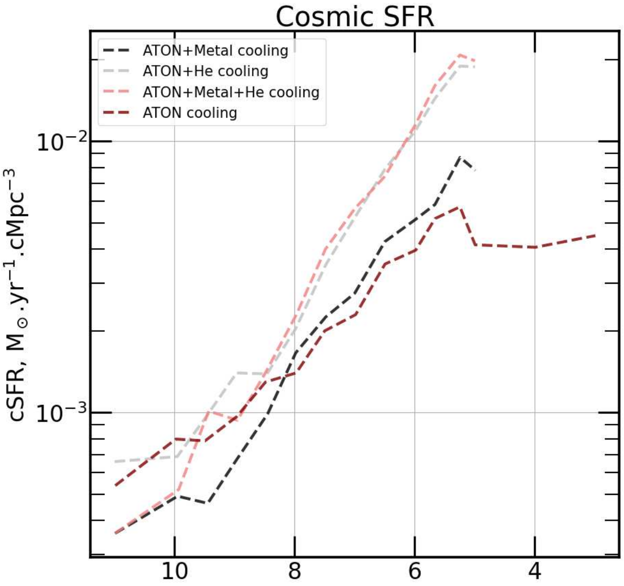


Figure 2.9: Cosmic star formation rates of several $4\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$ boxes, with different cooling channels active.

only and the ATON+He have very similar SFRDs. It could be that although the abundance of Helium does not change over the course of the simulation, at early times the star forming gas' temperatures are not high enough for Helium cooling to have a substantial effect on star formation.

2.5 Dust model

2.5.1 With a semi-analytical fit

Since RAMSES already tracks metals in the gaseous phase, one simple route to obtaining dust masses would be implementing a model of the dust to metal ratio (DTM). Dust formation occurs when the metal rich gas of supernova ejecta and the ISM condense into grains. Dust destruction occurs due to thermal and inertial sputtering, as well as by sublimation due to radiation. This means that a single galaxy can have a wide distribution of DTM values which is difficult to observe [Rémy-Ruyer et al., 2014; Wiseman et al., 2017], but also to model [Mancini et al., 2015; Popping et al., 2017; Vijayan et al., 2019], especially in the high redshift Universe.

I incorporated the fits of the semi-analytical (SAM) work of Vijayan et al. [2019] to compute the dust to metal ratio throughout our simulation. They use semi-analytical models of galaxies, that include realistic models for the formation, growth, and destruction of dust.

$$\mathrm{DTM} = \mathrm{D}_0 + (\mathrm{D}_1 - \mathrm{D}_0)\left(1 - \exp[-\alpha \mathrm{Z}^{\beta}((\mathrm{Age}\mathrm{D}_0\mathrm{Z})/\tau_{\mathrm{growth}})^{\gamma}]\right) , \tag{2.3}$$

with DTM the cell dust to metal ratio, D_1=0.339, D_0=8.75×10^{-4}, α=1.078 , Z the absolute metallicity, β=-1.083 , Age the mass weighted mean of stellar particles in the cell in Myr, τ_{growth}=5×10^4 Myr the growth scale for dust grain accretion, and γ=1.863.

The fitted function (Eq. 2.3) gives the DTM of a simulated cell as a function of gas metallicity, and of the average stellar populations age. We assume the conservative value of $\tau_{\mathrm{growth}} = 5 \times 10^4$ Myr given by the authors as it gives the best fit to observations.

This simple analytical relation between metallicity and stellar age allows both the post-processing of simulations, and computing a dust field on the fly. I first added fields to RAMSES that track the total stellar mass and mass weighted stellar age of each cell. The former is updated when stars form, and when they undergo supernovas, whereas the latter is computed at the same time as the source grid. Then, towards the end of every radiative step(before the source term is computed and sent to ATON on the GPUs), a grid of dust mass is computed using Eq. 2.3, the metallicity grid, the gas density grid, and the average stellar population age grid according to the following formula (Eq. 2.4):

$$\rho_{\mathrm{dust}} = \mathrm{Z}\rho_{\mathrm{gas}}\mathrm{DTM}\,, \tag{2.4}$$

2.5.2 With a physical model

Model

Yohan Dubois assisted me in implementing the physical model of Dubois et al. (in prep), that tracks dust creation and destruction on the fly. In practice, he adapted a version of his model to our version of RAMSES, that I implemented to function within the Vijayan et al. [2019] SAM, and then finished the merger with RAMSES-CUDATON (particularly for the feedback routines that are different from the RAMSES or RAMSES-RT codes, and for the coupling with radiation in ATON). The model assumes a standard solar chemical composition of dust[9] with a single average grain size of 0.1 µm, and includes the following mechanisms of dust production and destruction: Dust is released by supernova explosions (a fraction $\mathrm{f_{cond}}$, or the dust condensation fraction, of the released metal mass condenses into dust grains). It is also created by the accretion of gas phase metals, in the cooling step of RAMSES [taken from Eqs. 32 and 34 in Dwek, 1998] in Eq. 2.5 :

$$\dot{\mathrm{M}}_{\mathrm{growth}} = \left(1 - \frac{\mathrm{M_d}}{\mathrm{M_{metal}}}\right)\frac{\mathrm{M_d}}{\mathrm{t_{growth}}}\,\mathrm{M_\odot.Myr^{-1}}\,, \tag{2.5}$$

where $\mathrm{M_d}$ is the dust mass, $\mathrm{M_{metal}}$ is the total (gas and dust) metal mass, and $\mathrm{t_{growth}}$ is the growth time scale.

$$\mathrm{t_{growth}} = 100\alpha^{-1}(\mathrm{T})\mathrm{a_{0.1}}\mathrm{n_{gas}^{-1}}\left(\frac{\mathrm{T}}{20\mathrm{K}}\right)^{-0.5}\mathrm{Myr}\,, \tag{2.6}$$

with $\mathrm{a_{0.1}}$ the dust grain size (a representative value of 0.1 µm is chosen), $\mathrm{n_{gas}}$ the gas density in $\mathrm{g \cdot cm^{-3}}$, and T the temperature in Kelvin. $\alpha(\mathrm{T})$ is the dimensionless sticking coefficient of gas particles onto dust [Chaabouni et al., 2012]. In our application, it is 1 up to T=10^5 K, following $\alpha(\mathrm{T}) = (1 + 10(\frac{\mathrm{T}}{10^6\mathrm{K}}))^{-1}$ as in Novak et al. [2012].

Dust is destroyed by the inertial shock waves of supernovas following Eq. 2.7 :

$$\Delta\mathrm{M_{dest,SN}} = 0.3\frac{\mathrm{M_{s,100}}}{\mathrm{M_g}}\mathrm{M_d}\,\mathrm{M_\odot}\,, \tag{2.7}$$

where $\mathrm{M_g}$ is the cell gas mass and $\mathrm{M_d}$ is the cell dust mass. $\mathrm{M_{s,100}} = 6800\mathrm{E_{SN,51}}\mathrm{M_\odot}$ is the mass

---

[9] in terms of the ratio of silicates, carbons, etc

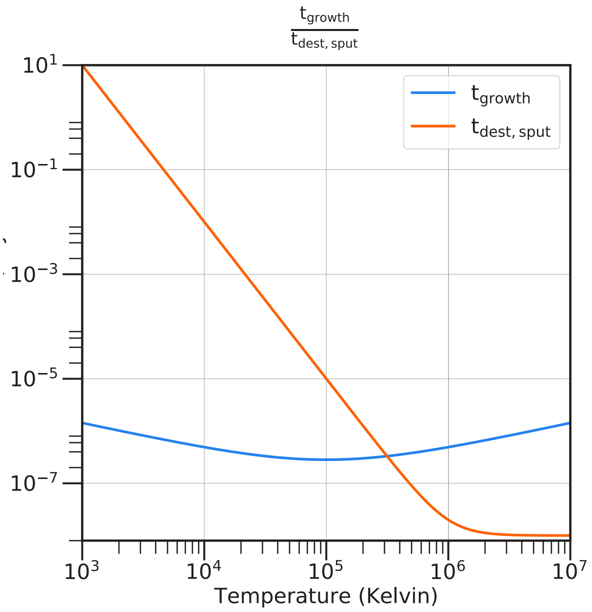


Figure 2.10: Timescales in Myr for both dust grain growth and dust destruction due to sputtering

of shocked gas at velocities larger than $100\,\mathrm{km}\cdot\mathrm{s}^{-1}$ (where inertial sputtering is sufficiently efficient). Finally, $E_{SN,51}$ is the SN explosion energy normalised by 10^{51} ergs. Thermal sputtering also destroys dust. One of the first accurate calculations of the rate of thermal sputtering was given by Draine and Salpeter [1979]. Here we use a fit to that result that originates from Novak [2013] :

$$t_{dest,sput} = 0.1 a_{0.1} n_{gas}^{-1} \Big(1 + (\frac{10^6 \mathrm{K}}{\mathrm{T}})^3 \Big)\,\mathrm{Myr}\,, \tag{2.8}$$

Fig. 2.10 shows both t_{growth} and $t_{dest,sput}$ as a function of temperature at fixed gas density. At 10^3 K the timescale for dust destruction via thermal sputtering $t_{dest,sput}$ is over e5 times longer than the time scale for dus t grain growth t_{growth}. However, whereas the former decreases with increasing temperature until 10^6 K , the latter changes only by a factor of ≈ 2. This means that dust destruction via thermal sputtering outpaces dust grain growth in cells hotter than 4×10^5 K. Based on the analysis of CoDa I and CoDa II [Ocvirk et al., 2016, 2020], such hot cells correspond to partly to gas cells heated by shocks, but mostly to the heating from supernova events. Hence, in most dense star forming cells that tend to remain cooler dust destruction should be negligible.

At the end of the cooling step, each cell's dust to metal ratio is checked against a maximum parameter : max(DTM). The dust density grid that is updated during the feedback and hydro cooling steps is passed to ATON along with the other hydrodynamical and source terms, where it is used in the radiative transfer computations.

Model parameter exploration

The main free parameters of the model are the dust condensation fraction f_{cond}, that sets the fraction of metals expelled by supernova events that condense into dust, and the maximum dust to metal ratio for every cell (max(DTM)). Fig. 2.11 illustrates the impact of these two parameters on the relationship

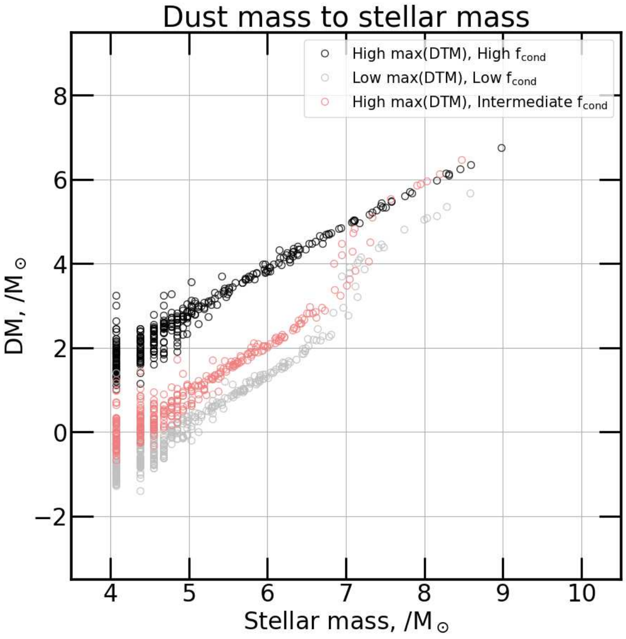


Figure 2.11: Log-log of galactic dust mass as a function of galactic stellar mass in three 256 cell,4 cMpc $\cdot$ h^{-1} boxes at z=6 . The black simulation has the highest f_{cond} (0.5) and the second highest max(DTM) (0.5). The grey simulation has the lowest f_{cond} (0.001) and the lowest max(DTM) (0.1). The pink simulation's parameters are : f_{cond} = 0.005, max(DTM) = 0.9.

between galactic dust mass and stellar mass in three L4 N256 simulations. In all three cases, dust mass rises with stellar mass. The grey locus (lowest max(DTM), lowest f_{cond}) is under the other two. The pink points (highest max(DTM), intermediate f_{cond}) are above the grey data for all stellar masses, and above the black and grey points for the highest stellar masses. Finally, the black data (second highest max(DTM), highest f_{cond}) are above all the others except at the highest stellar masses, where the pink galaxies. Summarily, the effects of the two free model parameters can be described by their effect on the loci of points below and above $M_\star \approx 10^7\,M_\odot$: the greater the f_{cond} of a simulation, the higher the dust masses below $M_\star \approx 10^7\,M_\odot$; whereas, the higher the max(DTM) ratio, the more dust can accumulate in the galaxies with the most stellar mass. This can be understood quite intuitively, since the highest stellar mass galaxies have the longest and strongest history of star formation, and therefore have the necessary age and SFR for cells to reach the maximum DTM ratio. Hence, their dust masses are far more sensitive to the maximum possible accumulation than to f_{cond} which sets the rate of production per stellar generation (not accounting for destruction). This is not the case of the lower stellar mass haloes, for which no or little cells have formed enough stars (and thus dust) to reach the maximum DTM ratio, rendering their dust masses insensitive to the max(DTM), and more sensitive to f_{cond}. Interestingly, the dust mass to stellar mass ratio changes abruptly in the pink data and in the grey data at roughly $M_\star \approx 10^7\,M_\odot$ in both cases. For these two data sets lower stellar mass galaxies appear less able to accumulate dust than high stellar mass galaxies. This is not the case of the black dataset, whose f_{cond} seems high enough that the dust masses of all galaxies follow the trend of high stellar mass galaxies where dust mass is limited by the maximum cell DTM ratio.

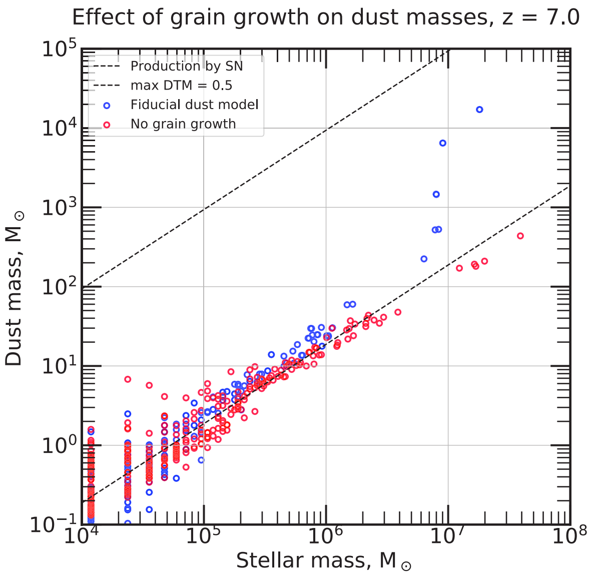


Figure 2.12: Halo dust mass as a function of their stellar mass at z = 7 in two L4 N256 simulations. The first (blue points) feature both dust production via condensation in the ejecta of supernovas and the growth of dust grains, whereas the second (red points) features only the former dust production process (grain growth was turned off manually in the simulation code). The lower dashed line shows the expected dust mass based only on condensation of dust in supernova ejecta. The upper dashed line shows the total possibly dust mass if all galaxy cells had the max(DTM) = 0.5.

As the dust masses of low stellar mass galaxies appears to be set by $f_{\rm cond}$ (i.e. : the dust masses in these galaxies are produced by condensation in the ejecta of supernovas), and the high dust masses in high stellar mass galaxies are limited by max(DTM), the observed sudden increase in the dust mass to stellar mass ratio could be due to dust grain growth only being effective in high stellar mass haloes. To explore this possibility I ran a new test simulation in which dust grain growth was deactivated. Fig. 2.12 shows the resulting dust masses versus stellar masses. As surmised, when dust grain growth is disabled, dust masses follow the prediction based on the ejected dust mass[10]. When it is enabled, we again observe a sudden uptake in dust galactic dust masses near $M_\star = 10^7\,M_\odot$. This shows the role of dust grain growth in dramatically increasing the dust mass to stellar mass ratio in high stellar mass galaxies.

We now investigate investigate the effects of other adjustable parameters; though they are not model parameters, the gas resolution, stellar particle mass, and yield simulation parameters can be expected to affect the dust masses of galaxies in our simulations. We explore this in Fig. 2.13. In the left panel, the grey points are at the same grid resolution as the black points, but have a lower yield parameter. Decreasing the yield parameter directly decreases the mass of metals and dust ejected by supernovas, thus causing the grey dust masses to be lower for all stellar masses. The pink data points have the same yield as the black ones but have a higher resolution. Increasing the resolution does not

---

[10] computed using the mass of stars that have already undergone supernova in the galaxies

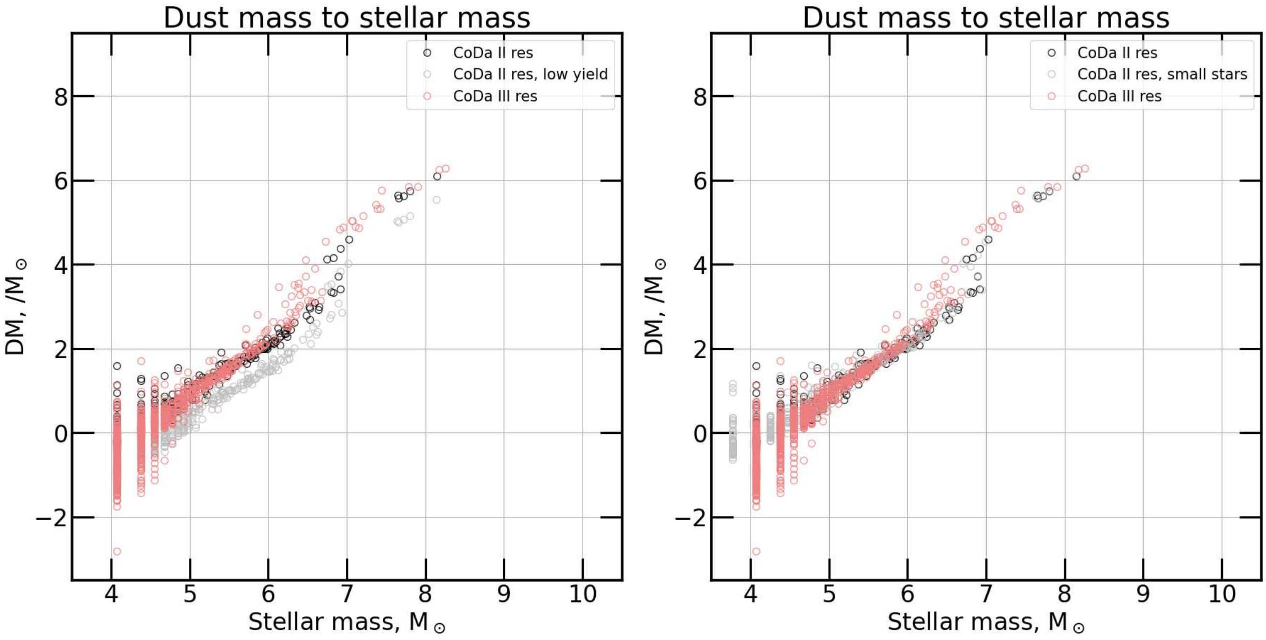


Figure 2.13: Log-log of galactic dust mass as a function of galactic stellar mass in three boxes at z=6. Points with the label "CoDa II res" are L4 N256 boxes. The points labelled"CoDa III res" are L4 N512. Finally, the points bearing the label "small stars" have stellar particles that are half as massive as in CoDa II and CoDa III.

significantly affect the dust masses of the highest stellar mass or lowest stellar mass galaxies. However, it increases the dust masses from stellar masses of $2 \times 10^6\,M_\odot$ to $10^7\,M_\odot$, increasing the number of galaxies whose dust masses are limited by max(DTM). It could be that increasing the grid's resolution tends to allow for the higher density cells to emerge, plausibly increasing the accumulation of dust in galaxies (up to the cell by cell limit of max(DTM)). for instance, we saw in Fig. 2.10 that the lower cell temperature is, the faster dust grain growth is relatively to thermal sputtering. In the higher gas density cells brought on by the higher resolution, gas cooling could be more effective giving rise to lower temperatures and more dust grain growth. Another important choice in our simulations is the mass of stellar particles, as more massive particles result in greater masses of ejected dust, but also greater radiative and supernova feedback and could affect gas density and temperature and therefore dust densities. The grey points in the right panel shows dust mass as a function of stellar mass in the case where stellar particles are half as massive as in CoDa II and in CoDa III. There is no significant difference with the CoDa-like stellar particle mass cases.

Fig. 2.14 shows the relation between halo dust mass and halo stellar mass (red points) in a L16 N1024 test box. It shows the tentative agreement between our physical model and the dust values obtained by post processing the simulation using Eq. 2.4 and Eq. 2.3 from Vijayan et al. [2019]. Both data sets show the strong correlation between dust mass and stellar mass, and agree fairly well. Below $_\star$ $=10^7\,M_\odot$, both distributions have large scatters of up to 4 dex in dust mass for the lowest stellar masses. The post-processed blue points show that dust masses broadly follow a power-law of stellar mass as they increase from around $2\,M_\odot$, at $M_\star$=$10^4\,M_\odot$, to about $10^5\,M_\odot$ at $M_\star$=$10^8\,M_\odot$.

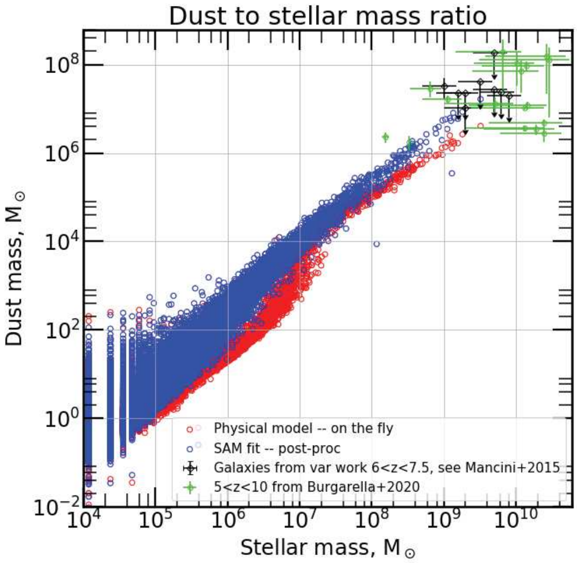


Figure 2.14: Dust to stellar mass ratio of galaxies in a L16 N1024 box at z=6. The blue circles are from post-processing the simulation using the fitting formulas from Vijayan et al. [2019], and the red circles are the values computed by our physical model developed in Dubois et al. (in prep).

The evolution of the red physical model points is different. Indeed, before $\approx 3 \times 10^6\,M_\odot$, the red dust masses appear scattered around a similar power law of stellar mass. However, between $3 \times 10^6\,M_\odot$ and $10^7\,M_\odot$, the red physical model points jump abruptly and come closer to the blue post-processed dust masses reaching $\approx 10^4\,M_\odot$. Beyond this mass, the two loci split again, with the physical model yielding lower dust masses than the post-processing SAM approach :for the highest stellar mass galaxies, the physical model predicts dust masses that are almost 10 times lower. While there is a disagreement between the two computations, this comparison is encouraging as it shows that two separate approaches yield the same global trend with stellar mass, and both predict believable dust masses in high stellar mass galaxies. We find that we are able to tune the dust masses of high stellar mass galaxies in our physical model using the max(DTM) parameter, allowing us to produce high stellar mass galaxies with dust masses that are more compatible with the constraints from Mancini et al. [2015]; Burgarella et al. [2020], permitting both implemented models to produce credible dust masses in high stellar mass galaxies at high redshifts.

Where does dust form?

So as to gain insight into the behaviour of the physical dust model of Dubois et al. (in prep), we study the physical properties of simulation cells according to their dust densities, painting a rough picture of where dust forms and why. Fig. 2.15 shows (from top to bottom, left to right) the distributions of cell gas density, absolute metallicity, temperature, SFR, stellar mass, oldest stellar particle of cells, cell magnitude in a L16 N1024 box for three ranges of dust density at z=6, that we shall call low (blue curves, 10^{-35} to $10^{-32}\ \mathrm{g \cdot cm^{-3}}$), medium (orange curves, 10^{-32} to $10^{-30}\ \mathrm{g \cdot cm^{-3}}$), and high (green curves, 10^{-30} to 10^{-27} $\mathrm{g \cdot cm^{-3}}$). The distributions are normalised by the total distribution of cells with $\mathrm{n_{dust}} > 10^{-35}\,\mathrm{g \cdot cm^{-3}}$.

The top left panel shows that gas density distributions are tightly linked to dust density distributions : below the density criterion for star formation, close to all cells are in the low dust bin. Above this threshold, the fraction of cells in the low dust bin rapidly drops off, as the fraction in the medium dust bin grows. By around $3 \times 10^{-26}\,\mathrm{g \cdot cm^{-3}}$, the medium dust bin fraction is at its maximum, and begins to be overtaken by the high dust bin. Beyond $10^{-25}\,\mathrm{g \cdot cm^{-3}}$, close to all cells belong to the high dust bin. This shows that dust either forms more effectively or is destroyed less effectively in cells with higher gas densities. As speculated when investigating model parameters, it may be that higher density cells form more stars and therefore have formed more dust through supernova and condensation, or that the high density cells cool more efficiently, thereby maintaining lower temperatures. Indeed, Fig. 2.10 showed us that decreasing gas temperature could affect favour the growth of grains over thermal sputtering .

Returning to Fig. 2.15: for all metallicities, >80% of cells are in the low dust bin. In fact, the fraction of cells in the low dust bin is close to one everywhere except between $\mathrm{Z}=3 \times 10^{-5}$ and $\mathrm{Z}=3 \times 10^{-3}$, where there is a small fraction of cells in the medium and high dust bins. Both distributions resemble one another, with a single maximum around $\mathrm{Z}=10^{-4}$, and a tail that extends to higher metallicities. However, the medium dust bin has a much greater fraction of cells (>5 times more). It is possible that this range in metallicity could evolve with time, gradually reaching higher dust densities. Its interesting to note that the highest metallicity cells are all in the low dust bin. It could be the sign that the highest metallicities reside in low density and high temperature gas, where dust is produced inefficiently and destruction processes are effective. This dust bin is quite probably associated with galactic winds. This means that although gas metallicity plays an important role in determining the dust density, it is not indicative of the dust density on its own.

At the lowest temperatures ($<4 \times 10^{3}\,\mathrm{K}$) cells belong mostly (>60%) to the high dust bin, and to the medium dust bin (>15%). The high dust bin fractions peak at around 85% near $3 \times 10^{3}\,\mathrm{K}$,whereas the maximum of the medium dust bin is nearer to $2 \times 10^{3}\,\mathrm{K}$, where the fractions peak at close to 40%. At temperatures higher than $4 \times 10^{3}\,\mathrm{K}$, almost all cells belong to the low dust bin. Indeed, dust destruction efficiency rises faster with temperature than dust accretion does. There is a very faint peak of medium dust bin cells near $4 \times 10^{4}\,\mathrm{K}$. At temperatures $>3 \times 10^{4}\,\mathrm{K}$ in this simulation, star formation is not permitted. This peak might be the sign that a nearby supernova event has occurred, raising both the dust density and the temperature. The faintness of the peak might indicate that the dust generated following these events tends to be destroyed in cells such as these, hence dust not accumulating. However, it is more likely simply the sign that these heated medium dust cells are very rare with respect to all cells with $\mathrm{n_{dust}} > 10^{-35}\,\mathrm{g \cdot cm^{-3}}$.

Cells with the lowest possible star formation rate (due to resolution), are overwhelmingly (>82%) in the low dust bin. However, the fraction of cells in this bin rapidly drops off, and by $2 \times 10^{-3}\,\mathrm{M_{\odot} \cdot yr^{-1}}$, the low dust bin has been overtaken by the medium bin that fluctuates between 30% and 40% from this point onwards. The fraction of cells in the high dust bin increases smoothly from the lowest SFRs till $1.0\,\mathrm{M_{\odot} \cdot yr^{-1}}$, reaching close to 80% of cells. At such high values of SFR, the low dust bin is comprised of less than 10% of the cells. Strikingly, although the greater the SFR the higher the fraction of cells with higher dust densities, there is a considerable amount of fluctuation : between 10^{-1} to 1 $M_{\odot} \cdot yr^{-1}$, the fraction of cells in the low dust bin varies between 0% to 10%. This indicates that there are cells with considerable star formation that retain low and medium dust densities. It could be that a 10 Myr average SFR is only a partial indicator of the dust content, and that the full star formation history must be considered to understand dust densities.

Almost all cells in which the oldest stellar particle is younger than 70 Myr belong to the low dust bin. Beyond this threshold in age, a considerable fraction of cells are in the medium dust bin. Between 300 Myr and 600 Myr, the fraction of cells in the medium and high dust bins increases sharply to around 35% and 50 % respectively. At the same time, the number of cells in the low dust bin decreases equally abruptly from close to 90% to just below 20 %. The age of the oldest stellar particle is linked to the cell's star formation history. The older the oldest particle, the longer the period of time during which there has potentially been star formation, and therefore dust generation, within cells.

In the final panel, we examine the fraction of cells belonging to each dust bin, depending on their magnitudes. The low dust bin resembles a high plateau of values near 90 % for the faintest cells (>-11). In brighter cells, the fraction belonging to the low dust bin rapidly drops, reaching values of just over 10 % by -15. Of the rare cells that are brighter than this, hardly any belong to the low dust bin. However, there is a large spike of over 20 % near -19. The medium dust bin behaves much as the low dust bin for the faintest cells (>-11), albeit much less populated as it contains merely 10 % of the faint cells. These faint but relatively dusty cells could be the sign of cells where star formation has been suppressed, and that maintain some dust content. The sudden decrease seen in the fraction of cells belonging to the low dust bin, is accompanied by a corresponding sharp increase in the fraction in the medium dust bin. At around -13, the fraction of cells in the medium dust bin reaches close to 50 %. Between -14 and -20, the fraction of cells in the medium dust bin falls to just over 20% at -20. Around -19, there is a massive peak that reaches 80 %. Almost no faint (>-11) cells belong to the high dust bin. In bright cells, the fraction of cells in the high dust bin rises steeply from a few percent at -11, to just over 60 % at -15. From there on its rise is steadier. In the brightest bin, all cells belong to the high dust bin. However, between -20 and -19, there is a drastic decrease in the fraction of cells in the high dust bin, reaching a minimum of 20 %. The large fluctuations at the low magnitude end are most probably noise due to the rarity of such bright cells.

On the whole, we find that dust accumulates the most in cells where the gas is dense, cool, and metal rich, and where there is a high SFR, and old stellar populations. This means that the highest dust masses are probably mostly found in the inner-most regions of galaxies, where the gas is the densest and the most stars form. This is in-line with simulations such as Aoyama et al. [2020], and observations such as Ménard et al. [2010] that observed dust surface densities decrease with distance from galactic centres. At the same time, the results from the physical model of Dubois et al. (in prep) show that the cells with the highest dust masses do not have the highest metallicities, showing that prescriptions based on metallicity alone fail to match results from physical models. Because of this, we will exclusively use the physical dust model from Dubois et al. (in prep) throughout the rest of the presented work.

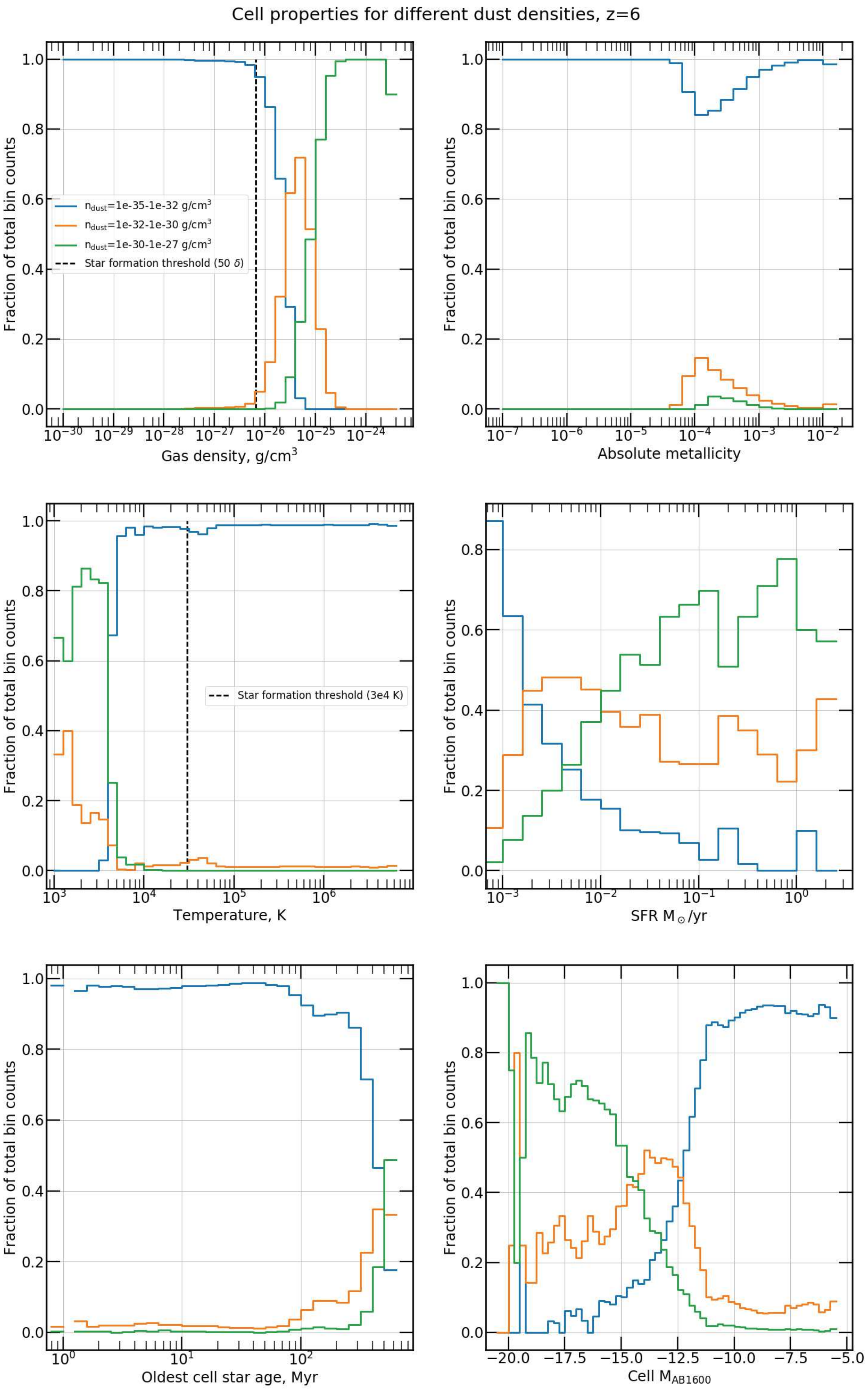


Figure 2.15: Distributions of cell properties (gas density, absolute metallicity, temperature, neutral fraction, stellar mass, oldest stellar particle) for three dust density bins. The distributions are given as a fraction of the total cells per bin.

2.6 Radiative transfer through dust

2.6.1 LyC photon absorption by dust in ATON

In order to implement dust absorption into ATON, we must re-derive the radiative transfer equations as done in Sec. 3.3 of Aubert and Teyssier [2008]. Starting from the set of transport equations (Eqs. 32 and 33 in Aubert and Teyssier [2008]), we add an absorption term that is analogous to the neutral Hydrogen one, and obtain Eq. 2.9.

$$\begin{cases} \frac{\mathrm{dN}_\gamma}{\mathrm{dt}} = \mathrm{n_e n_{HII}}(\alpha_\mathrm{A} - \alpha_\mathrm{B}) - \mathrm{n_{HI}}\sigma_\gamma \mathrm{cN}_\gamma - \rho_\mathrm{d}\kappa_\mathrm{dust}\mathrm{cN}_\gamma \, , \\ \\ \frac{\mathrm{d}\mathbf{F}_\gamma}{\mathrm{dt}} = -\mathrm{n_{HI}}\sigma_\gamma \mathrm{c}\mathbf{F}_\gamma - \rho_\mathrm{d}\kappa_\mathrm{dust}\mathrm{c}\mathbf{F}_\gamma \, , \end{cases} \tag{2.9}$$

here, $\mathrm{n_e}$ is the electron density, $\mathrm{n_{HII}}$ is the proton density, α_A is the case A recombination coefficient, α_B is the case B recombination coefficient, $\mathrm{n_{HI}}$ is the neutral Hydrogen density, $\sigma_\gamma = 2.93055 \cdot 10^{-22}\mathrm{m}^{-2}$ is the effective ionizing cross section for 20.28 eV photons[11] N_γ the ionizing photon density, $\mathbf{F}_\gamma$ the ionizing photon flux, and c the speed of light, ρ_d is the dust mass density in a cell in $\mathrm{g \cdot m^{-3}}$, and κ_dust =8.85 $\mathrm{m^2 \cdot g^{-1}}$ is the dust mass attenuation coefficient at $\lambda = 700$Å [12] derived from SMC[13] data by from Draine and Li [2001].

Then we can follow the same steps as Aubert and Teyssier [2008] (to reach an equivalent of their Eqs. 36 and 37), and obtain an implicit numerical scheme for updating the number density and flux of photons in a cell in Eq. 2.10.

$$\begin{cases} \mathrm{N}_\gamma^{\mathrm{p+1}} = \frac{\Delta_\mathrm{t}\left(\mathrm{n_e^{p+1} n_{HII}^{p+1}}[\alpha_\mathrm{A}^{\mathrm{p+1}} - \alpha_\mathrm{B}^{\mathrm{p+1}}]\right) + \mathrm{N}_\gamma^\mathrm{p}}{1 + \mathrm{c}\Delta_\mathrm{t}[\mathrm{n_{HI}^{p+1}}\sigma_\gamma + \kappa_\mathrm{dust}\rho_\mathrm{d}^{\mathrm{p+1}}]} \, , \\ \\ \mathbf{F}_\gamma^{\mathrm{p+1}} = \frac{\mathbf{F}_\gamma^\mathrm{p}}{1 + \mathrm{c}\Delta_\mathrm{t}[\sigma_\gamma \mathrm{n_{HI}^{p+1}} + \kappa_\mathrm{dust}\rho_\mathrm{d}^{\mathrm{p+1}}]} \, , \end{cases} \tag{2.10}$$

Where Δ_t denotes the ATON time step that separates steps p and p+1.

Having identified the following lines in ATON where the number density and flux of photons are updated as per Eq. 2.10, all that remains is to add the extra term that comes from dust absorption.

We do not track the dust temperature in our implementation, implying that the radiative energy from the interactions of UV photons with dust is lost. This also means that there is no interaction between the gas and the dust, i.e. : dust is merely advected with the Hydrogen gas as a passive scalar field (like metallicity field).

---

[11] The effective energy for the ionising photon group used in RAMSES-CUDATON was computed using the SEDs of the BPASSv2.2.1 stellar evolution model [Eldridge and Stanway, 2020]: the effective energy is the average ionising energy of a stellar population, weighted by the SED at each energy, and integrated over the stellar population's lifetime.

[12] very roughly the wavelength that corresponds to the energy of 20.28 eV photons, in fact it should be closer to 611 Å but was set at 700 Å where the attenuation coefficient is higher during development of the code so as to confirm that the new absorption was indeed working. In the end the difference with 611 Å is slight, nevertheless it has since been patched to the correct value.

[13] The use of values from the Local Universe can be surprising at first, it is however a compromise: first, because only the dust content and extinction of local objects such as the SMC are well constrained enough for our purposes; second, because the SMC and LMC are relatively low metallicity satellite galaxies, they can be thought of on some levels as analogues for the reionising galaxies we wish to study.

2.6.2 Post-processing with dust

Extinction

To produce dust extinction values for computing the UV continuum slopes and UV magnitudes of galaxies, I employ a simple approach to mimic observations : dust optical depths are computed[14] along a single line of sight per star forming cell halo in one direction. Since the crossing of another dusty object by a single line of sight is unlikely, and since we cannot account for the breadth and evolution of the universe between the redshift of emission and the redshift of observation, dust opacities are computed up to a distance of 2× r_{200} from the source cell (r_{200} is determined using the mass of the halo that each star forming cell is associated with). r_{200} is chosen as a limit so that stellar particles on the opposite side of the galaxy from the observation point, have extinction computed based on the galaxy's full LoS column density of dust. Finally, every stellar particle is attributed the extinction of its host cell. The left panel of Fig. 2.16 is a representative diagram of this process, where the mock observation happens on the left of the halo in the centre of the figure. Not only does this approach yield extinctions that are more faithful to observations than angular averaged ones, but it is also computationally cheaper.

The right panel of Fig. 2.16 compares the distribution of uncorrected magnitudes and the dust extincted magnitudes at z=6 in a L16 N1024 box, using the physical dust model, and with a plausible calibration. The dust extincted histogram's brightest bins (<-20) count less than half of the objects as in the uncorrected case, thereby illustrating that our dust model can produce significant extinction.

Dust escape fractions

To compute the dust escape fractions of our haloes, or the fractions of produced LyC photons that traverse the colums densities of dust within r_{200} in every halo, we reprise the method developed for the HI escape fractions in Lewis et al. [2020], that we briefly recalled in Sec. Escape fractions. Summarily : $f^{\mathrm{ray}}_{\mathrm{esc,dust}}$ is computed using the transmission along 768 rays from each star forming cell to r_{200} . $f^{\mathrm{ray}}_{\mathrm{esc,dust}}$ is then obtained for each halo by averaging over the transmissions, weighting each cell's contribution by its total ionising luminosity (or photon production). To compute the transmissions, an opacity is required; we pick the dust mass attenuation coefficient of Draine and Li [2001] using data from the SMC at $\lambda \approx 700\,\text{Å}$ (or $\kappa_{\mathrm{dust}} = 8.85 \times 10^4\,\mathrm{cm}^2 \cdot \mathrm{g}^{-1}$).

Dependence on the mass attenuation coefficient (κ)

The computed values of $f^{\mathrm{ray}}_{\mathrm{esc,dust}}$ are susceptible to the value of dust mass attenuation coefficient , or κ_{dust} that we take from Draine and Li [2001]. Indeed, for two transmissions, T_1 and T_2, that differ only in the chosen κ_{dust} κ_1 and κ_2, then $T_1 = T_2^{\frac{\kappa_1}{\kappa_2}}$. Tab. 2.6.2 shows the different opacities in Draine and Li [2001]. Throughout this work, we use the SMC values. If we had picked another, then a transmission T=0.5 with the SMC κ_{dust}, would be ≈0.31 for the LMC values, or 0.68 for the MW ones. Massive haloes ($\geq 10^{11}\,M_{\odot}$) in CoDa 2.5 have $f^{\mathrm{ray}}_{\mathrm{esc,dust}}$ values between 0.1 and 0.01 between

---

[14] Using dust mass attenuation coefficient derived from observations of the SMC from Draine and Li [2001] shown in Tab. 2.6.2.

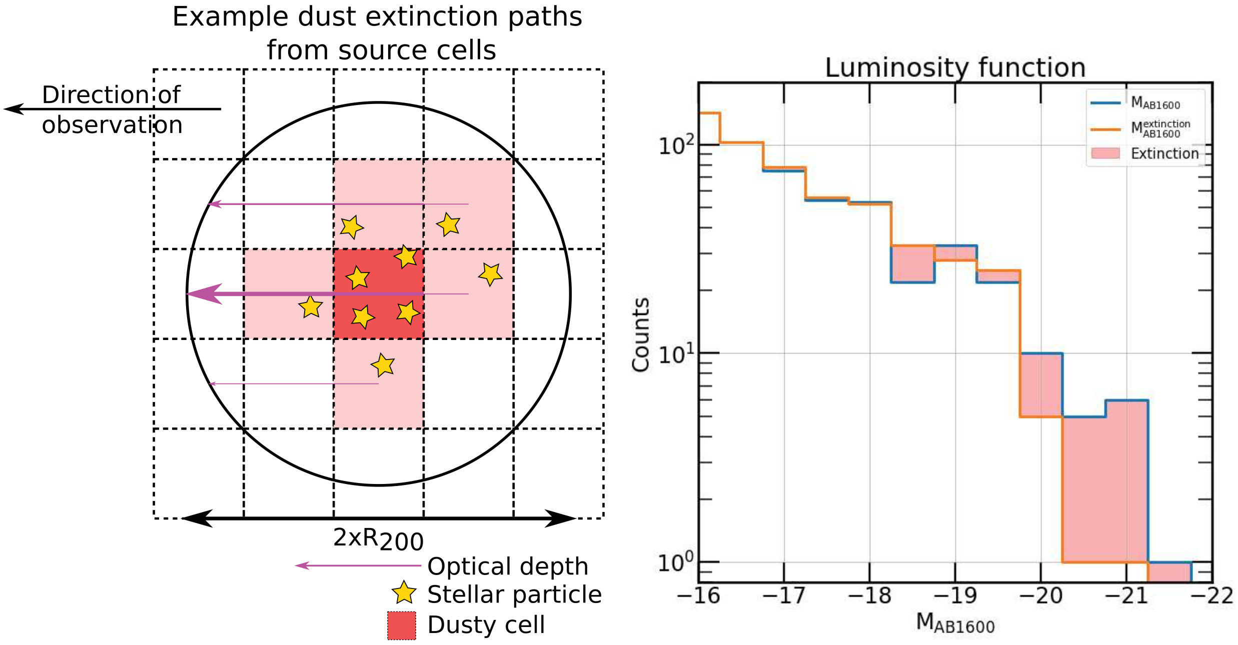


Figure 2.16: *Left*: Schematic of the computation of extinction. Paths are cast from the centres of star forming cells along an observation direction. Dust optical depths are computed along said paths until r_{200} (our definition of the limit of the galaxy) using the mass attenuation coefficient at $\lambda = 1600\,\text{Å}$ derived by Draine and Li [2001] using data from the SMC. The shading of cells represents an arbitrary dust distribution, with redder hues denoting high dust densities, and white symbolising no dust. The purple arrows represent the LoS cast to compute extinction. Their width represents the cumulative optical depth from the origin to r_{200}. *Right*: Luminosity function with and without extinction in a L16 N1024 box with per side at z=6.

	$\kappa_{\rm dust}({\rm cm}^2 \cdot {\rm g}^{-1})$			
Wavelengths	700 Å	1492 Å	1600 Å	2621 Å
LMC	2.19×10^5	4.89×10^4	4.16×10^4	2.2×10^4
SMC bar	1.3×10^5	4.98×10^4	3.83×10^4	8.21×10^3
MW	7.15×10^4	2.33×10^4	2.3×10^4	1.58×10^4

Table 2.1: Mass attenuation coefficient ($\kappa_{\rm dust}$) taken from tables of Draine and Li [2001]. For the LMC values : `ftp://ftp.astro.princeton.edu/draine/dust/mix/kext_albedo_WD_LMCavg_20`. For the SMC bar values : `ftp://ftp.astro.princeton.edu/draine/dust/mix/kext_albedo_WD_SMCbar_0`. For the MW values : `ftp://ftp.astro.princeton.edu/draine/dust/mix/kext_albedo_WD_MW_5.5B_30`. 700 Å very roughly corresponds to the effective wavelength of the ionising photon group tracked in CUDATON. 1600 Å is used to compute the $M_{\rm AB1600}$ of galaxies (and the UVLF). 1492 Å and 2621 Å are used to compute the slope of the UV continuum for galaxies.

z=6 and z=5, meaning there are probably individual dust transmissions much lower than 0.5 in those haloes. If T=10^{-2} with the SMC $\kappa_{\rm dust}$, it would be 4.3×10^{-4} with the LMC $\kappa_{\rm dust}$, or 7.9×10^{-2} with the MW $\kappa_{\rm dust}$. Since massive haloes have very large intrinsic luminosities, such a large variation in $f^{\rm ray}_{\rm esc,dust}$ within a single cell can dramatically change the number of photons that reach the IGM. It would definitely be worthwhile to explore the effects of different assumptions on dust grain $\kappa_{\rm dust}$ in our work. In particular, letting the mass attenuation coefficient vary with temperature, which would follow the local UV density.

2.7 Evaluating the validity of the physical dust model : comparison to observations and semi-analytical models

The two parameters of the physical dust model must be set to match the few observational constraints available. Recent observational breakthroughs (with ALMA for instance) give us access to probes of the dust content in galaxies earlier than z=4 [for instance Burgarella et al., 2020; Schaerer et al., 2015; Béthermin et al., 2015]. At the same time, we check how our results compare to semi-analytical modelling and other simulations [such as Popping et al., 2017; Vijayan et al., 2019; Mancini et al., 2015; Graziani et al., 2020]. Throughout the rest of this section, we shall examine our best matching run : a L16 N1024 box, with $f_{\rm cond}$ of 0.001, and a max(DTM) of 0.5. The resolution and set up of the box are almost identical to that of CoDa II, though in a box 4 times smaller on each side. In our best matching run, star formation is twice as efficient as in CoDa II ($\epsilon_\star$=0.04), there is a star formation temperature criterion of $T_\star=3 \times 10^4$ K, and a slightly higher sub-grid escape fraction of 0.5.

2.7.1 Dust mass to stellar mass

Fig. 2.17 shows the average dust mass to stellar mass in our simulation at several epochs. Generally, dust mass increases on average with increasing stellar mass. As observed in smaller scale simulations, between $2 \times 10^6\,M_\odot$ and $2 \times 10^7\,M_\odot$, the slope of the average relation increases sharply before falling back to returning to gentler values. More directly, this means that haloes with stellar masses >

$2 \times 10^7\,\mathrm{M}_\odot$ have, on average, dust masses equal to roughly 10^{-2} times their stellar masses, whereas haloes below $2 \times 10^6\,\mathrm{M}_\odot$ of stellar mass have average dust mass to stellar ratios of between 10^{-4} to 10^{-5}. Interestingly, the average relation does not change considerably over time. Although, it seems that the stellar mass threshold separating low dust mass to stellar mass halos from high dust mass to stellar mass haloes, moves to higher stellar mass as the simulation progresses. Finally, the scatter around the average is very low, as indicated by the 25th and 75th percentile lines. As in Sec. Dust model (Fig. 2.14), we can match a mix of constraints and upper limits gathered by Mancini et al. [2015], as well as those from Burgarella et al. [2020]. This means our results also line up reasonably well with their predictions for the relation between dust mass and stellar mass. For stellar masses from $10^7\,\mathrm{M}_\odot$ to $10^{10}\,\mathrm{M}_\odot$, our average curves agree well with the average relation between dust mass and stellar mass reported by Popping et al. [2017] in their high-condensation semi analytical model. Our results also agree with their fiducial SAM, but only near $10^9\,\mathrm{M}_\odot$ of stellar mass. Further, it appears that their average stellar mass to dust mass relation also varies little with redshift. Vijayan et al. [2019] report lower (by ≈ 0.5 dex) dust masses versus stellar masses for $z>3$ (and at $z=6$ as shown here). Interestingly, if we recall Fig. 2.14, our physical model matches the dust mass to stellar mass given by post-processing our simulation with the fit to their model. However, due to our box size, we cannot compare our results with the higher stellar mass portions of these SAM results. Graziani et al. [2020] use the SPH simulation code dusty gadget to investigate the dust masses of early ($z>3.5$) galaxies. They report less dusty galaxies than ours. In particular, the abrupt transition in dust mass as a function of stellar mass that we find near $2 \times 10^6\,\mathrm{M}_\odot$, occurs at stellar masses of almost $4 \times 10^8\,\mathrm{M}_\odot$ in their work. This means that our galaxies from 2×10^7 to $4 \times 10^8\,\mathrm{M}_\odot$ have dust masses up to 100 times greater than their equivalents in Graziani et al. [2020]. It is noteworthy that they also find this 'S shape' like curve with an abrupt shift in the average dust mass to stellar mass ratio. This suggests that the models in our simulations are very similar, but that our results diverge at low stellar mass due to different parametrisations and/or resolution. Despite the outlined differences, for masses $>10^9\,\mathrm{M}_\odot$, our results agree fairly well. Finally, their relation between stellar mass and dust mass shows little to no redshift evolution, which is similar to our findings.

2.7.2 Dust to metals and dust to gas

The left panel of Fig. 2.18 shows the average halo dust to metal ratios for various redshifts. Average DTMs are strikingly constant for stellar masses below $2 \times 10^6\,\mathrm{M}_\odot$, and for those above $2 \times 10^7\,\mathrm{M}_\odot$. Over this interval, the average DTM rises from just over 10^{-3} to around 0.3. This regime change in DTM corresponds to the one that occurs in the stellar mass to dust mass relations: the low stellar mass galaxies have low DTMs set by $\mathrm{f}_{\mathrm{cond}}$ (the fraction of metal mass ejected as dust by supernovas), whereas the high stellar mass galaxies have high DTMs set by max(DTM). Accordingly, the height of the first plateau is set by the value of $\mathrm{f}_{\mathrm{cond}}=10^{-3}$, and the height of the second plateau sits near the maximum cell DTM that we set in our model (0.5) (but underneath it as not all galactic cells are likely to reach the maximum DTM). The fact that the low stellar mass plateau is exactly $\mathrm{f}_{\mathrm{cond}}$ again shows that the dust masses in these galaxies come almost exclusively from supernova ejecta. At the same time, the scatter around the averages is very small as indicated by the 25 and 75 percentile lines

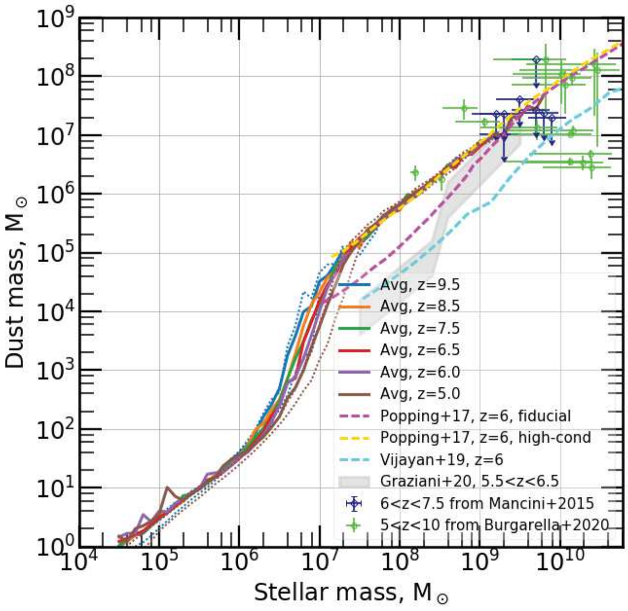


Figure 2.17: Average total dust mass of haloes as a function of their average total stellar mass at various redshifts. The dotted lines denote the 75th and 25th percentiles for z=9.5 and for z=5. The blue upper limits are from Mancini et al. [2015], the green data are from Burgarella et al. [2020]. The gold (magenta) curve shows the z=6 high-cond (fiducial) SAM result of Popping et al. [2017]. The cyan curve shows the z=6 average result from the SAM of Vijayan et al. [2019].

(coloured dotted lines) for z=5 and for z=9.5. Again, there is little to no difference in the behaviour of the different curves as a function of time. However, it does seems that as time goes on, the reported stellar mass transition interval shifts slightly to higher stellar masses. In their equivalent curves, the average DTMs of the fiducial model of Popping et al. [2017] also appear to increase rapidly between two plateaus of constant average DTM. But, in their results, this occurs at much higher stellar masses, between 5×10^8 M$_\odot$ and 5×10^9 M$_\odot$. Our results do not posses such a feature at those masses. Further, we overshoot the average DTMs of the galaxies in their fiducial model by about a factor of 2. However our high M$_\star$ galaxies' DTMs are a good match to the predictions of their high-cond and no-acc results. Vijayan et al. [2019] find much lower (up to 1 dex lower) DTM ratios at similar times and masses. For z>4 the average DTM in their SAM increases between 10^8 to 10^{10} M$_\odot$. By z=3, the average DTM that they report has increased substantially and is close to our own. At these later times, they also report close to no evolution of the average DTM with halo mass $\gtrsim 10^7$ M$_\odot$. In both of these SAMs, the transition in average DTM between lower stellar mass galaxies and higher stellar mass galaxies is explained by highlighting a corresponding increase in the stellar mass weighted stellar age in galaxies.

The right panel of Fig. 2.18, shows the average halo dust to gas ratio as a function of stellar mass for various epochs. Interestingly, the average DTG exhibits trends that are exceedingly similar to those of the average dust mass to stellar mass relation, and presents an initial slow rise with stellar mass from 2×10^{-8} at 4×10^4 M$_\odot$ to 10^{-7} at 2×10^6 M$_\odot$, followed by a sharp increase to 6×10^{-5} at 2×10^7 M$_\odot$, and finally ending in another slow rise till 4×10^{-4} at 2×10^9 M$_\odot$. Once again, the percentile curves (coloured dotted lines) show very little scatter around the average relation as a function of stellar

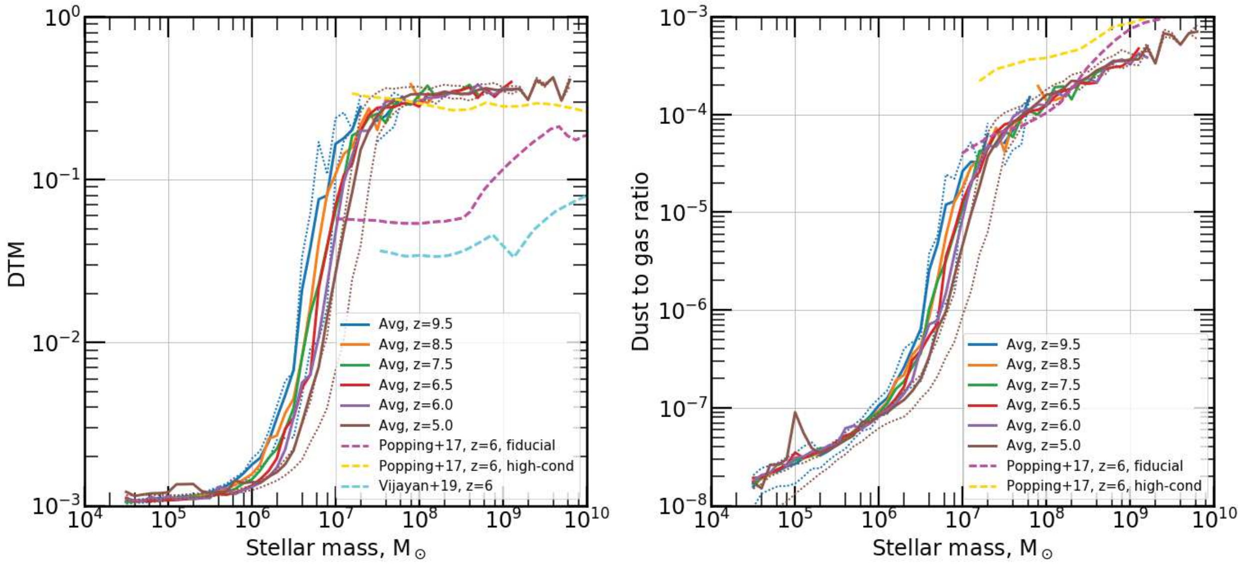


Figure 2.18: *Left*:Average halo dust to metal ratios as a function of halo stellar mass for various redshifts. The dotted lines denote the 75th and 25th percentiles for z=9.5 and for z=5. The gold (magenta) curve shows the z=6 high-cond (fiducial) SAM result of Popping et al. [2017]. The cyan curve shows the z=6 average result from the SAM of Vijayan et al. [2019]. *Right*:Average halo dust to gas ratios as a function of halo stellar mass for various redshifts. Same as on the left for the dotted lines. The gold (magenta) curve shows the z=6 high-cond (fiducial) SAM result of Popping et al. [2017]

mass. The results of Popping et al. [2017] are close to ours : their fiducial average DTG is very similar to our own from $10^7\,M_\odot$ to $2 \times 10^8\,M_\odot$, but overshoots our result at higher masses. Their high-cond SAM is systematically higher (by up to ≈ 0.5 dex) than our average DTGs at similar masses. They also report little change over the course of Reionization.

Over all, the agreement with the findings of Popping et al. [2017] is good, particularly with their high-cond SAM, except for the average DTG of galaxies that we predict to be lower for our galaxies with the most stellar mass. We find a lesser agreement with their fiducial SAM, especially when comparing our average DTMs, as theirs are much lower. As with the average dust mass to stellar mass relation, we overshoot the DTM results of Vijayan et al. [2019] during Reionization.

2.7.3 Dust mass function

Fig. 2.19 shows the dust mass function at several redshifts. The simulation is too small to feature the largest dust masses in the comparable plot of Popping et al. [2017] (Their Fig. 7 spans a higher dust mass range from 10^6 to $10^{10}\,M_\odot$). That being said, our dust mass functions (DMFs) are comparable to those of their models. In particular, at z=5, our DMF is near 10^{-2} at $10^6\,M_\odot$, as are their high condensation and no accretion models. However, the slope in our DMF is slightly more pronounced at z=5 : our DMF reaches just above 10^{-3} at $10^7\,M_\odot$ where theirs are just over 3×10^{-3} , though this could be a limitation due to our box size, visible by the strong cut off in dust masses near $\approx 2 \times 10^7\,M_\odot$. The match between DMFs at z=6 and $10^6\,M_\odot$ is good, as our DMF is around 10^{-2}, which is close

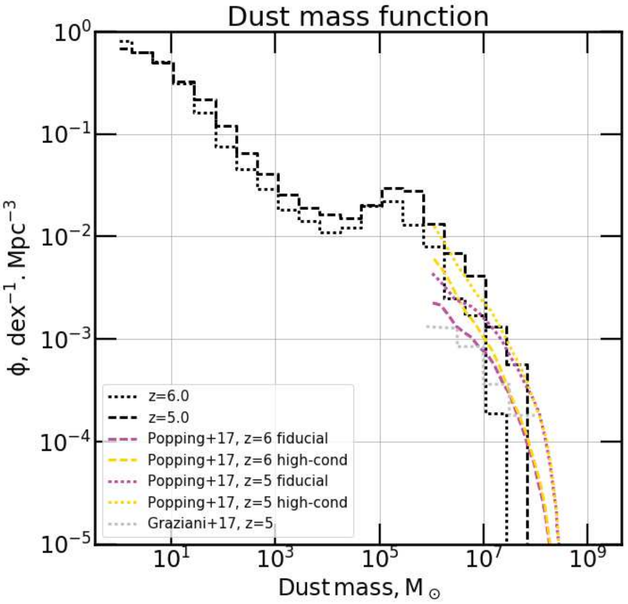


Figure 2.19: Dust mass function of galaxies at several redshifts. z=5 results are plotted as dotted lines, whereas dashed lines show findings from z=6. Our findings are in black. The gold (magenta) curve shows high-cond (fiducial) SAM predictions of Popping et al. [2017]. The silver curve shows the z=5 average result from the simulations of Graziani et al. [2020].

to their models. The DMF of Graziani et al. [2020] is much flatter than ours, probably due to much lower dust masses in all but their highest stellar mass galaxies. Their DMF is also up to ≈ 0.7 dex lower than ours before 10^7 , plausibly for similar reasons.

The bump like feature near $10^5\,\mathrm{M}_\odot$ is intriguing, and nothing comparable can be seen at the DMF at higher masses in SAMs. The dust masses at which the feature appears are within the range at which a fast transition in dust mass seen in the average dust mass to stellar mass relation (and in the DTM versus stellar mass relation) occurs. Therefore, this bump is most likely associated with the previously described jump in the capacity of galaxies to accumulate dust (there is no corresponding jump in the stellar mass function, or the halo mass function).

It would be of great interest to perform the same comparison with L32 N2048 or L64 N4096 boxes. In particular, because our most massive star forming galaxies have the best agreement with SAM and other simulation DMFs. This could be achieved cheaply by running AMR enabled hydro+DM only simulations.

2.7.4 UV continuum slopes

It has been shown that slope of the UV continuum (or β) is sensitive to the presence of dust, due to the reddening that it produces. We compute the slope of the UV continuum between $\lambda,\mathrm{high} = 2625$Å and $\lambda,\mathrm{low} = 1492$Å[15]. Using the fluxes $\mathrm{f}^{\lambda}(\mathrm{age, metallicity})$ of our stellar particles from our stellar

---

[15] In observations, β is determined using a linear fit to the UV spectra of galaxies. We choose to determine the luminosity of galaxies at two UV wavelengths and use them to infer a slope that is close to the result of such a fit. The exact wavelength values were chosen so as to avoid typical spectral features that could invalidate our result.

emissivity model (BPASS V2.2.1 Eldridge and Stanway [2020]), and the dust absorption computed as in Sec. Post-processing with dust, we can then compute β as follows Eq. 2.11:

$$\beta = \frac{\log_{10}\Big(\sum_{\mathrm{i}}^{\mathrm{N}_\star} \mathrm{m_i f^{\lambda_{high}}(A_i, Z_i, T^{\lambda_{high}}_{dust,i})} \Big) - \log_{10}\Big(\sum_{\mathrm{j}}^{\mathrm{N}_\star} \mathrm{m_j f^{\lambda_{low}}(A_j, Z_j, T^{\lambda_{low}}_{dust,j})} \Big)}{\log_{10}\Big(\lambda_{\mathrm{high}}/\lambda_{\mathrm{low}}\Big)} \tag{2.11}$$

With i,j the indices we use to sum over the $\mathrm{N}_\star$ stellar particles in galaxies, m the stellar particle mass, A the stellar particle age, and $\mathrm{T}^{\lambda}_{\mathrm{dust}}$ the transmission through galactic dust at the wavelength λ.

Fig. 2.20 shows the relation between the slope of the UV continuum and magnitude at 1600Å of our simulated galaxies, against observations from Finkelstein et al. [2012b]. Before z=5, low brightness galaxies (>-12.5) have highly scattered β values. In fact, since the highest magnitude galaxies are those with only a few stellar particles, they follow discrete tracks that correspond to the reddening of galactic light as the stellar particles age. Young faint galaxies appear in β-magnitude space at low β when their first stars forms. As they age, the galactic β and magnitude both increase. One might expect there to be a discrete track for every possible number of stellar particles. That this is not always visible here is partly due to the logarithmic scale of magnitudes, and to the fact the magnitude and β of galaxies are scattered by the metallicity of stellar particles. Between z=6 and z=5, the region with β<-2.25 and $\mathrm{M_{AB1600}}$ >-11 is emptied of galaxies. This area of the plot corresponds to low mass galaxies with one or a few young and blue stellar particles. The dramatic depopulation of such galaxies is a tell tale sign of star formation suppression.

For galaxies brighter than -12.5, at all times and when not accounting for the UV reddening due to dust, the simulated β-magnitude relation is flat and near -2.75. However, when we account for the reddening produced by dust, a marked upwards trend in β appears for galaxies brighter than -16. This evolution with magnitude places the β within range of the observations between z=8.5 and z=7.5, that also present an upwards trend with magnitude between -16 and -19. Despite this success, there appears to be too much reddening for some of the brightest galaxies in the simulation, meaning the brightest observed galaxies have less reddening due to dust on average than we predict. Indeed, whereas our simulated βs grow at an exponential rate from -16 and brighter, the average points from Finkelstein et al. [2012b] oscillate around $\beta = -2.0$ for $\mathrm{M_{AB1600}}$<19. Already at z=7.5, the reddened β of some haloes is outside of the errorbars (standard deviation in the magnitude bin) of the points from Finkelstein et al. [2012b], with several galaxies almost reaching 0.25. By z=6.0, there are many points above the average constraints near -19 in magnitude, they span from β=-2.75 to almost 0.5 At z=5, the high scatter is maintained amongst the present brighter galaxies. Interestingly, the maximum β does not increase significantly between z=6 and z=5. Note, that when comparing the two loci of points at the bright end (<-17), it may appear as though some or most of the black points have no analogues in the red distribution. However, this is simply because the red β are plotted against extincted halo magnitude, whereas the black points do not account for extinction.

Although many of our β values are outside of the observation errobars, the scatter in the Finkelstein et al. [2012a] is considerable. In fact, the observed β values cover a range all the way from -2.5 to 3.0 (for z≥5.0). This means that the very high βs in our simulation aren't incompatible with the observation data per se.

The few galaxies between M_{AB1600}=-18 and -23 have both betas within the observational error bars and within the locus of un-reddened points. It is probably the case that our physically sound, single LoS approach, to measuring extinction and reddening further increases scatter. Indeed, dust free lines of sight could explain the very bright galaxies with close to no reddening, as could be possible in observations.

A noteworthy remark is that the reddening shown here is very susceptible to the assumptions made regarding the opacity of dust and its dependence on wavelength. Using dust opacities that vary less with wavelength could lessen the reddening in the brightest galaxies, and therefore bring our results closer to the observed βs. For instance, it would be good to experiment with the other dust opacities from Draine and Li [2001].

The interpretation of these results in a larger box with a greater sample of very bright galaxies would be clearer.

2.8 Dusty massive haloes

Fig. 2.21 and Fig. 2.22 presents a selection of $\approx 10^{11}\,\mathrm{M}_\odot$ haloes at z=6, from a L16 N1024 box, using the evolving BPASS stellar population model, metal and Helium cooling, the physical dust model of Dubois et al. (in prep) with radiative transfer through the dust in ATON. From top to bottom, and from left to right, the maps show $\log_{10}$(neutral fraction), $\log_{10}$(gas overdensity), $\log_{10}$(gas temperature), $\log_{10}$(young stellar mass (<10 Myr)), $\log_{10}$(stellar mass), $M_{AB16000}$, $\log_{10}$(absolute metallicity), $\log_{10}$(dust density), and the maximum absorption across r_{200}. Each galaxy's r_{200} is represented by a circle centred on the pFoF halo centre. We define the absorption shown in the bottom left of each set of maps as : $A = 1 - \exp[-\tau_{\mathrm{dust}}]$, where τ_{dust} is the optical depth along r_{200} taken perpendicularly to the plane of the map.

Although metals are effectively distributed by the supernova explosions during which they are injected into expanding gas, dust seems to cluster in the densest centre most regions of galaxies. Consequently, dust extinction only happens in the centre most regions of our galaxies, which are the densest, most opaque, and most star forming regions. In fact absorption of LyC photons by dust is only consequential (>0.1) in the most dust dense cells that have more than $10^{-28}\,\mathrm{g}\cdot\mathrm{cm}^{-3}$ of grains. That dust absorption is only effective in the centre most cells of galaxies means that extinction could plausibly flatten the luminosity profile of our massive galaxies by dimming their cores in the UV. The much lower densities of dust present in the hot, ionised, low density regions that surround these galaxies, in spite of their high metallicities, could be a combination of two aspects of our dust model : first, dust can only accrete in sufficiently cool ($T<10^5$ K) cells; second, dust destruction via thermal sputtering scales with $(1 + \frac{T}{10^6})^3$, meaning that the dust expelled from haloes in the super heated supernova blast waves can be efficiently destroyed. It may be that dust could propagate more readily to the IGM and CGM in models where the dynamics of dust are followed consistently with the hydrodynamics. Computing the accretion rate of dust in the same maps could allow us to differentiate between in-situ formation of dust, and the accumulation of dust blown out of galaxies by supernova events.

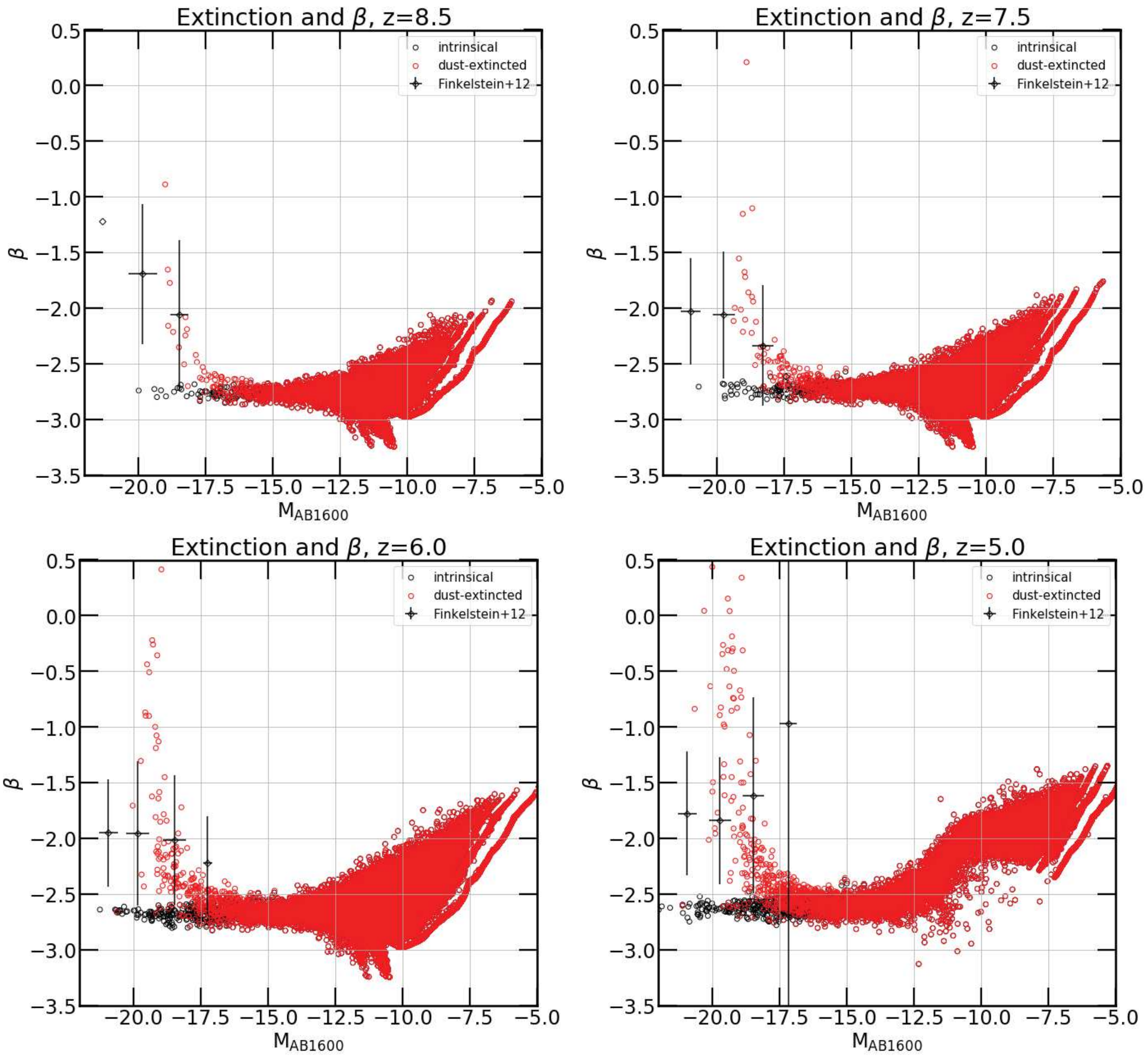


Figure 2.20: Halo UV continuum slopes of simulated galaxies in a L16 N1024box. Red circles represent β when post-processing for dust reddening and extinction, black circles represent the same case with no dust reddening or extinction. Black crosses are data from Finkelstein et al. [2012b], that were binned according to their magnitude. All error bars represent bin standard deviation.

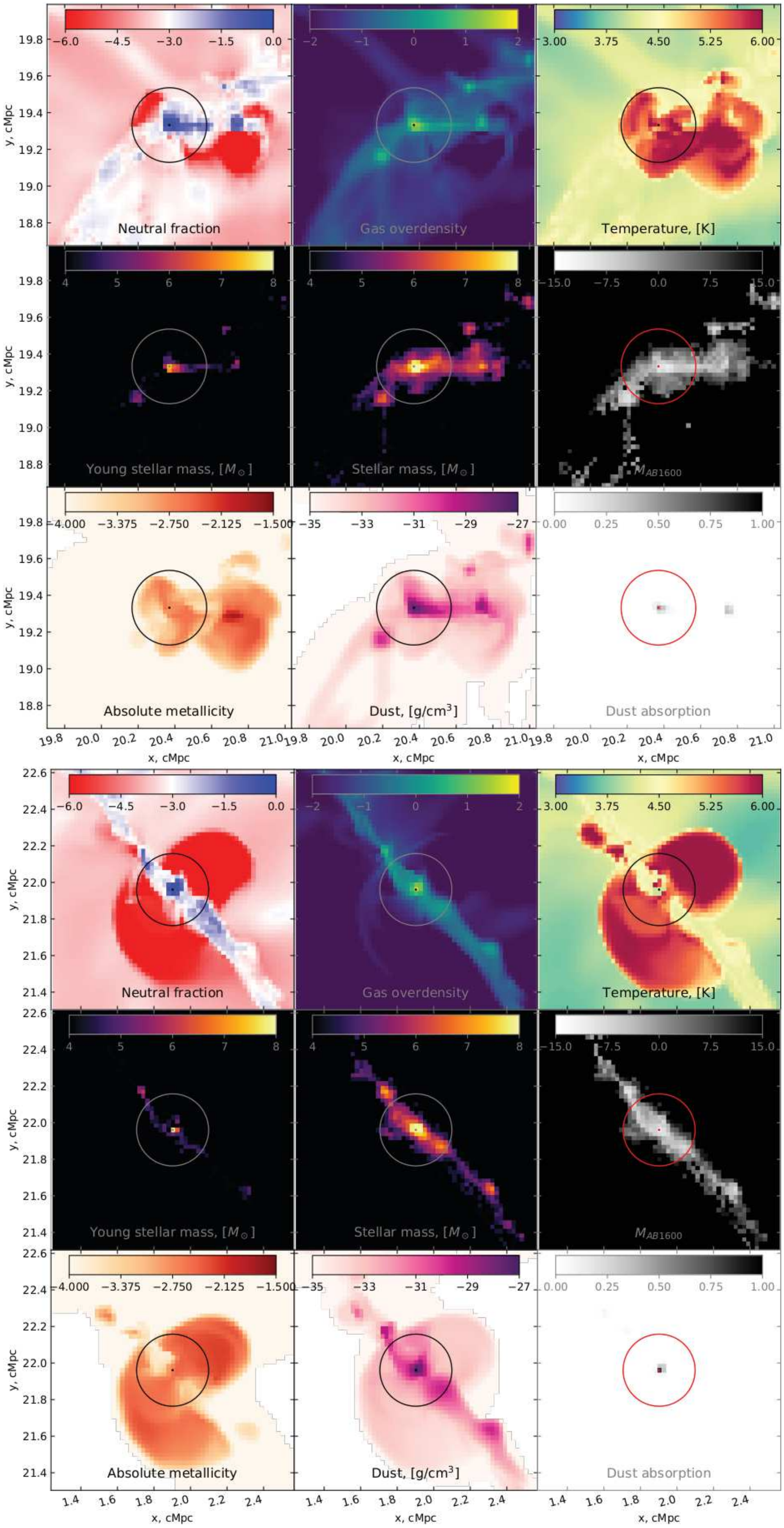


Figure 2.21: Two sets of maps of two massive haloes.

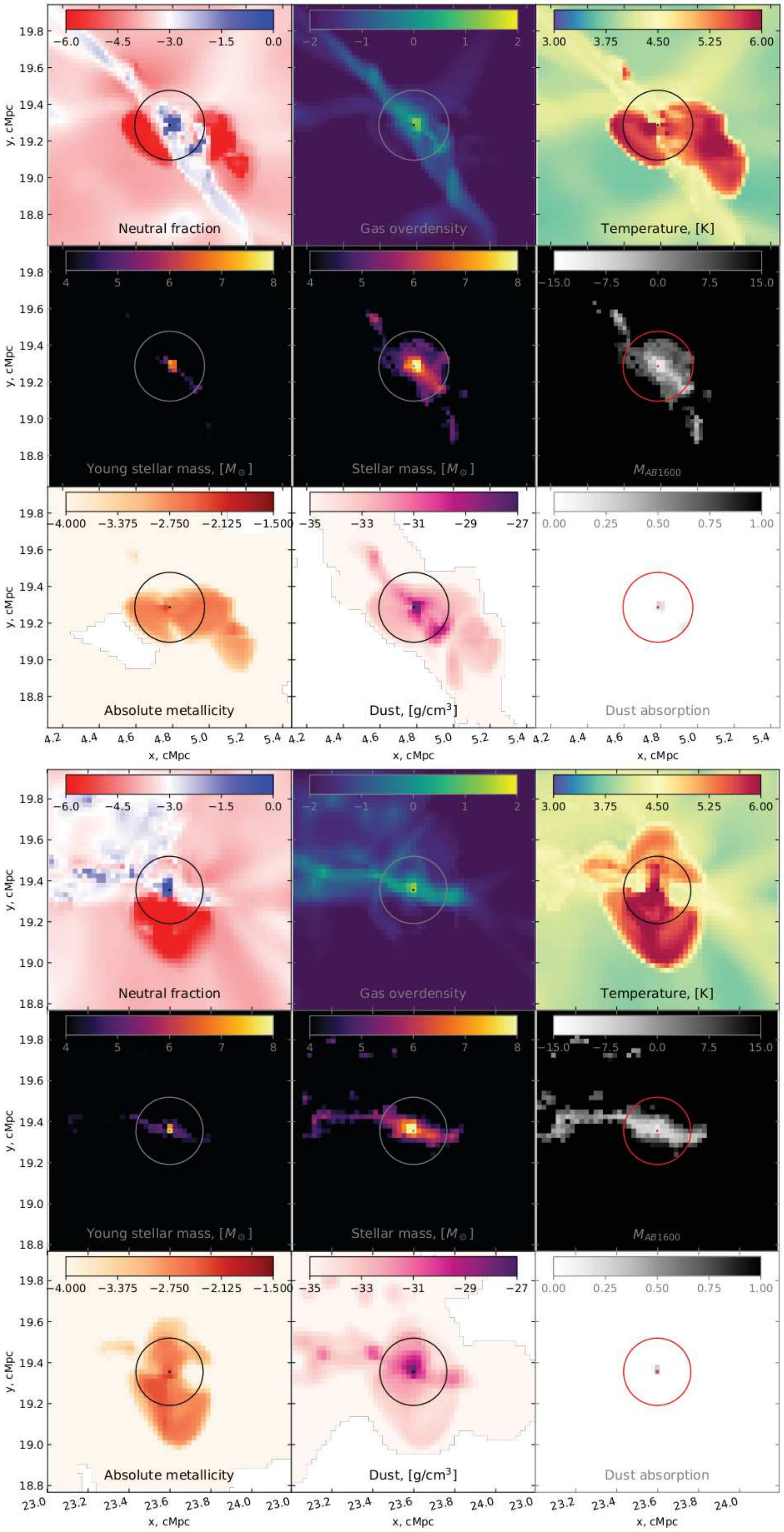


Figure 2.22: Two sets of maps of two massive haloes.

2.9 Résumé des résultats

Dans ce deuxième chapitre, j'ai exposé différents modèles physiques qui faisaient défaut dans CoDa II, puis j'ai motivé et détaillé mon implémentation de ces modèles dans le code de simulation RAMSES-CUDATON. Enfin, j'ai examiné leurs effets sur la réionisation.

En premier lieu, j'ai évoqué la sur-abondance de photons ionisants et d'ionisations à la fin de la Réionisation dans CoDa II par rapport aux contraintes connues. Par ailleurs, l'ionisation du gaz d'hydrogène dans le milieu intergalactique est trop forte à la fin de la Réionisation, et des résultats récents montrent que la Réionisation dans CoDa II est peut-être trop précoce [comme suggéré par Kulkarni et al., 2019].

L'évolution stellaire Une possibilité pour pallier à ces problèmes est de modèliser l'évolution des émissivités ionisantes des populations stellaires en fonction de leur age et de leur métallicité. En effet, en vieillisant et en s'enrichissant en métaux, les modèles d'évolution stellaire prédisent que les populations stellaires émettent moins de photons ionisants par masse stellaire. J'ai montré qu'inclure ces effets pouvait effectivement réduire l'émissivité ionisante moyenne par masse stellaire, retardant la Réionisation, mais conservant l'évolution de la fraction neutre moyenne d'hydrogène.

Refroidissement dû à l'hélium et aux métaux L'accrétion de gaz et la formation stellaire dépendent des mécanismes de réfroidissement du gaz. Bien que le processus principal de refroidissement du gaz (dû à l'hydrogène atomique) ait été représenté dans CoDa II, ni le refroidissement lié à l'hélium, ni celui lié aux atomes métalliques n'y étaient modélisés. Or, le refroidissement dû à l'hélium prend le dessus sur celui dû à l'hydrogène autour de $10^5\,\mathrm{K}$, et le refroidissement lié aux métaux peut être important aux très grandes températures et en présence de fortes métallicités ($\mathrm{T}{>}10^5\,\mathrm{K}$, $\mathrm{Z}\gtrsim 0.1\mathrm{Z}_\odot$). J'ai implémenté ces deux mécanismes en m'inspirant de ce qui a déjà été fait dans des codes similaires. J'ai montré que ces ajouts menaient à un refroidissement plus efficace des régions de gaz très denses ($>10^{-2}\,\mathrm{H}\cdot\mathrm{cm}^{-3}$), et à des densités maximales de gaz plus grandes. Ces différences engendrent d'avantage de formation stellaire dans les galaxies pendant la Réioinsation. L'effet de l'hélium semble être plus important que celui engendré par les métaux.

Poussière, extinction et rougissement Des études récentes ont signalé la présence d'importantes masses de poussière dans les galaxies massives à haut redshift, y compris pendant la Réioinsation [comme dans Burgarella et al., 2020; Schaerer et al., 2015; Béthermin et al., 2015]. De grandes masses de poussière pourraient affecter le déroulement de la Réionisation. En effet, la poussière rougit et diffuse les photons ionisants du continuum Lyman, de manière à potentiellement réduire le nombre de photons ionisants qui pourraient atteindre le milieu intergalactique et l'ioniser. En même temps, d'importantes quantités de poussière pourraient avoir de grandes implications pour l'observation des galaxies lointaines, en atténuant la fonction de luminosité ultra-violette des galaxies par exemple. Ainsi motivé, j'ai cherché à coupler des modèles de poussière au code RAMSES-CUDATON.

J'ai examiné deux modèles, le premier se basant sur un travail semi-analytique [décrit dans Vijayan et al., 2019], et le second utilisant le modèle physique de Dubois et al. 2020 (en préparation)[16]. J'ai montré qu'il était possible d'obtenir des masses de poussière plausibles avec les deux modèles différents. Comme le modèle de Dubois et al. 2020 (en préparation) est mieux motivé physiquement, j'ai préféré son utilisation. En étudiant la distribution de poussière, j'ai constaté qu'elle se forme plus densément dans les régions où le gaz est dense et froid, et où la population stellaire est vieille, ce qui est conforme aux attentes. En ajustant les paramètres du modèle (liés à la condensation de grains de poussière), j'ai abouti à des masses de poussière, des taux de métaux sous forme de poussière, des taux de poussière sur gaz et des fonction de masse de poussière compatibles avec des modèles semi-analytiques récents comme celui de Popping et al. [2017] (notamment le modèle 'high-cond' ou 'haute condensation'), et des résultats de simulations comme Graziani et al. [2020]. Finalement, j'ai étudié l'effet du rougissement sur la pente du spectre des galaxies dans l'UV (β). Il s'avère que même si les observations sont très dispersées, l'inclusion de poussières et du rougissement qu'elles produisent rapproche les valeurs de β de nos galaxies à celles des observations présentées par Finkelstein et al. [2012b].

Ces éléments montrant que la nouvelle description de la poussière est vraisemblable, je me suis intéressé à l'extinction qu'elle pouvait engendrer. Pour calculer l'extinction et le rougissement due aux poussières, j'ai choisi une approche sensible à l'angle de vue de chaque galaxie comme dans les observations : les effets de la poussière sont calculés le long d'un rayon entre chaque particule stellaire et r_{200} dans la direction d'un point d'observation (placé arbitrairement à l'infini, en pratique les lignes de visées suivent la grille cartésienne). Notamment, j'ai constaté que l'extinction due à la poussière pourrait significativement altérer la fonction de luminosité ultra-violette des galaxies pendant la Réionisation (par exemple à z=6, la poussière fait chuter les nombre de galaxies plus brillantes que -20.5 par un facteur d'environ 10).

---

[16] YD l'avait déjà implémenté dans RAMSES et RAMSES-RT. Mon travail a consisté à coupler le champ de poussière produit dans RAMSES avec la partie CUDATON du code, ainsi que d'y implémenter l'absorption due à la poussière.

BIBLIOGRAPHY

Aoyama S., et al.. Galaxy simulation with the evolution of grain size distribution. *Monthly Notices of the Royal Astronomical Society*, 491:3844–3859, 2020.

Aubert D. and Teyssier R. A radiative transfer scheme for cosmological reionization based on a local Eddington tensor. *Mon Not R Astron Soc*, 387(1):295–307, 2008.

Barkana R. and Loeb A. In the beginning: the first sources of light and the reionization of the universe. *Physics Reports*, 349(2):125–238, 2001.

Burgarella D., et al.. Observational and theoretical constraints on the formation and early evolution of the first dust grains in galaxies at $5 < z < 10$. *arXiv*, page arXiv:2002.01858, 2020.

Béthermin M., et al.. Evolution of the dust emission of massive galaxies up to $z = 4$ and constraints on their dominant mode of star formation. *A&A*, 573:A113, 2015.

Calverley A.P., et al.. Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect. *Monthly Notices of the Royal Astronomical Society*, 412:2543–2562, 2011.

Cen R. A hydrodynamic approach to cosmology - Methodology. *The Astrophysical Journal Supplement Series*, 78:341–364, 1992.

Chaabouni H., et al.. Sticking coefficient of hydrogen and deuterium on silicates under interstellar conditions. *Astronomy and Astrophysics*, 538:A128, 2012.

Draine B.T. and Li A. Infrared Emission from Interstellar Dust. I. Stochastic Heating of Small Grains. *The Astrophysical Journal*, 551:807–824, 2001.

Draine B.T. and Salpeter E.E. Destruction mechanisms for interstellar dust. *The Astrophysical Journal*, 231:438–455, 1979.

Dwek E. The Evolution of the Elemental Abundances in the Gas and Dust Phases of the Galaxy. *The Astrophysical Journal*, 501:643, 1998.

Eldridge J.J. and Stanway E.R. Population and spectral synthesis: it doesn't work without binaries. *arXiv e-prints*, 2005:arXiv:2005.11883, 2020.

Eldridge J.J., et al.. Binary Population and Spectral Synthesis Version 2.1: Construction, Observational Verification, and New Results. *Publications of the Astronomical Society of Australia*, 34:e058, 2017.

Fan X., et al.. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132:117–136, 2006.

Ferland G.J., et al.. The 2017 Release Cloudy. *Revista Mexicana de Astronomia y Astrofisica*, 53:385–438, 2017.

Finkelstein S.L., et al.. CANDELS: The Contribution of the Observed Galaxy Population to Cosmic Reionization. *The Astrophysical Journal*, 758:93, 2012a.

Finkelstein S.L., et al.. Candels: The Evolution of Galaxy Rest-frame Ultraviolet Colors from z = 8 to 4. *The Astrophysical Journal*, 756:164, 2012b.

Graziani L., et al.. The assembly of dusty galaxies at z $\geq$ 4: statistical properties. *Monthly Notices of the Royal Astronomical Society*, 494:1071–1088, 2020.

Hoag A., et al.. Constraining the Neutral Fraction of Hydrogen in the IGM at Redshift 7.5. *The Astrophysical Journal*, 878:12, 2019.

Hui L. and Gnedin N.Y. Equation of state of the photoionized intergalactic medium. *Mon Not R Astron Soc*, 292(1):27–42, 1997.

Katz N., et al.. Cosmological Simulations with TreeSPH. *The Astrophysical Journal Supplement Series*, 105:19, 1996.

Kulkarni G., et al.. Large Ly α opacity fluctuations and low CMB τ in models of late reionization with large islands of neutral hydrogen extending to $z < 5.5$. *Monthly Notices of the Royal Astronomical Society: Letters*, 485(1):L24–L28, 2019.

Laporte N., et al.. Dust in the Reionization Era: ALMA Observations of a z = 8.38 Gravitationally Lensed Galaxy. *The Astrophysical Journal Letters*, 837:L21, 2017.

Lewis J.S.W., et al.. Galactic ionising photon budget during the Epoch of Reionisation in the Cosmic Dawn II simulation. *Monthly Notices of the Royal Astronomical Society*, 2020.

Mancini M., et al.. The dust mass in z > 6 normal star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 451:L70–L74, 2015.

Mason C.A., et al.. The Universe Is Reionizing at z $\approx$ 7: Bayesian Inference of the IGM Neutral Fraction Using Lyα Emission from Galaxies. *The Astrophysical Journal*, 856(1):2, 2018.

Ménard B., et al.. Measuring the galaxy-mass and galaxy-dust correlations through magnification and reddening. *Monthly Notices of the Royal Astronomical Society*, 405:1025–1039, 2010.

Novak G. Radiative Transfer, Black Hole Growth, AGN Feedback in Galaxies. 221:103.06, 2013. Conference Name: American Astronomical Society Meeting Abstracts #221.

Novak G.S., et al.. Radiative transfer and radiative driving of outflows in active galactic nuclei and starbursts. *Monthly Notices of the Royal Astronomical Society*, 427:2734–2756, 2012.

Ocvirk P., et al.. Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe. *Monthly Notices of the Royal Astronomical Society*, 463(2):1462–1485, 2016.

Ocvirk P., et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 2020.

Ostriker J. and Silk J. Dust Cooling of Hot Gas. *The Astrophysical Journal Letters*, 184:L113, 1973.

Ouchi M., et al.. Statistics of 207 Lyα Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models. *The Astrophysical Journal*, 723(1):869, 2010.

Popping G., et al.. The dust content of galaxies from z = 0 to z = 9. *Monthly Notices of the Royal Astronomical Society*, 471(3):3152, 2017.

Rosdahl J., et al.. RAMSES-RT: radiation hydrodynamics in the cosmological context. *Monthly Notices of the Royal Astronomical Society*, 436:2188–2231, 2013.

Rosdahl J., et al.. The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Monthly Notices of the Royal Astronomical Society*, 479:994–1016, 2018.

Rosen A. and Bregman J.N. Global Models of the Interstellar Medium in Disk Galaxies. *The Astrophysical Journal*, 440:634, 1995.

Rémy-Ruyer A., et al.. Gas-to-dust mass ratios in local galaxies over a 2 dex metallicity range. *Astronomy and Astrophysics*, 563:A31, 2014.

Schaerer D., et al.. New constraints on dust emission and UV attenuation of z = 6.5–7.5 galaxies from millimeter observations. *A&A*, 574:A19, 2015.

Smith B., et al.. Metal cooling in simulations of cosmic structure formation. *Mon Not R Astron Soc*, 385(3):1443–1454, 2008.

Sutherland R.S. and Dopita M.A. Cooling functions for low-density astrophysical plasmas. *The Astrophysical Journal Supplement Series*, 88:253–327, 1993.

Vijayan A.P., et al.. Detailed dust modelling in the L-Galaxies semi-analytic model of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, page stz1948, 2019. ArXiv: 1904.02196.

Wise J.H., et al.. The Birth of a Galaxy - III. Propelling reionisation with the faintest galaxies. *Monthly Notices of the Royal Astronomical Society*, 442(3):2560–2579, 2014.

Wiseman P., et al.. Evolution of the dust-to-metals ratio in high-redshift galaxies probed by GRB-DLAs. *Astronomy & Astrophysics*, 599:A24, 2017.

Wu X., et al.. Photometric properties of reionization-epoch galaxies in the Simba simulations. *Monthly Notices of the Royal Astronomical Society*, page staa1044, 2020. ArXiv: 1911.06330.

Towards Cosmic Dawn III : Recalibration and matching the observed Lyα forest opacities

3

3.1 Beyond CoDa II and Lyα forest constraints

In Chp. 2, we discussed the shortcomings of CoDa II with respect to its agreement with constraints on the average neutral fraction at the end of Reionization. In short, <xHI> is significantly lower at the end of the EoR (by over a factor of ten at z<6). At the same time, the density of photons is too high. This is a common failing of Reionization simulations as pointed out in Ocvirk et al. [2020]. In Chp. 2, we detailed various new features implemented into the simulation code, that have the potential to address this excess ionisation. Here we tackle this issue from the perspective of advances in the community regarding constraints on Reionization, and explore different simulation calibrations.

One of the most cited efforts to constrain the neutral fraction of Hydrogen is Fan et al. [2006]. However, the oft used constraints therewithin are model dependant, and require assumptions about the density of the IGM. To forgo these assumptions, one must directly reproduce the observations of Fan et al. [2006] by computing effective optical depths due to Lyα absorption along simulated lines of sight, i.e.: Essentially producing mock Lyα forest spectra. Recently, observations such as Becker et al. [2018]; Bosman et al. [2018], have shown that there remain large Lyα effective optical depths even after z=6, highlighting the lingering existence of large bubbles of neutral Hydrogen. Subsequent simulations have shown that this has implications for the spatial distribution and luminosity of ionising sources, possibly requiring very bright and rare sources such as AGN to supplement Reionization by galaxies [as in Chardin et al., 2015, 2017]. Kulkarni et al. [2019] showed that by striving to match the average effective optical depths derived from Lyα observations, as well as their cumulative distribution, requires late Reionization (z≈5.5) in a galaxy only simulation (no AGN).

Fig. 3.1 shows the average effective Lyman alpha opacity in CoDa II near z=5.75 (orange diamond without errorbars), computed thanks to the method outlined in Theuns et al. [1998] (That we expand upon in Sec. New parametrisation). Constraints from various observations (see caption) are also shown. CoDa II effective optical depths are too small by a considerable margin. To better fit the data, we could reionize later as in Kulkarni et al. [2019]; Keating et al. [2019]. This is akin to shifting the CoDa effective Lyα points towards lower redshifts, and therefore reducing the gap between the data and the simulation. In another approach, we will experiment with the sub-grid star formation recipe, driving the amount of suppression in low mass haloes upwards. This would reduce the amount of ionising photons, thus lowering average ionisation of gas and raising the average encountered optical depths due to neutral hydrogen at a given redshift, which would also naturally delays Reionization.

3.2 The new calibration

3.2.1 New parametrisation

One effective approach to amplify feedback on star formation at the end of Reionization, is to introduce a temperature criterion for star formation. This would drastically increase star formation suppression by forbidding it in cells that have been heated and ionised by radiation or by supernova explosions. The physical motivation for a temperature switch for star formation is simple : We know stars form from cold gas, therefore star formation in a hot cell can be expected to be less efficient. The corresponding theoretical temperature limit for star formation is $2 \times 10^4\,\mathrm{K}$ [sub-grid recipes are used extensively for star formation Rosdahl et al., 2018; Trebitsch et al., 2020; Ocvirk et al., 2014]. Such a criterion was implemented in CoDa I, but was relaxed in CoDa II to make star formation more permissive and achieve earlier Reionization. Here, I experimented with a different set-up : with a temperature limit at $3 \times 10^4\,\mathrm{K}$ [1]. In order to produce enough stars in spite of our new stricter conditions for star formation, I doubled the star formation efficiency with respect to CoDa II ($\mathrm{eps}_\star$=0.04). Without this tweak, our predictions tend to undershoot the observed UVLF. Finally, setting a sub grid escape fraction of 0.5 leads to a satisfactory late timing of Reionization. At the same time, these new runs include the features I developed in Chp. 2 : evolving stellar emissivities, metal cooling, Helium cooling, dust production and dust RT absorption.

3.2.2 Matching the Lyα forest

Aiming to produce Lyα optical depths that are comparable to observations such as Bosman et al. [2018], I followed a similar procedure as Chardin et al. [2015], (originally outlined in more detail in Theuns et al. [1998]). For every simulation snapshot, I computed the average Lyα transmission along 12 lines of sight staggered throughout the simulation data cubes. These transmissions can then be converted into effective optical depths by averaging the transmission through the Lyα forest T for a given line of sight : $\tau^{\mathrm{eff}}_{\mathrm{ly}\alpha}$=$-\log(< \mathrm{T} >_{\mathrm{LoS}})$. To match the $50\,\mathrm{Mpc}\cdot\mathrm{h}^{-1}$ of integrated depth in the observations [as in Becker et al., 2015a] , I extended the lines of sight further than the data cubes' size of $16\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$ by exploiting their periodicity. In order to avoid sampling the same regions of simulated space multiple times, I shot lines of sight at a 40° angle from the x axis of the Cartesian grid, within the (x,y) plane. Fig. 3.2 shows an example Lyα path as described above.

The left panel of Fig. 3.1 shows the effective Lyα optical depth in a L16 N1024 box with the above parametrisation, that we shall call CoDa 2.5 from now on. As Reionization progresses, the average simulated effective optical depth, shown in diamonds, or $< \tau^{\mathrm{eff}}_{\mathrm{ly}\alpha} >$, decreases from over 7 at z=6 to just under 1 at z=4. Over the same interval of time, the scatter around $< \tau^{\mathrm{eff}}_{\mathrm{ly}\alpha} >$ decreases as remnant neutral regions ionise, and the number of neutral opaque LoS outliers decreases. At redshifts of 4, 5, and 5.3 the simulated average effective optical depths lie in the centre of the locus of observational points. However, at z=6 and z=5.7, the average simulated data lies at the opaque edge of the observation data, and exhibits high scatter. This could be the sign that although CoDa 2.5 finished Reionization

---

[1] I picked this temperature, which is higher than the temperature used in CoDa I, because I found it gave a better match to the observations of the Lyα forest

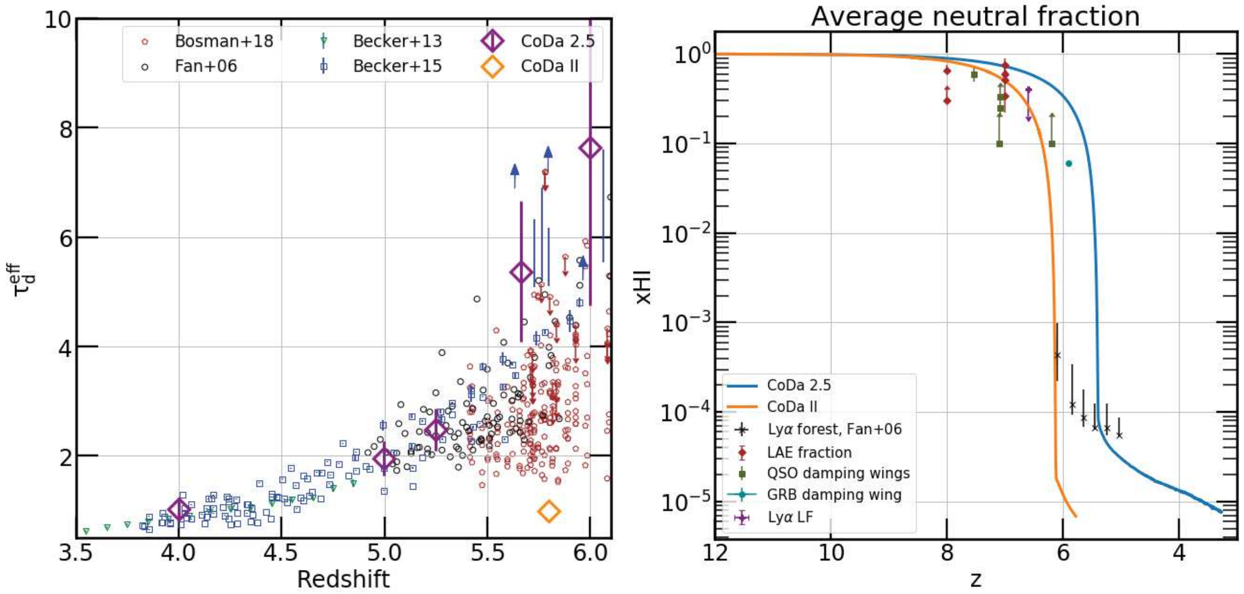


Figure 3.1: *Left*: Average effective Lyα optical depths as a function of redshift in a L16 N1024 box, with a temperature star formation criterion $\mathrm{T_{sf}} = 3 \times 10^4$ K. Simulation data are in magenta. The green points are from Becker et al. [2013], the blue squares from Becker et al. [2015a], the black points from Fan et al. [2006], the red points are from Bosman et al. [2018]. *Right*: Average xHI in CoDa II and in our CoDa 2.5 as a function of redshift.

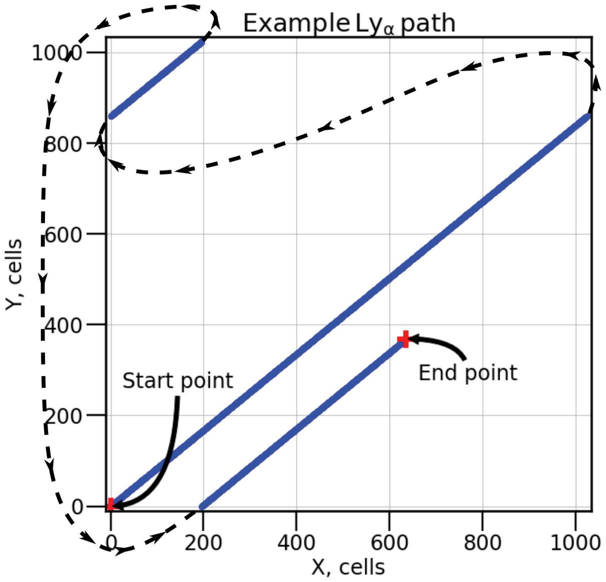


Figure 3.2: Figure demonstrating an example path for the starting point X,Y,Z=(0,0,0) and an angle of 40° in a 1024 cell sided box. The dotted lines with arrows show how the path continues once it reaches the box's edge, which it does twice before reaching the end point. Obviously, no integration is performed along these dashed paths.

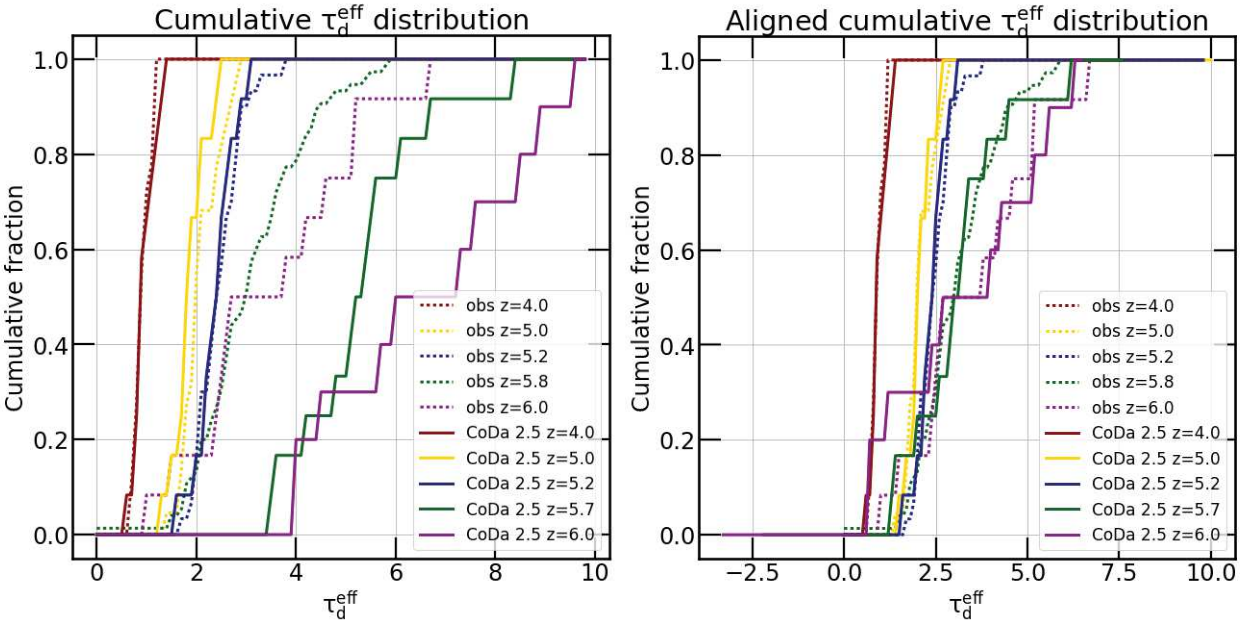


Figure 3.3: *Left*: Cumulative distribution of $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$ in CoDa 2.5, versus observations from Fan et al. [2006]; Becker et al. [2013, 2015a]; Bosman et al. [2018] *Right*: Same as the left but the simulated distributions are shifted to meet the observational curves at 50% .

convincingly on average, Reionization starts too late and is too abrupt, as our box is too neutral near z=6.

The right panel of Fig. 3.1 shows the evolution of the average neutral fraction in a late reionizing box, compared to that of CoDa II. The comparison between the two Reionization histories as seen through the average neutral fraction is striking : CoDa II's Reionization is over just before z=6, whereas in CoDa 2.5 overlap[2] appears to occur sometime around z≈5.5. Although both curves show exceedingly rapid decreases of xHI between 10^{-1} and 3×10^{-4}, the foot of the average xHI curve varies substantially. Not only does this foot lie at higher average xHI in the CoDa 2.5 (a few 10^{-4} versus almost 10^{-5} in CoDa II), but its slope is also much gentler. Most interestingly, CoDa 2.5 reionises too late and too strongly to match the constraints from Fan et al. [2006]. That such a box can match the average effective Lyα opacities so well, is further suggestive that the average xHI predictions from Fan et al. [2006] should be taken with a grain of salt. Finally, since CoDa 2.5 reionises so late and so rapidly, it is a poorer match to the locus of xHI constraints that are in the xHI>0.1,z>6 region of the plot. This further suggests that Reionization starts too late and is too sudden in CoDa 2.5.

The left panel of Fig. 3.3 shows the cumulative distribution of $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$ at various redshifts. The lower the redshift, the more complete Reionization is, and the more uniform the distribution of $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$. In turn, this means that the lower the redshift, the more abrupt the cumulative distribution of $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$ is. It also means that as redshift increases, the cumulative distributions shift to high values of $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$. At z=4.0, 5.0, and 5.2 the $\tau_{\mathrm{ly}\alpha}^{\mathrm{eff}}$s from CoDa 2.5 simulation are a very good match to data gathered from Fan et al. [2006]; Becker et al. [2013, 2015a]; Bosman et al. [2018], indicating that not only is the average

---

[2] Or when <xHI> ≈0.5

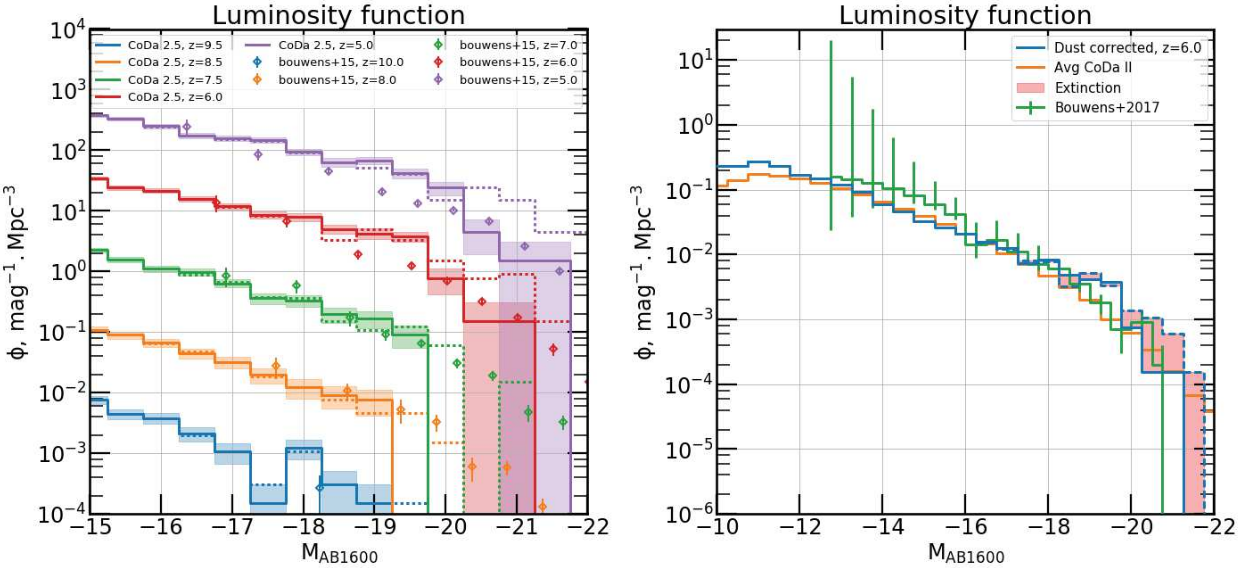


Figure 3.4: *Left*: UVLF of CoDa 2.5 simulation at z=9.5, 8.5, 7.5, 6.5, 6, 5 and the nearest (within dz=0.5) constraints from Bouwens et al. [2015]. Dotted lines indicate the UVLF without extinction, whereas full lines show the UVLF when considering extinction due to dust. The shaded regions show the poissonian error around the UVLF with extinction. For clarity, the data for each successive redshift is shifted upwards by 1 dex. *Right*: UVLF of the simulation when accounting for extinction (full blue line), and when no dust extinction is considered (dashed blue line) at z=6. Also plotted are the UVLFs from Bouwens et al. [2017] and Ocvirk et al. [2020] at z=6.

$\tau_{\rm ly\alpha}^{\rm eff}$ realistic at the end of Reionization and beyond, but that the distribution of neutral and ionised regions is realistic as well. At higher redshifts (z≥5.7), before the end of Reionization, the simulated distribution of $\tau_{\rm ly\alpha}^{\rm eff}$ is a very poor match to the observational data, and is up to two times too opaque. Nonetheless, we seek to compare the shape of the cumulative distributions of $\tau_{\rm ly\alpha}^{\rm eff}$ at these higher redshifts. To that end we shift the simulated distributions so that they intersect the distributions from observations at 50%. The result is shown in the right panel of Fig. 3.3, and demonstrates that although our simulation is on average too neutral at z≥5.7, the shape of the distribution of $\tau_{\rm ly\alpha}^{\rm eff}$ is a very good match to observations at z=5.7. At z=6, there is an excess of low $\tau_{\rm ly\alpha}^{\rm eff}$s in the simulated distribution when compared to observations. All of the above hints at CoDa 2.5 simulation containing a realistic distribution of ionising sources, both in relative luminosity and in space. However, it is also yet another sign that Reionization in CoDa 2.5 happens too suddenly, and is too neutral for z≥5.7.

Content that the new simulation is a much superior match to observational data on the Lyα forest, we now turn our attention to other typical constraints to assure ourselves that CoDa 2.5 simulation is a convincing simulation of Reionization.

3.2.3 UVLF

The left panel of Fig. 3.4 shows the UVLF of CoDa 2.5 at several redshifts. Note that for z>6, the data from Bouwens et al. [2017] are at least 0.5 in redshift away from the simulation snapshots we analysed, potentially reducing or improving the apparent quality of the match with observations. When compared with the UVLFs of Bouwens et al. [2015] at similar times, the CoDa 2.5 UVLFs with extinction are in relatively good agreement with the observational data. For the brightest bins, the UVLF with extinction is a slightly better match to the observations. This is an encouraging sign for our dust model, particularly because we do not attempt to calibrate our extinction post-processing in order to match the UVLF. There are a few regions where the agreement with observations is less satisfactory than elsewhere. For instance, around -19 at z=6 and z=5, the CoDa 2.5 UVLF is significantly higher than that reported by Bouwens et al. [2015], though this excess disappears at higher magnitudes, in part due to dust extinction. Interestingly, as redshift decreases and brighter galaxies make their appearance, the magnitude threshold brighter than which dust has a marked effect on the UVLF shifts to lower magnitudes. This means that extinction has a surprisingly important role even at early times. For instance, at z=8.5, the $\mathrm{M}_{\mathrm{AB1600}}$=-20 and -19.5 bins are completely suppressed by extinction, as is the -21 bin at z=7.5.

The right panel of Fig. 3.4 shows the UVLF at z=6 with and without extinction. Extinction has a marked effect on the UVLF at z=6, particularly from -20 and brighter. For instance, at magnitudes near -20 extinction reduces the UVLF by up to almost 1 dex. Overall, the CoDa 2.5 UVLF is a relatively good match to the observations from Bouwens et al. [2017] and the simulated UVLF of CoDa II. Of particular interest is the fact that the UVLF with extinction is a much better match to the two other data sets from $\mathrm{M}_{\mathrm{AB1600}}$ = -20 to -21. Near -19, the CoDa 2.5 simulation over predicts the UVLF when compared to both of the other data sets. Elsewhere, it undershoots the data from Bouwens et al. [2017] between -16 and -11, but is a very good match to CoDa II. When compared to results from Wu et al. [2020], the impact of dust in our simulation is similar. Their z=6 UVLF is diminished by extinction from -18 and onwards. At -20.5 and z=6, dust extinction reduces the UVLF by around 0.8 dex in CoDa 2.5, which is close to their result simulation. That being said, it is difficult to compare our extinction as they do not use the computed dust density field, but rather the metallicity along lines of sight to compute extinction via empirical scaling relations.

3.2.4 Further constraints on sources

The left panel of Fig. 3.5 shows the cosmic star formation rate density in CoDa 2.5. Both the fullbox SFRD and the SFRD of stars associated with haloes are in close agreement with the observations of Bouwens et al. [2014a]. However, the SFRD of stars associated with haloes brighter than -17 undershoots the shaded area between z=9 and z=7, showing that CoDa 2.5 could be under-predicting the amount of star formation during Reionization. Increasing star formation during this time frame could potentially improve our agreement with z>6 constraints on the neutral fraction of Hydrogen.

The right panel of Fig. 3.5 shows the average relation between stellar mass and halo mass in CoDa 2.5 at various redshifts. For all redshifts, the average stellar mass to halo mass relation is a power law of halo mass (roughly, $\mathrm{M}_\star \propto \mathrm{M}_\mathrm{h}^{1.6}$). Though the average does not change with redshift, the the 25^{th} and 75^{th} percentile curves show that scatter around the average does slightly : Below $10^9\,\mathrm{M}_\odot$ the 25^{th} percentile curve is lower at z=5 than at z=11, probably due to star formation suppression in low mass haloes. Above this mass the opposite seems true. All of our average curves are a reasonably good

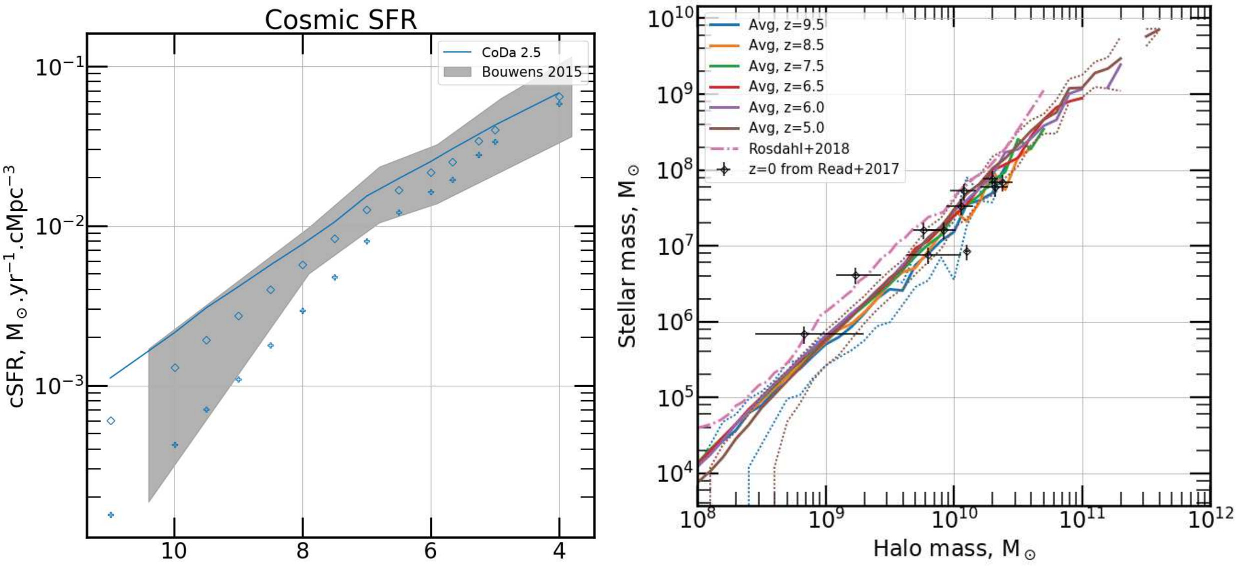


Figure 3.5: *Left*: Cosmic star formation rate density of CoDa 2.5 (blue), and (grey shaded area) from Bouwens et al. [2015]. The width fo the grey shaded area corresponds to the combined errorbars of both the dust corrected and the dust uncorrected results from Bouwens et al. [2015]. The full blue line shows the SFRD of the simulated box, the diamonds show the SFRD inside halo r_{200}s, and the crosses show the SFRD inside halo r_{200}s of haloes brighter than -17. *Right*: Stellar mass as a function of halo mass for various redshifts. The full lines show the CoDa 2.5 averages, whereas the dotted lines show the 25th and 75th percentiles for the distributions at z=5,11. Black crosses show z=0 constraints from Read et al. [2017]. The pink dot dashed line shows the z=6 average from the SPHINX simulation [Rosdahl et al., 2018].

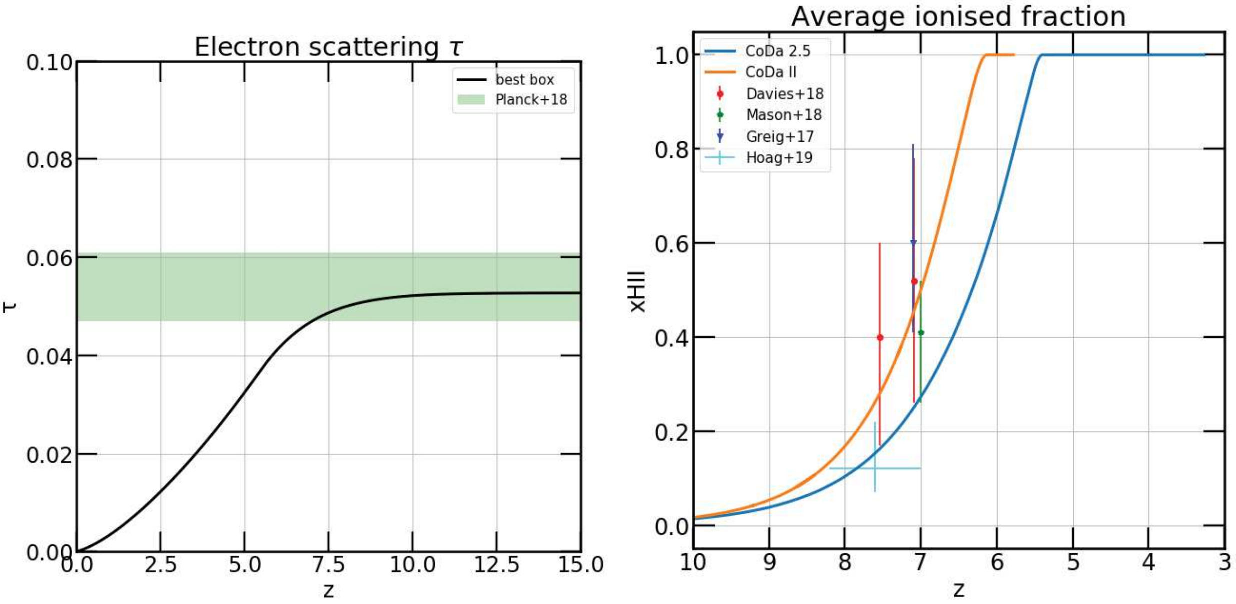


Figure 3.6: *Left*: Integrated Thomson optical depth from electronic scattering as a function of redshift in CoDa 2.5 simulation, compared to constraints from Planck Collaboration et al. [2018]. *Right*: Average ionised fraction of Hydrogen gas in CoDa 2.5 and CoDa II versus several observational constraints from Hoag et al. [2019]; Greig and Mesinger [2017]; Mason et al. [2018]; Davies et al. [2018]

fit to the z=0 observations of local dwarf galaxies from Read et al. [2017]. The SPHINX simulation [Rosdahl et al., 2018] also showed a good match to the same observations. Their predicted Stellar masses are similar to ours, albeit higher by ≈0.1-0.2 dex. This difference could stem from both their higher resolution, and their different recipes for star formation and supernova feedback.

3.2.5 Further constraints on ionisation

The left panel of Fig. 3.6 shows the integrated optical depth due to scattering of the CMB photons by free electrons in the IGM in CoDa 2.5, as well as constraints from Planck Collaboration et al. [2018]. In particular, it serves as a test of the timing of Reionization in the simulation. Too late and τ would be too small, too early and it would be too big. The agreement between our predictions and the Planck data lends further credibility to the realism of the reionising universe as depicted by CoDa 2.5. The right panel of Fig. 3.6 shows the volume averaged ionised fraction as a function of redshift for both the CoDa II and the CoDa 2.5 simulations. In both cases, the average fraction of ionised gas rises from z=10 onwards, reaching close to 1 at z=6.2 (z=5.5) in CoDa II (CoDa 2.5) when the ionised bubbles percolate at overlap. In fact, the increase in average xHII accelerates as time goes by. The difference between the two curves is yet another illustration of the difference in the timing of Reionization between CoDa II and CoDa 2.5. The later Reionization of the latter simulation, means that it is a better fit to the constraint on average xHII from Hoag et al. [2019] (who favour late Reionization in their interpretation), but a poorer match to the constraints from Greig and Mesinger [2017]; Mason et al. [2018]; Davies et al. [2018]. Again, this suggests that although the timing of the end

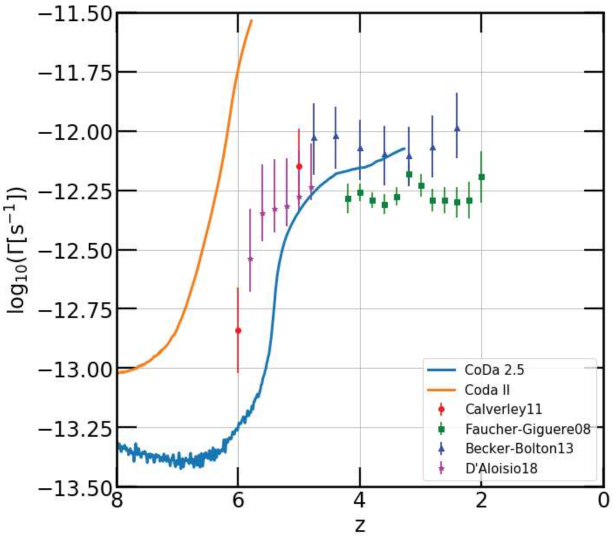


Figure 3.7: Average photo-ionisation rate (Γ) as a function of redshift in the CoDa 2.5 simulation and in CoDa II. Also plotted are various constraints from the literature (observations : Calverley et al. [2011]; Becker and Bolton [2013] and simulations : Faucher-Giguère et al. [2008]; D'Aloisio et al. [2018]).

of Reionization may be realistic in the CoDa 2.5 simulation, the number of ionising photons released to the IGM between z=10 and z=6 could be insufficient (and therefore, the average xHII too low) to properly match observational constraints.

Fig. 3.7 shows the CoDa 2.5 average photo-ionisation rate as a function of redshift. The rate of ionisation increases rapidly from z=6 till z=5, after which the increase is steadier (trends which are echoed in the average neutral fraction in Fig. 3.1). On the whole, the CoDa 2.5 average Γ is low when compared to the displayed data points. After z=4, the agreement with the observationally derived constraints of Becker and Bolton [2013] is quite good. Earlier, between z=4 and z=5 the average Γ is just under the data from Becker and Bolton [2013] and Calverley et al. [2011]. The agreement with observations deteriorates measurably between z=5 and z=6, where $<\Gamma>$ falls off far too rapidly to match the z=6 constraints. In fact, at z=6 the CoDa 2.5 photo-ionisation rate is over 1.5 dex lower than the closest constraint, again demonstrating that we lack enough ionising photons before z=5.7. That being said, the match between CoDa 2.5 and observations is far better than between CoDa II and said observations. Indeed, at all times, the CoDa II $<\Gamma>$ is more than one dex above the CoDa 2.5 result. The CoDa II curve also lacks the knee like feature of the CoDa 2.5 average Γ.

All in all, we have shown that the new set-up and calibration is an improved match with observational constraints when compared to CoDa II. However, there remains room for improvement, especially during Reionization when there seems to be too few ionisations, and the IGM seems to transparent. Ideally we would need to converse the good match of CoDa 2.(with observations for z<5.5 whilst slightly increasing the number of ionisations that take place for z>5.5.

3.3 Résumé des résultats

Le troisième chapitre est consacré à l'étude d'éventuels changements dans la paramétrisation des simulations pour améliorer l'accord avec les contraintes observationnelles de la Réionisation. Principalement, je discute des résultats récents [Kulkarni et al., 2019, en particulier], qui rappellent que certaines contraintes sur l'ionisation moyenne du mileu inter-galactique (fondées sur des mesures de la forêt Lyα de quasars [Fan et al., 2006]) dépendent d'hypothèses importantes. Ils mettent en avant la production d'observation Lyα factices comme une meilleure contrainte de l'ionisation du gaz d'hydrogène dans le milieu intergalactique. En ce faisant, ils obtiennent des simulations favorisant une Réionisation plus tardive ($z \approx 5.5$) que dans CoDa II, où le milieu inter-galactique est trop ionisé à la fin de la Réionisation.

Pour réduire cet excès de photons à la fin de la Réionisation, une approche possible est de chercher à amplifier la suppression de formation stellaire qui a lieu de plus en plus vers la fin de la Réionisation. Inclure un critère de temperature dans la recette sous-grille de formation stellaire accentue la suppression de formation stellaire à la fin de la Réionisation, quand le gaz a été chauffé par des photons ionisants et par des explosions de supernovas. J'ai montré qu'en ayant recours à un tel critère (plus précisément de 3×10^4 K), l'accord des simulations avec les observations de la forêt Lyα était grandement amélioré, et la Réionisation terminait plus tard (vers $z \approx 5.7$) En particulier, j'ai trouvé que quand $4<z<5.2$, l'opacité effective moyenne Lyα et la distribution des opacités Lyα était en bon accord avec les observations de Bosman et al. [2018]; Becker et al. [2015b]; Fan et al. [2006]. Pour des redshifts plus élevés, l'accord est moins bon : les opacités effective moyennes Lyα sont trop grandes, traduisant un milieu inter-galactique trop neutre. Cependant, un ajustement des courbes montre que bien que la moyenne est trop élevée dans les simulations, les distributions sont très similaires même pour $z > 5.2$, suggérant que la géométrie des régions ionisées est réaliste.

Par la suite, j'explore les effets combinés des nouvelles physiques et du critère de température sur l'accord avec les autres contraintes typiquement utilisées. Notamment, je montre que la fonction de luminosité ultra-violette des galaxies est réaliste, étant comparable à la fois à celle de CoDa II et aux observations de Bouwens et al. [2015]. Conjointement, j'ai mis en evidence l'extinction importante qui se produit dès $z = 8$ pour les magnitudes plus lumineuses que -19. J'ai aussi montré que l'accord avec les mesures du taux de formation stellaire cosmique de Bouwens et al. [2014b] était bon. En même temps, le rapport entre masses en matière noire et masses stellaires des galaxies est semblable aux contraintes à bas redshift de Read et al. [2017], et aux résultats de la simulation SPHINX [Rosdahl et al., 2018]. De plus, l'épaisseur optique Thomson rencontrée par les photons du CMB à cause des électrons libres dans la simulation est compatible avec les observations de Planck [?]. J'ai aussi étudié l'accord entre la nouvelle simulation et les contraintes sur le taux d'ionisation gaz comparé à celui de CoDa II : en dépit de la Réionisation plus tardive dans la nouvelle simulation, l'accord avec l'ensemble des contraintes est similaire. Finalement, j'ai montré que la nouvelle simulation est beaucoup plus compatible avec les contraintes sur le taux d'ionisation de l'Hydrogène que CoDa II.

BIBLIOGRAPHY

Becker G.D. and Bolton J.S. New measurements of the ionizing ultraviolet background over $2 < z < 5$ and implications for hydrogen reionization. *Monthly Notices of the Royal Astronomical Society*, 436:1023–1039, 2013.

Becker G.D., et al.. A refined measurement of the mean transmitted flux in the Lyα forest over $2 < z < 5$ using composite quasar spectra. *Mon Not R Astron Soc*, 430(3):2067–2081, 2013.

Becker G.D., et al.. Evidence of patchy hydrogen reionization from an extreme Lyα trough below redshift six. *Monthly Notices of the Royal Astronomical Society*, 447:3402–3419, 2015a.

Becker G.D., et al.. Reionisation and High-Redshift Galaxies: The View from Quasar Absorption Lines. *Publications of the Astronomical Society of Australia*, 32:e045, 2015b.

Becker G.D., et al.. Evidence for Large-scale Fluctuations in the Metagalactic Ionizing Background Near Redshift Six. *The Astrophysical Journal*, 863:92, 2018.

Bosman S.E.I., et al.. New constraints on Lyman-{\alpha} opacity with a sample of 62 quasars at $z > 5.7$. *arXiv:1802.08177 [astro-ph]*, 2018.

Bouwens R.J., et al.. A Census of Star-forming Galaxies in the Z ~ 9-10 Universe based on HST+Spitzer Observations over 19 Clash Clusters: Three Candidate Z ~ 9-10 Galaxies and Improved Constraints on the Star Formation Rate Density at Z ~ 9.2. *The Astrophysical Journal*, 795(2):126, 2014a.

Bouwens R.J., et al.. UV-continuum Slopes of >4000 z ~ 4-8 Galaxies from the HUDF/XDF, HUDF09, ERS, CANDELS-South, and CANDELS-North Fields. *The Astrophysical Journal*, 793:115, 2014b.

Bouwens R.J., et al.. UV Luminosity Functions at Redshifts $z \sim 4$ to $z \sim 10$: 10,000 Galaxies from HST Legacy Fields. *ApJ*, 803(1):34, 2015.

Bouwens R.J., et al.. The $z \sim 6$ Luminosity Function Fainter than -15 mag from the Hubble Frontier Fields: The Impact of Magnification Uncertainties. *The Astrophysical Journal*, 843:129, 2017.

Calverley A.P., et al.. Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect. *Monthly Notices of the Royal Astronomical Society*, 412:2543–2562, 2011.

Chardin J., et al.. Calibrating cosmological radiative transfer simulations with Ly α forest data: evidence for large spatial UV background fluctuations at $z \sim 5.6$–5.8 due to rare bright sources. *Mon Not R Astron Soc*, 453(3):2943–2964, 2015.

Chardin J., et al.. Large-scale opacity fluctuations in the Lyα forest: evidence for QSOs dominating the ionizing UV background at $z \sim 5.5$-6? *Monthly Notices of the Royal Astronomical Society*, 465:3429–3445, 2017.

D'Aloisio A., et al.. Large fluctuations in the high-redshift metagalactic ionizing background. *Monthly Notices of the Royal Astronomical Society*, 473:560–575, 2018.

Davies F.B., et al.. Quantitative Constraints on the Reionization History from the IGM Damping Wing Signature in Two Quasars at z $>$ 7. *The Astrophysical Journal*, 864:142, 2018.

Fan X., et al.. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132:117–136, 2006.

Faucher-Giguère C.A., et al.. Evolution of the Intergalactic Opacity: Implications for the Ionizing Background, Cosmic Star Formation, and Quasar Activity. *The Astrophysical Journal*, 688:85–107, 2008.

Greig B. and Mesinger A. The global history of reionization. *Monthly Notices of the Royal Astronomical Society*, 465:4838–4852, 2017.

Hoag A., et al.. Constraining the Neutral Fraction of Hydrogen in the IGM at Redshift 7.5. *The Astrophysical Journal*, 878:12, 2019.

Keating L.C., et al.. Long troughs in the Lyman-α forest below redshift 6 due to islands of neutral hydrogen. *arXiv e-prints*, page arXiv:1905.12640, 2019.

Kulkarni G., et al.. Large Ly α opacity fluctuations and low CMB τ in models of late reionization with large islands of neutral hydrogen extending to $z < 5.5$. *Monthly Notices of the Royal Astronomical Society: Letters*, 485(1):L24–L28, 2019.

Mason C.A., et al.. The Universe Is Reionizing at z $\approx$ 7: Bayesian Inference of the IGM Neutral Fraction Using Lyα Emission from Galaxies. *The Astrophysical Journal*, 856(1):2, 2018.

Ocvirk P., et al.. The Reionization of Galactic Satellite Populations. *The Astrophysical Journal*, 794:20, 2014.

Ocvirk P., et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. *Monthly Notices of the Royal Astronomical Society*, 2020.

Planck Collaboration, et al.. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, 1807:arXiv:1807.06209, 2018.

Read J.I., et al.. The stellar mass-halo mass relation of isolated field dwarfs: a critical test of ΛCDM at the edge of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 467(2):2019, 2017.

Rosdahl J., et al.. The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Monthly Notices of the Royal Astronomical Society*, 479:994–1016, 2018.

Theuns T., et al.. P^3M-SPH simulations of the Lyalpha forest. *Monthly Notices of the Royal Astronomical Society*, 301:478–502, 1998.

Trebitsch M., et al.. The Obelisk simulation: galaxies contribute more than AGN to HI reionization of protoclusters. *arXiv e-prints*, 2002:arXiv:2002.04045, 2020.

Wu X., et al.. Photometric properties of reionization-epoch galaxies in the Simba simulations. *Monthly Notices of the Royal Astronomical Society*, page staa1044, 2020. ArXiv: 1911.06330.

Ionising photon budget of dusty galaxies with late Reionization

4

In Chp. 2, we showed that RAMSES-CUDATON can produce realistic dust distributed throughout reionising galaxies, thanks to a physical model (designed and partly implemented by Yohan Dubois (Dubois 2020, in prep.) that I coupled to RAMSES-CUDATON. The dust densities computed are used on the fly to absorb LyC photons as they propagate through the densest gas where dust forms preferentially. At the same time I also introduced additional cooling produced by metals and Helium, which I have shown can have a significant impact on star formation of galaxies, as well as on the history of Reionization. In parallel, in Chp. 3, I outlined a new calibration of our simulation code, that produces late Reionization scenarios, and is a much better match to observational constrains from the Lyα forest. Immediately, one is curious as to the effects of this new setup on the conclusions we drew in Chp. 1 from my analysis of the ionising photon budget of galaxies in CoDa II [see Sec. Galaxy sample and Lewis et al., 2020]. To seek to answer such questions we will split our analysis into two parts: Understanding the new $f_{\rm esc}^{\rm ray}$ results that combine both neutral gas and dust opacities, and understanding the new photon budget.

4.1 Escape fraction

4.1.1 Escape fraction versus mass

Fig. 4.1 shows the distribution of haloes in (halo mass,$f_{\rm esc}^{\rm ray}$) space, in CoDa 2.5, at four epochs. Globally the maps and curves resemble the CoDa II results, with high $f_{\rm esc}^{\rm ray}$ values at low masses, that decrease with mass and increase with decreasing redshift.

In the top left plot, the density map of $f_{\rm esc}^{\rm ray}$ as a function of mass at z=11, is very similar to the equivalent map at z=10 in CoDa II. Most of the haloes are concentrated below $\approx 10^9\,{\rm M}_\odot$, and between $f_{\rm esc}^{\rm ray} = 0.1$ and $f_{\rm esc}^{\rm ray} = 1.0$, with some low mass haloes having $f_{\rm esc}^{\rm ray}$ as low as 3×10^{-5}. At this early time, less than 10% of haloes have $f_{\rm esc}^{\rm ray} \approx 1.0$. The average halo $f_{\rm esc}^{\rm ray}$ decreases slightly by from ≈ 0.6 to near 0.3 from $10^8\,{\rm M}_\odot$ to $10^9\,{\rm M}_\odot$, and then more dramatically by around 2 dex between $10^8\,{\rm M}_\odot$ and a few $10^{10}\,{\rm M}_\odot$. The scatter around this descending average is quite high. For instance, at $\approx 4\times10^9$, $f_{\rm esc}^{\rm ray}$ varies by around two dex between around 0.6 and 4×10^{-4}. A fraction of the low mass haloes have very low $f_{\rm esc}^{\rm ray}$ values, that almost extend down to 10^{-5} in the map. We showed that in CoDa II and Fig. 1.18, haloes that occupied this space had just formed a stellar particle, and had lower $f_{\rm esc}^{\rm ray}$ values because there had not been enough time for them to ionise their gas more.

At z=8.5, there are more star forming haloes, and more massive ones as well. Indeed, the map now extends to almost $10^{11}\,{\rm M}_\odot$, and continuing the trend seen at z=11, down to lower $f_{\rm esc}^{\rm ray}$ values of just under 10^{-3}. The fraction of haloes in the highest $f_{\rm esc}^{\rm ray}$ bin increases between z=11 and z=8.5. Though the trend of the average $f_{\rm esc}^{\rm ray}$ curve is preserved, it lies higher by about one bin or 0.1 dex, reaching near 0.8 at $10^8\,{\rm M}_\odot$. Due to the higher number of star forming haloes, the average curve is not as noisy

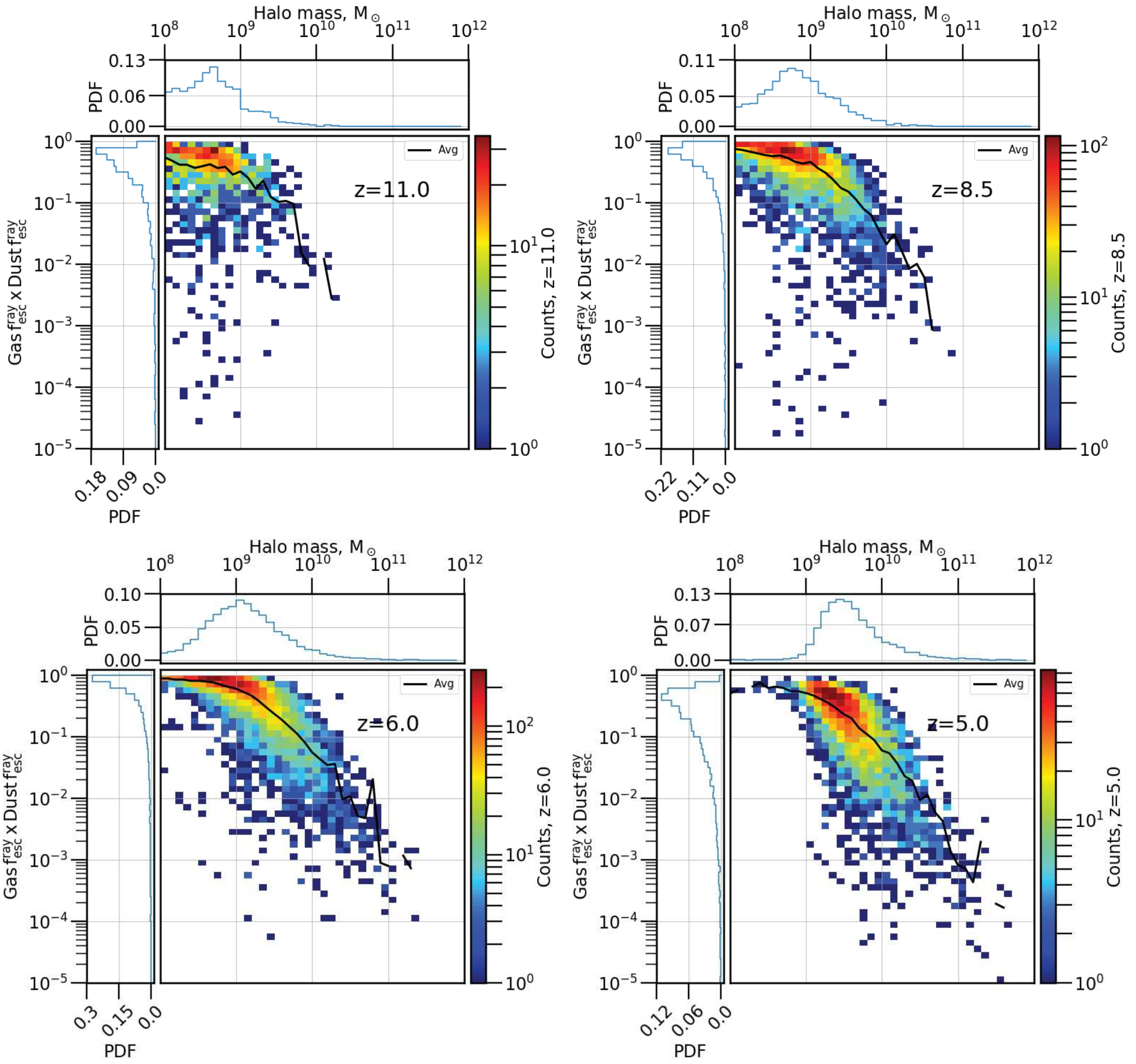


Figure 4.1: Density maps of $f_{esc,total}^{ray}$(accounting for both gas and dust opacities) versus halo mass, for star forming (With a least one stellar particle older than 10 Myr) haloes in CoDa 2.5. The average f_{esc}^{ray} of star forming haloes are in full black lines. Four redshifts are shown; clockwise, and starting from the top left they are: z=11, 8.5, 6, 5.

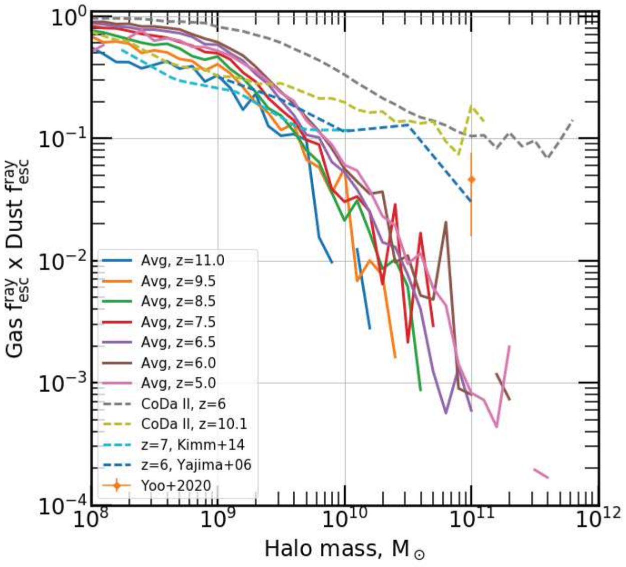


Figure 4.2: Average galactic escape fraction through resolved simulation gas and dust in the CoDa 2.5 simulation, for haloes with SFR>0. Also shown are the equivalent averages from CoDa II at z=6.0 and z=10.1, as well as simulation results from Kimm and Cen [2014]; Yajima et al. [2011]; Yoo et al. [2020].

as it was at z=11. The haloes with very low f_{esc}^{ray} that were below $\approx 3 \times 10^8\,M_\odot$ vanish between z=11 and z=8.5.

Between z=8.5 and z=6, the previously described redshift evolution continues: The counts of galaxies increase for all masses and f_{esc}^{ray}s, with masses higher than $10^{11}\,M_\odot$ appearing and continuing the downward trend of $f_{esc}^{ray}$with halo mass. By z=6, average curves reach all the way down to $f_{esc}^{ray}=8 \times 10^{-4}$ at $10^{11}\,M_\odot$. The highest f_{esc}^{ray} bin is now the most populated one, containing just under 30% of all haloes. The average curve has risen to ≈ 0.9 at $10^8\,M_\odot$, increasing by about close to 0.1 dex for most masses.

One striking difference between the z=5 map and the former maps, is the almost complete disappearance of the plateau feature of f_{esc}^{ray} under $8 \times 10^8\,M_\odot$ due to star formation suppression. Above that mass, f_{esc}^{ray} values form a downwards stretching plume. The vanishing of high f_{esc}^{ray} values due to suppression is dramatic: the highest f_{esc}^{ray} bin is now almost empty, and the second highest f_{esc}^{ray} bin has less that 6% of haloes. In Fig. 2.20 we also saw a strong imprint of star formation suppression in CoDa 2.5 between z=6 and z=5. In CoDa II, star formation was much less dramatic as shown in Fig. 1.16 and in Fig. 1 of Lewis et al. [2020]: at z=6 in CoDa II, when Reionization is complete, there remains a large population of star forming haloes producing ionising light and with high f_{esc} values. Although the average values for the highest mass haloes are similar, the scatter of f_{esc}^{ray} for the highest masses now extends further, to 10^{-5}.

4.1.2 Average escape fraction versus mass

Fig. 4.2 shows the averaged total (accounting for HI and dust absorption) $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ of star forming haloes at various redshifts. The curves at z=11, 8.5, 6, 5 correspond to the full curves of the four corresponding panels of Fig. 4.1. At all times, the averages take the shape of a high $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ plateau at low masses, that fades into a descending power law with mass above $\approx 10^9\,\mathrm{M}_\odot$. Although this is qualitatively similar to the CoDa II averages (and the plotted results from Yajima et al. [2011]; Kimm and Cen [2014] above this mass), the CoDa 2.5 average slope is steeper than in CoDa II, giving $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ averages that are substantially lower than in CoDa II: at z=6 the lowest average $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ value is $\approx 10^{-3}$ near $2 \times 10^{11}\,\mathrm{M}_\odot$, versus 0.1 at the same mass and redshift in CoDa II, and ≈ 0.03 in Yajima et al. [2011]; Yoo et al. [2020]. This is both because of the inclusion of dust absorption, but also due to the new cooling and star formation temperature criterion in CoDa 2.5. From z=11 to z=6 the average increases steadily for masses $< 10^9\,\mathrm{M}_\odot$ by around 0.5. The low mass CoDa 2.5 average is quite a good match to the CoDa II equivalent at z=10.1. Though the CoDa II averages at low mass also increase between z=10.1 and z=6, they do not change as much and reach higher values of practically 1. This difference in evolution could be linked to the different timing of Reionization in the boxes. That the CoDa 2.5 $\mathrm{f}_{\mathrm{esc}}$ at low masses is lower on average at later times than in CoDa II, could explain the much better agreement with Γ and the Lyα Forest constraints after z=6. The figure shows that the slopes of the power law with mass are almost identical between redshifts, whereas in CoDa II, the behaviour with mass of the average $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ changes substantially between z=10.1 and z=6. The changes at higher masses are less clear cut, partly due to the low numbers of massive haloes at the highest redshifts. Nevertheless, there does appear to be a similar shift towards higher $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$ between z=11 and z=6. From z=6 to z=5, there is no significant change in the averages. At the highest masses ($\geq 2 \times 10^{10}\,\mathrm{M}_\odot$) the noise in the curves tends to mask any slight increase, however the curves do not appear to join near $10^{11}\,\mathrm{M}_\odot$ as they do in CoDa II.

4.1.3 Total average $\mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}}$

An oft reported quantity, especially in models of Reionization is the star formation weighted average escape fraction (or $< \mathrm{f}_{\mathrm{esc}} >_{\mathrm{SFR}_{10}}$) Fig. 4.3 shows the $< \mathrm{f}_{\mathrm{esc,total}}^{\mathrm{ray}} >_{\mathrm{SFR}}$, $< \mathrm{f}_{\mathrm{esc,gas}}^{\mathrm{ray}} >_{\mathrm{SFR}}$, and $< \mathrm{f}_{\mathrm{esc,dust}}^{\mathrm{ray}} >_{\mathrm{SFR}}$ in CoDa 2.5. This quantity can be useful guidance for groups running SAMs, willing to assume a single unique number for $\mathrm{f}_{\mathrm{esc}}$ throughout their model. This suggests the average global fraction of produced photons are lost inside galactic haloes (though with a strong bias towards star forming haloes). The CoDa 2.5 $< \mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}} >_{\mathrm{SFR}_{10}}$ decreases from about 0.2 at z=11 to under 2×10^{-2} at z=5, mainly driven by the evolution of $\mathrm{f}_{\mathrm{esc,gas}}^{\mathrm{ray}}$ which is almost identical. The small gap that appears between the $< \mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}} >_{\mathrm{SFR}_{10}}$ and the $< \mathrm{f}_{\mathrm{esc,gas}}^{\mathrm{ray}} >_{\mathrm{SFR}}$ curves between z=9.5 and z=5 is due to dust. Indeed, $< \mathrm{f}_{\mathrm{esc,dust}}^{\mathrm{ray}} >_{\mathrm{SFR}}$ decreases much more gently from just over 0.9 at z=11 to 0.4 at z=5. Based on Fig. 4.2, that showed that the average relation between $\mathrm{f}_{\mathrm{esc,dust}}^{\mathrm{ray}}$, $\mathrm{f}_{\mathrm{esc,gas}}^{\mathrm{ray}}$ and halo mass actually increases over time, the decrease seen in the SFR$_{10}$ weighted average for the different $< \mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}} >_{\mathrm{SFR}_{10}}$ most probably comes from the build up of more massive haloes with lower $\mathrm{f}_{\mathrm{esc,dust}}^{\mathrm{ray}}$ and lower $\mathrm{f}_{\mathrm{esc,gas}}^{\mathrm{ray}}$, as well as increased suppression (although that mostly affects the low mass haloes). The decrease of $< \mathrm{f}_{\mathrm{esc}}^{\mathrm{ray}} >_{\mathrm{SFR}_{10}}$ in CoDa 2.5 with time is in stark contrast to the equivalent curves from CoDa II,

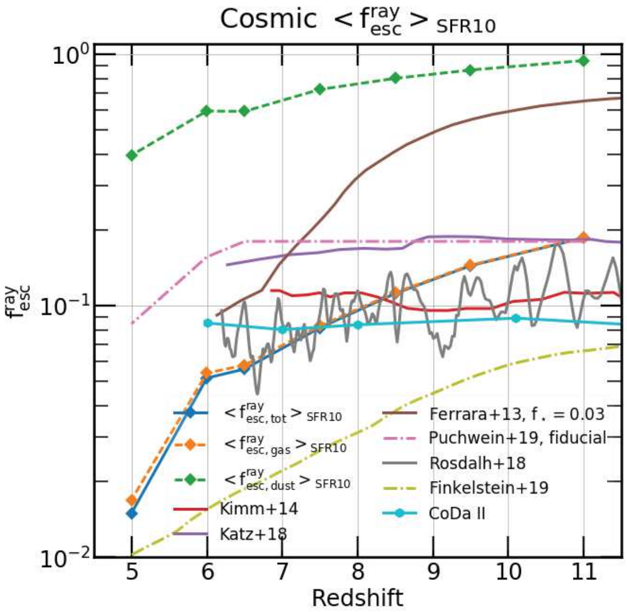


Figure 4.3: SFR_{10} weighted mean of $\mathrm{f^{ray}_{esc}}$, $\mathrm{f^{ray}_{esc,gas}}$, and $\mathrm{f^{ray}_{esc,dust}}$ in CoDa 2.5, CoDa II as well as from the simulations of Kimm and Cen [2014]; Katz et al. [2018]; Rosdahl et al. [2018]; Kulkarni et al. [2019]; Keating et al. [2019], the models of Puchwein et al. [2019]; Madau [2017]; Finkelstein et al. [2019] (in dot dashed curves), and the SAM of Ferrara and Loeb [2013].

and from Kimm and Cen [2014]; Katz et al. [2018]; Rosdahl et al. [2018]. The latter 4 average curves fluctuate but do not evolve at the redshifts that we study, and stay between ≈ 0.2 and 0.09. Interestingly enough, the CoDa 2.5 SFR_{10} weighted average intersects the curves from simulations of Katz et al. [2018] and the model of Puchwein et al. [2019] at z=11. The trend with redshift reported by Ferrara and Loeb [2013] shows a strong decreases over time, which is similar to the CoDa 2.5 case. However, their average escape fraction is higher than the CoDa 2.5 equivalent at all times. The model of Puchwein et al. [2019] is flat until z$\approx$6.5. After z=6.5, their model shows a sudden downturn in $<\mathrm{f^{ray}_{esc}}>_{\mathrm{SFR}_{10}}$ which has a similar slope to that of CoDa 2.5 at the same time. The predictions of Finkelstein et al. [2019] have a very similar trend with redshift to our own, but are consistently lower. In fact, they found that a downturn in $<\mathrm{f^{ray}_{esc}}>_{\mathrm{SFR}_{10}}$ near the end of Reionization is necessary to reproduce the constraints on the total ionising emissivity (or total photon budget). This is similar to the findings of Madau [2017]; Finkelstein et al. [2019], who require a decrease in $<\mathrm{f_{esc}}>_{\mathrm{SFR}_{10}}\times\xi_{\mathrm{ion}}$ to match the changing neutrality of the IGM in observations. At the same time, Kulkarni et al. [2019]; Keating et al. [2019] found that they required a decrease in the total photon budget of their simulations with redshift, in order to reproduce observations of the Lyα forest opacities.

To better comprehend the observed differences, we shall study the gas and dust components of $\mathrm{f^{ray}_{esc}}$ ($\mathrm{f^{ray}_{esc,gas}}$ and $\mathrm{f^{ray}_{esc,dust}}$) individually.

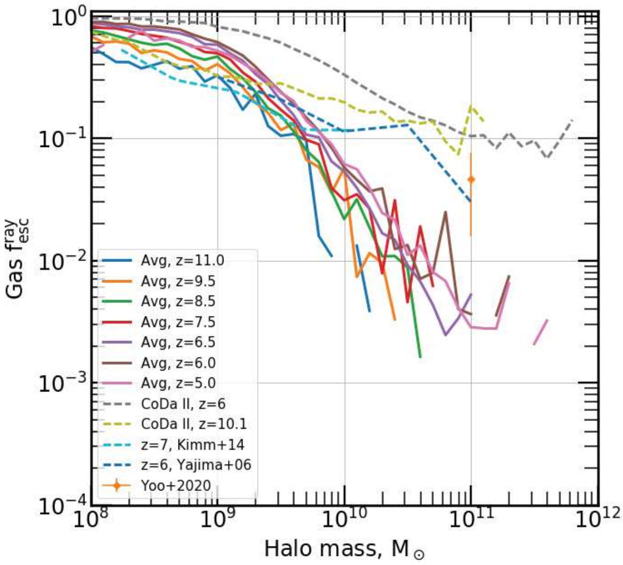


Figure 4.4: Average galactic escape fraction through resolved simulation gas of star forming haloes, in CoDa 2.5. Also shown are the equivalent averages from CoDa II at z=6.0 and z=10.1, as well as simulation results from Kimm and Cen [2014]; Yajima et al. [2011]; Yoo et al. [2020].

4.2 Gas f_{esc}^{ray}

4.2.1 Average gas f_{esc}^{ray} versus mass

Fig. 4.4 shows the average $f_{esc,gas}^{ray}$ of SFR>0 haloes in CoDa 2.5, at several redshifts. When accounting only for HI absorption in galactic gas, the average results are quasi-identical to those that account for both HI and dust absorption between 10^8 and $\approx 2 \times 10^{10}\,M_\odot$. For greater masses, differences emerge between the average $f_{esc,gas}^{ray}$ and the total f_{esc}^{ray} of Fig. 4.2. From that mass and onwards, the $f_{esc,gas}^{ray}$ averages are higher at all times. For instance, where the lowest average values of f_{esc}^{ray} are around 10^{-3} at z=6, the averages of $f_{esc,gas}^{ray}$ are never lower than 2×10^{-3}. This is the imprint of dust absorption in CoDa 2.5 galaxies, that is clearly smaller than absorption from HI at all times and masses. That being said, the average $f_{esc,gas}^{ray}$ are still noticeably lower than the average f_{esc}^{ray} of CoDa II or of Kimm and Cen [2014]; Yajima et al. [2011]; Yoo et al. [2020] that are more than one dex greater near $10^{11}\,M_\odot$, meaning that our new setup has changed the relation between f_{esc}^{ray} and halo mass. We investigate the probable cause of this in the next subsection.

4.2.2 Effect of $T_\star$ on gas f_{esc}^{ray}

Our new setup has significantly altered the behaviour of $f_{esc,gas}^{ray}$ (and f_{esc}^{ray} by extension) as a function of halo mass. One modification that might have obvious ramifications on $f_{esc,gas}^{ray}$, is the addition of a temperature criterion for star formation. Indeed, in Lewis et al. [2020], we showed that haloes' escape fractions were set by the relative amount of star formation between neutral opaque regions, and hot ionised transparent regions. Therefore, by disallowing the formation of stellar particles in the hot ionised transparent regions, it is natural to expect lower escape fractions. To test whether the

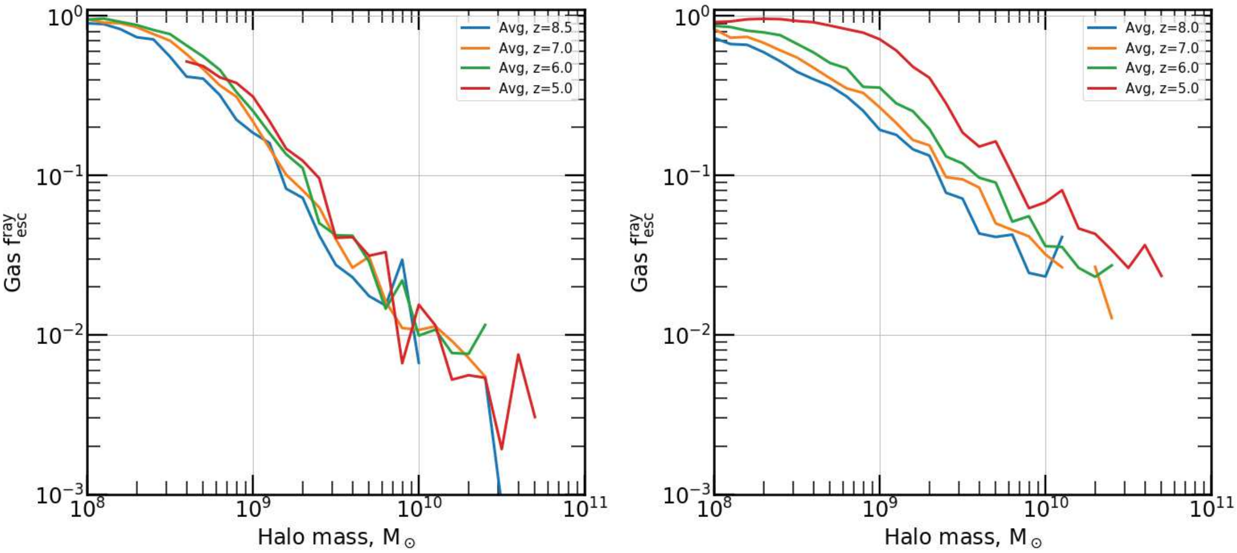


Figure 4.5: Average galactic escape fraction through resolved simulation gas, for haloes with stars, in two separate 1024 cell, $8\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$ simulations, that used the same simulation code. THey are identical except for the rules for star formation. The left panel corresponds to a simulation with a star formation criterion of $T_\star = 2\times10^4\,\mathrm{K}$, whereas the right panel has no such temperature criterion.

introduction of the star formation criterion is the determining factor in the change in the behaviour of the average escape fractions between CoDa 2.5 and CoDa II, I computed the f_{esc}^{ray} in two existing 1024 cell, $8\,\mathrm{cMpc}\cdot\mathrm{h}^{-1}$ boxes (ran with the same code) with and without a temperature criterion for star formation. Fig. 4.5 shows the resulting average $f_{esc,gas}^{ray}$: both with a star formation criterion (left panel) and without such a criterion (right). Although the resolution and exact setup are different than in the CoDa 2.5 simulation, the behaviour of the average $f_{esc,gas}^{ray}$ in the left is similar to that of the CoDa 2.5 and is close to a power law with a low mass plateau, albeit with a faster transition between the two and a steeper power law. In the right panel the average curves have a much gentler slope than in the left, leading to higher average $f_{esc,gas}^{ray}$ values for $M_{halo} \geq 5\times10^8\,M_\odot$ at z=8,7,6. At z=5, the difference is particularly striking. Whereas in the left panel, the average curves are similar for all redshifts, the right panel curves do not overlap as much. In particular, there is a considerable step between the z=6 and the z=5 curve, for which the transition mass between the the plateau and the power law happens around 0.5 dex later in mass, yielding much higher $f_{esc,gas}^{ray}$ than at earlier redshifts. In CoDa II, for which there was no temperature criterion for star formation, we observed a similar change in the mass threshold between a plateau like feature and a power law after z=7. The above observations suggest that the introduction of a temperature criterion for star formation strongly impacts the average $f_{esc,gas}^{ray}$ of haloes. Not only does it alter the slope of the average $f_{esc,gas}^{ray}$ curves with mass, but it also reduces their evolution with redshift, explaining the major differences when comparing the average behaviour of f_{esc} in CoDa II and CoDa 2.5.

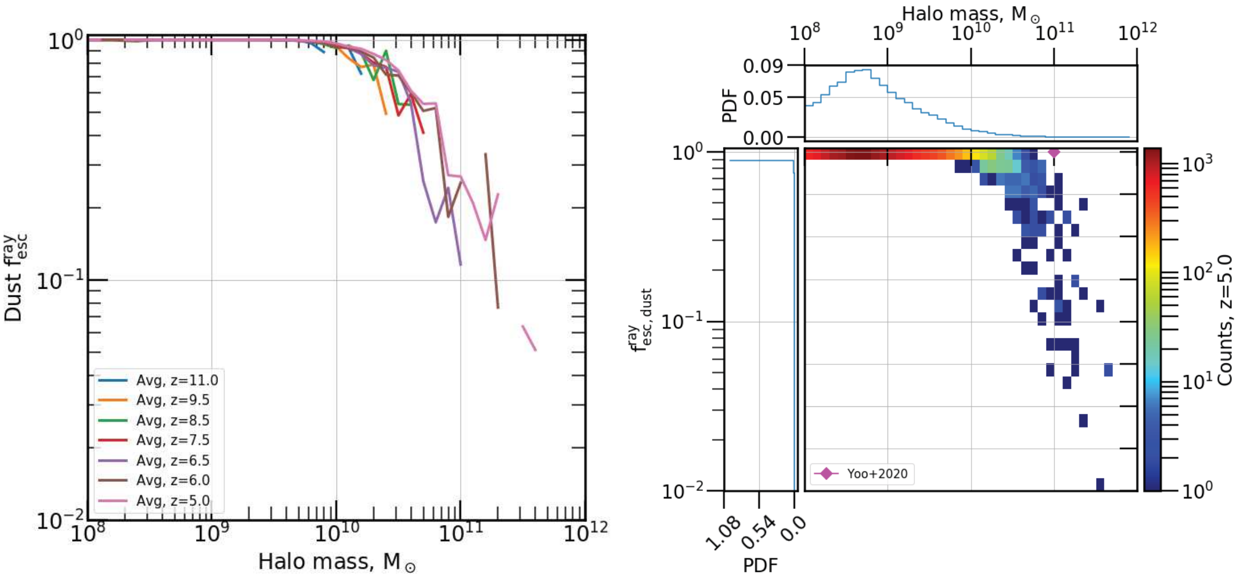


Figure 4.6: *Left*: Galactic escape fraction through resolved simulation dust in the CoDa 2.5, as a function of halo mass scaled for various epochs. *Right*: Density map of halo galactic escape fraction through resolved simulation dust in the CoDa 2.5 as a function of mass, at z=5.0, for haloes with stars.

4.3 Dust f_{esc}^{ray}

4.3.1 Average dust f_{esc}^{ray} versus mass

The left panel of Fig. 4.6 shows the average $f_{esc,dust}^{ray}$ in star forming haloes as a function of halo mass, in CoDa 2.5 at several redshifts (full curves). At all times the average $f_{esc,dust}^{ray}$ due to dust absorption within galaxies is flat from $10^8\,M_\odot$ to $7\times10^9\,M_\odot$. Above these masses, the averages behave as power laws of halo mass, with an exponent of $\approx$ -0.8. This leads to values of less than 0.1 on average for haloes more massive than $10^{11}\,M_\odot$ at z=5, signifying that massive haloes have low dust escape fractions ($\gtrsim$ 90% of LyC photons are reprocessed by dust grains). Naively, one expects dust to build up over time in galaxies, giving higher dust masses and lower $f_{esc,dust}^{ray}$ over time. However, the downwards slope does not change much over time. Further unintuitively, the higher the redshift, the lower the halo mass at which average f_{esc}^{ray} values start to be <1. Meaning that, for a given halo mass, the later average curves show higher $f_{esc,dust}^{ray}$ values, which is again curious as one might expect there to be more dust and less $f_{esc,dust}^{ray}$ as time and star formation go on. That being said, Fig. 2.17 showed that the ratio of dust mass to stellar mass at fixed stellar mass does not change measurably over time, in turn suggesting that the average dust f_{esc}^{ray} could also not change appreciably over time. It could be that this change is due to the cosmological expansion of cells, as $\tau_{dust}\propto a^{-2}$.

The right panel shows the density map of $f_{esc,dust}^{ray}$ as a function of halo mass at z=5. The scatter around the average curves is quite low until $\approx 3\times10^{10}\,M_\odot$, after which it spans almost one dex. The magenta diamond shows the average $f_{esc,dust}^{ray}$ reported by Yoo et al. [2020] in zoom-in[1] simulations of

---

[1] zoom-in simulations use a lower resolution simulation as initial conditions to simulate an object or a group of

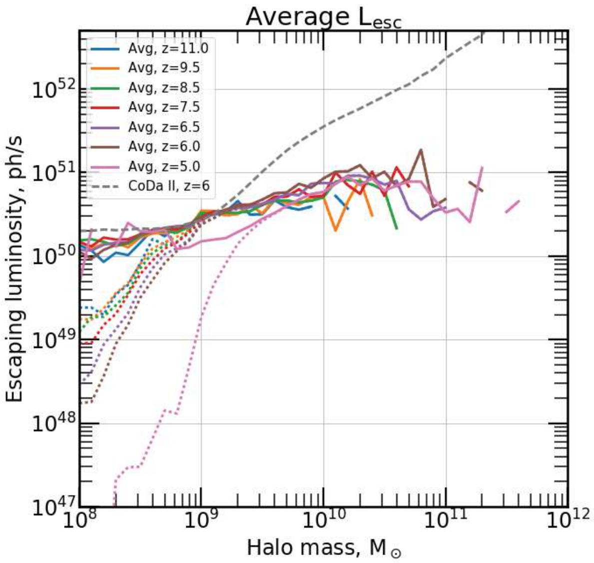


Figure 4.7: Full lines (dotted lines) show the average L_{esc} when accounting for HI and dust absorption, for the star forming (whole) sample of haloes in CoDa 2.5. The equivalent of the full lines in CoDa II at z=6 is shown in dashed grey lines.

a $10^{11}\,M_\odot$ z=4-3 galaxy analogue (≈ 0.999). Our highest $f^{ray}_{esc,dust}$ at z=5 is much lower, and is closer to 0.7, while the CoDa 2.5 average is almost one dex lower. However, the simulations of Yoo et al. [2020] are quite limited in scale and don't cover the individual zoom-in galaxies all the way to r_{200}. Furthermore, their DTM ratio is fixed based on galaxy integrated values from Rémy-Ruyer et al. [2014], avoiding the possibility of very high DTM cells as in our simulation.

4.4 Escaping luminosities in CoDa 2.5

Fig. 4.7 shows the L_{esc}[2](or the halo luminosity that escapes through galactic HI and dust to r_{200}) as a function of halo mass. The full curves show the average for the star forming population of haloes, whereas the dotted curves show the average for the whole sample. Both sets of average curves increase monotonically with halo mass. The star forming average L_{esc} appear s to behave as a power-law of halo mass. At masses lower than $10^9\,M_\odot$, the CoDa 2.5 star forming average curve is very close to the z=6 CoDa star forming average L_{esc}. At higher masses, the latter splits from the CoDa 2.5 curves, and increases at a much greater rate, reaching values of close to $3 \times 10^{52}\,\mathrm{ph}\cdot\mathrm{s}^{-1}$ at $10^{11}\,M_\odot$, versus $\approx 10^{51}\,\mathrm{ph}\cdot\mathrm{s}^{-1}$ at the same mass in the CoDa 2.5. This difference between CoDa 2.5 and CoDa II is both due to the lower $f^{ray}_{esc,gas}$ values in the former simulation, and the introduction of $f^{ray}_{esc,dust}\leq$ in CoDa 2.5. Before z=5, the star forming curves are almost equal, except at high mass where there is substantial noise. Near $10^9\,M_\odot$, both sets of curves join for redshifts >5, and are interchangeable at higher masses.

---

objects in more detail and possibly with different physics.

[2] $L_{esc}= L_{intr}\times f^{ray}_{esc,dust}\times f^{ray}_{esc,gas}\times f^{sub}_{esc}$, where L_{intr} is the intrinsic photon production or intrinsic luminosity, and f^{sub}_{esc} is the stellar particle escape fraction which is 0.5 in CoDa 2.5

Below this mass, the full sample L_{esc} average curves are below the star forming averages, and have considerably stronger slopes. As time progresses from z=11 to z=6, the slopes become stronger: the full sample average decreases from $2 \times 10^{49}\,\mathrm{ph \cdot s^{-1}}$ to $2 \times 10^{48}\,\mathrm{ph \cdot s^{-1}}$ at $10^{8}\,\mathrm{M_{\odot}}$. This is the sign of star formation suppression: Indeed, galaxies where star formation is suppressed have less stars and older redder stellar populations, which both lead to lower intrinsic ionising luminosities. Between z=6 and z=5, the full sample average breaks away dramatically by close to 1.5 dex from the other full sample averages, and only meets the star forming sample average at $\approx 3 \times 10^{9}\,\mathrm{M_{\odot}}$. The sudden change corresponds to a sudden rise in star formation suppression: As the box finishes reionising, and the ionised regions percolate, the UVBG increases, substantially increasing radiative feedback in the lowest mass haloes that cannot cool enough to keep forming stars at the same rate. In fact, the star forming sample average also changes between z=6 and z=5. It decreases by around 0.2 dex from $5 \times 10^{8}\,\mathrm{M_{\odot}}$ to $8 \times 10^{9}\,\mathrm{M_{\odot}}$.

4.5 Galactic ionising photon budget in CoDa 2.5

4.5.1 Photon budget

As a function of mass Fig. 4.8 shows the galactic ionising photon budget in CoDa 2.5. The left panel shows the intrinsic luminosity per mass bin. The total ionising intrinsic luminosity per mass bin is a rising function of mass. This distinctive shape is given by the decrease in halo abundance and increase in average halo L_{esc}(as per Fig. 4.7) as a function of mass. As time progresses, the curves' maxima shift from $10^{10}\,\mathrm{M_{\odot}}$ at z=11.0 to $4 \times 10^{10}\,\mathrm{M_{\odot}}$ at z=6.0, as more and more massive bright efficient star forming galaxies appear. The distribution has a sharp cut off at the massive end which is not seen in the equivalent CoDa II curves of Fig. 1.11. The under-representation of massive galaxies, and the absence of the most massive galaxies from CoDa II is simply due to the much smaller box size. The z=5.0 curve is singular in its behaviour: There is a dramatic reduction in the intrinsic emissivity between 10^{8} and $2 \times 10^{9}\,\mathrm{M_{\odot}}$, as if the distribution has been slashed when compared to the z=6 curve. This is most likely yet another imprint of the strong star formation suppression in our new set-up.

The right panel shows the rate of photons escaping to the IGM when considering the absorption by neutral Hydrogen and dust. The curves have a normal-like shape, with modes near $10^{9}\,\mathrm{M_{\odot}}$, except at z=5.0 when the mode is closer to $3 \times 10^{9}\,\mathrm{M_{\odot}}$. The modes of the CoDa 2.5 distributions are at lower masses than in CoDa II (the shift is ≈0.3-0.5 dex in mass). At the same time, the shapes of the distributions in CoDa 2.5 are much narrower and shallower: at z=9.5 in CoDa 2.5, the budget's low mass wing is higher by up to a factor of 2 than in CoDa II's at z=10.1. Accordingly, the high mass wing of the budget in CoDa 2.5 is substantially lower than in CoDa II at z=10.1 (by between 0.5 and 1 dex). Similarly, the z=6 CoDa II budget is up to 1.5 dex higher than in CoDa 2.5 at z=6, and also has less photons from low mass haloes, but many more from the high mass wing of the distributions. This can be expected as we find f_{esc}^{ray} to be up to 100 times lower in CoDa 2.5 for the highest mass haloes. In summary, there are more photons escaping to the IGM in CoDa II, and CoDa II's budgets have a higher relative contribution from high mass haloes.

When comparing the escaping emissivities to the intrinsic ones, one clearly sees the marked impact

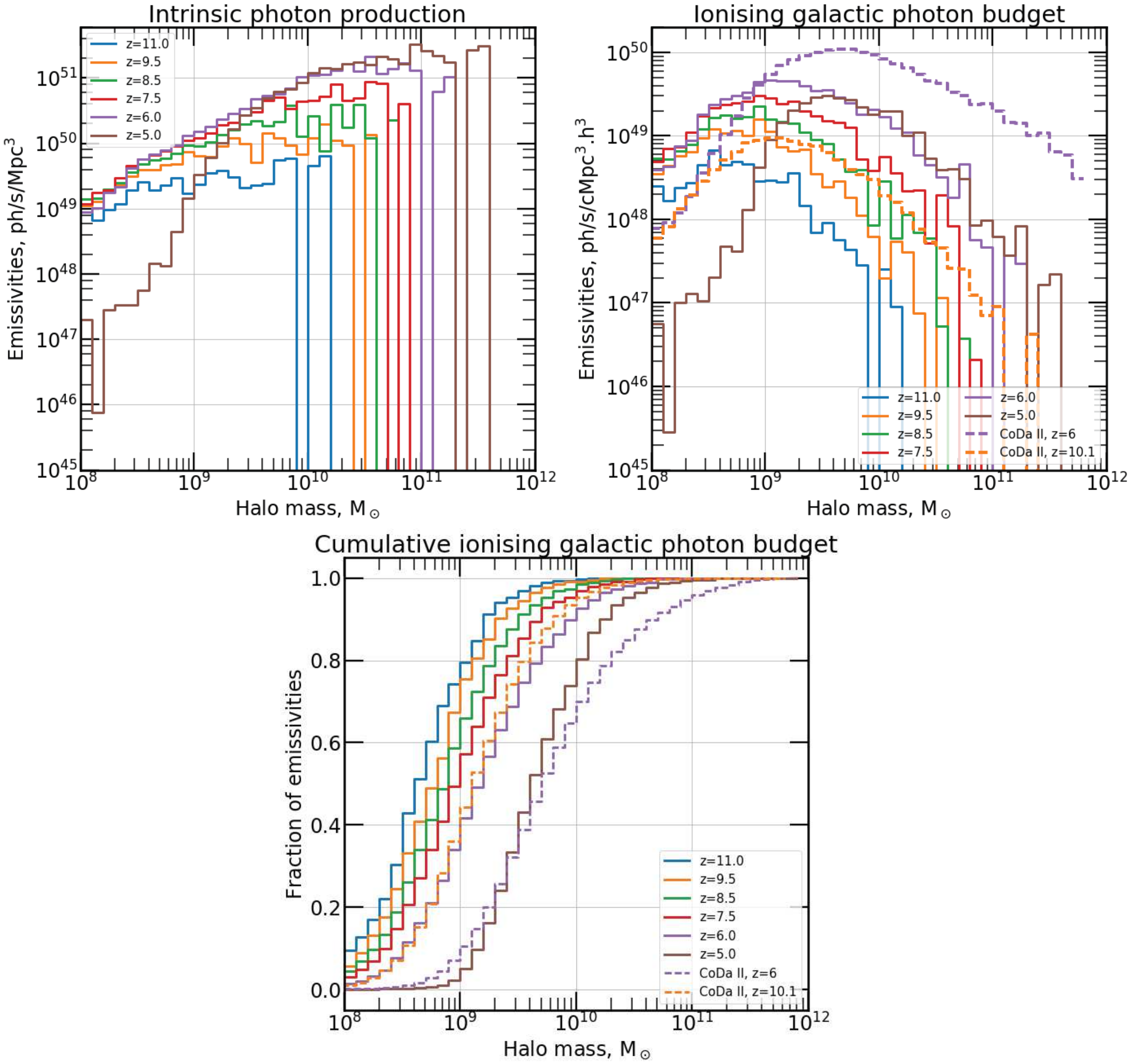


Figure 4.8: *Left*: Intrinsic photon budget at several epochs the mass bin total density of photons produced within haloes. *Right*: Galactic ionising photon budget of the CoDa 2.5 simulation, in mass bin total rate of photons that escape to the IGM through neutral galactic gas per Mpc^3. Dotted lines show the curves from CoDa II at z=10.1 and at z=6. *Bottom*: Cumulative version of the galactic ionising photon budget (full lines of the left panel) as a function of mass. The dashed lines in both panels show results from CoDa II.

of neutral gas and dust column densities on the photon emissivity that reaches the IGM. The stellar particle f_{esc}^{sub}=0.5 shifts the budget curves downwards from the intrinsic curves for all masses, the amplitude of the shift is seen clearly for low mass haloes at the end of Reionization for which both $f_{esc,dust}^{ray}$ and $f_{esc,gas}^{ray}$ are close to 1. From $\approx 10^{9}\,M_{\odot}$ ($3\times10^{9}\,M_{\odot}$ at z=5.0) onwards, the sets of curves diverge, with the escaping emissivity curves decreasing monotonically from that threshold onwards. For instance, at masses of close to $10^{11}\,M_{\odot}$, the absorption of LyC photons by neutral gas is such that the escaping photon curves are over 1000 times lower than the photon production curves. At z=6, $10^{11}\,M_{\odot}$ haloes produce less than 20 times less photons than $10^{9}\,M_{\odot}$ haloes. In CoDa II, the ratio was closer to 5 times less. This again shows that in our new calibration and set-up, massive bright galaxies are less important than they were in CoDa II.

The right panel of Fig. 4.8 shows the cumulative galactic ionising photon budget as a function of mass. It shows that around 80 % of the ionising photons that reach the IGM between z=5 and z=11 came from galaxies with halo masses within $10^{8}\,M_{\odot}$ to $2\times10^{10}\,M_{\odot}$. In CoDa II, the corresponding interval is $6\times10^{8}\,M_{\odot}$ to $3\times10^{10}\,M_{\odot}$, once more showing that the new calibration for late Reionization, and the new physics favour the contribution of the lowest mass haloes over the more massive ones. The CoDa 2.5 z=6 curve is very close to the z=10.1 curve from CoDa II, which attests to how the evolution of the photon budget with time is also vastly different between the two simulations. The CoDa II z=6 cumulative curve is the closest to the z=5 CoDa 2.5 curve. That they are similar is not so surprising, as they are both at redshifts just after their respective box's Reionization. However, the CoDa II curve shows that the relative contribution to photon budget of high mass haloes is around $\approx$2-3 times greater than in CoDa 2.5, but to a lesser degree so is that of the $\leq 2\times10^{9}\,M_{\odot}$ haloes (this is probably due to more suppression in the CoDa 2.5). As time and Reionization progress, the importance of low mass haloes diminishes (the cumulative distribution shifts to higher masses), because of the formation of brighter haloes and the effects of star formation suppression on low mass haloes. As seen before, the difference between z=6 and z=5 is more pronounced than between other redshifts. Between these last two snapshots, not only is the cumulative distribution shifted more than at previous times, but the slope of the distribution also changes due to the significant reduction in L_{intr} that occurs (as per Fig. 4.7).

As a function of magnitude The left panel of Fig. 4.9 shows the galactic ionising photon budget of CoDa 2.5 as a function of halo M_{AB1600}. At all times, the distributions resemble asymmetric normal functions with a sharp cut on the faint side ($M_{AB1600}\approx$-10), and a noisy cut on the bright side that moves to brighter M_{AB1600} as time goes by. The mode at z=6 is M_{AB1600}= -13. Between z=11 and z=6, the distributions are shifted towards higher emissivities, and the maximum increases from near $4\times10^{50}\,\mathrm{ph}\cdot\mathrm{s}^{-1}\cdot\mathrm{Mpc}^{-3}$ to $2\times10^{51}\,\mathrm{ph}\cdot\mathrm{s}^{-1}\cdot\mathrm{Mpc}^{-3}$. Between z=6 and z=5, there is little change for $M_{AB1600}\leq$ -15. But, at fainter magnitudes, the whole distribution is substantially diminished in between. For instance, the maximum is cut to $\approx 10^{51}\,\mathrm{ph}\cdot\mathrm{s}^{-1}\cdot\mathrm{Mpc}^{-3}$. This is another very suggestive clue as to the increase in star formation suppression between z=6 and z=5: The faint haloes that are the most susceptible to star formation suppression (having the least star formation, and the least gas to shield themselves from SNe and radiative suppression), form no new stellar particles, and as their stellar population ages, so does the luminosity decrease. The z=6 budget is close to the CoDa

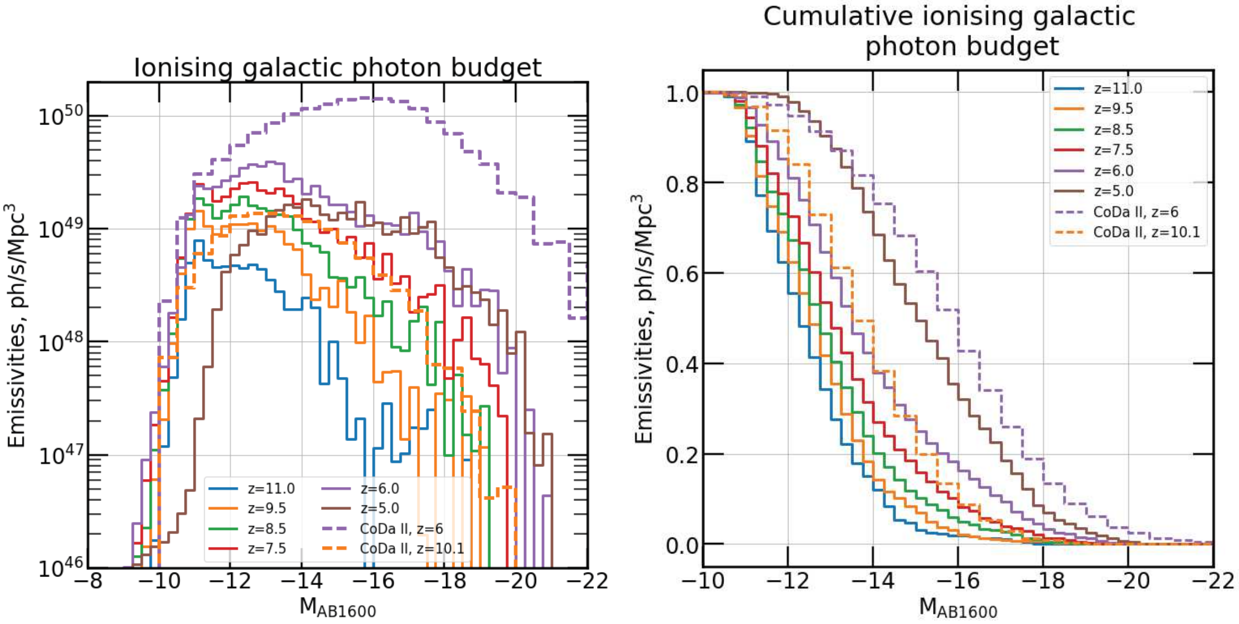


Figure 4.9: *Left*: Galactic ionising photon budget of the CoDa 2.5 simulation. Full lines denote the M_{AB1600} bin total rate of photons that escape to the IGM through neutral galactic gas per Mpc^3. Dotted lines show the magnitude bin total density of photons produced within haloes. *Right*: Cumulative version of the galactic ionising photon budget (full lines of the left panel) as a function of M_{AB1600}. The dashed lines in both panels show results from CoDa II.

II budget for $M_{AB1600} \geq$ -11, albeit with a slightly softer faint cut-off. In brighter bins, the CoDa II budget is significantly more emissive, especially for the brightest galaxies, where it is over 1.25 dex greater. The budgets in CoDa at z=10.1 and in CoDa 2.5 at z=9.5 differ similarly, however the gap between the two curves is smaller, once more demonstrating that CoDa II's Reionization is much more reliant on the photons from (massive) bright galaxies, and that many more photons are generated in CoDa II towards the end of Reionization than in CoDa 2.5.

The right panel shows the cumulative version of the galactic ionising photon budget as a function of magnitude. It shows that between z=11 and z=7.5, $\approx$80 % of the ionising photons that reionised the universe came from haloes within M_{AB1600}=-16.25-11.25. By z=6 the same fraction of photons came from the magnitude interval M_{AB1600}=-17,-11.25, and at z=5 the corresponding interval was M_{AB1600}=-17.5,-13.75. At later times, the brighter galaxies that form shift the interval to brighter magnitudes. Between z=5 and z=6, the ongoing star formation suppression shuts down the L_{esc} in haloes with $M_{AB1600} \gtrsim$ -12, further favouring the contribution of bright haloes. As in Fig. 4.8 we see that the CoDa II curves at z=6 and z=10.1 are the closest to the CoDa 2.5 cumulative budget curves are z=9.5 and z=5, respectively. At z=9.5, haloes brighter than $\approx$-13.75 contribute around 20 % of the photons that reach the IGM in CoDa 2.5, versus close to 50% in CoDa II. The gap between the cumulative budget at z=6 is even greater, illustrating that though in both simulations, as time goes on the relative contribution of (massive) bright galaxies increases, it does so more rapidly in CoDa II.

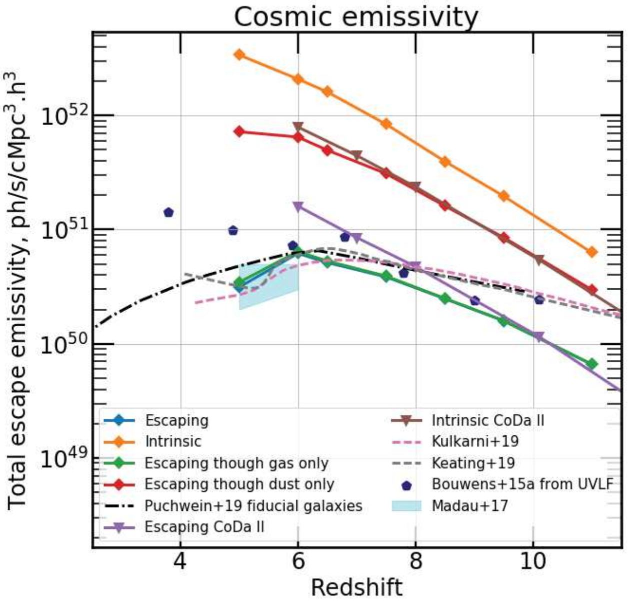


Figure 4.10: Box total emissivity of CoDa 2.5, and CoDa II (both escaping and intrinsic, i.e.: that contribute to the IGM, and that are produced). For comparison, the model predictions of Puchwein et al. [2019]; Madau [2017], the simulation results of Kulkarni et al. [2019]; Keating et al. [2019], and the observationally derived points of Bouwens et al. [2015] are provided.

4.5.2 Ionising emissivities

Fig. 4.10 shows the evolution of CoDa 2.5 total ionising emissivity as a function of redshift. Although the intrinsic emissivity and the escaping through dust emissivity increase monotonically with decreasing redshift, and unlike in CoDa II, the total escaping and the escaping through gas emissivities decrease between z=6 and z=5. Despite the latter trends being stronger in CoDa 2.5, a similar turning point is reported by Puchwein et al. [2019], whose galaxy only model curve we intersect at z=6. Kulkarni et al. [2019]; Keating et al. [2019]; Finkelstein et al. [2019] also find similar decreases in the total escaping galactic ionising emissivity after z≈7.5. This means that CoDa 2.5 is also in agreement with the constraints from Madau [2017], who report requiring such a trend to reproduce constraints on the ionisation of the IGM. Our total escaping emissivity is also relatively close to observations from Bouwens et al. [2015] before z=6 (though their observations should be compared with a magnitude cut curve). Dust lowers the total amount of photons that escape to the IGM by a quasi constant shift that is several times lower than the effect of HI gas, but nonetheless represents a sizeable difference. Note that the total escaping and the escaping through gas curves are more similar than the difference between the intrinsic and the escaping through gas curves might suggest at first glance. This is probably because the intrinsic and dust only curves are biased towards dense regions with massive dusty galaxies that produce the most photons, whereas the total and gas only curves are biased towards lower density (and dust) regions that contribute the most photons to the IGM. There are several interesting differences between CoDa 2.5 and CoDa II: Though the total escaping emissivity of CoDa 2.5 is close to the CoDa II result at z=10, CoDa II's escaping emissivity increases faster with redshift, and ends up almost twice as high as the curve from CoDa 2.5 at z=6. This is most likely because of the lower total f_{esc}^{ray} we measure in CoDa 2.5.

4.6 Discussion

When compared to CoDa II, the galactic ionising photon budget in CoDa 2.5 favours the contribution of lower mass and fainter galaxies to reionising the IGM. This is due in part to the inclusion of dust absorption in galaxies, but much more so due to the lower values of $f_{esc,gas}^{ray}$ indicating greater absorption by the HI galactic gas. The resulting photon budget shows that the mass (M_{AB1600}) range $10^8\,M_\odot$ to $2 \times 10^{10}\,M_\odot$ ($M_{AB1600}\approx$-17.5,-11.25) dominates Reionization by producing over 80% of the photons that escape to the IGM from galaxies during the EoR.

Other numerical studies of the f_{esc} of galaxies report higher values for $M_{halo} \gtrsim 10^{10}\,M_\odot$, as well as higher values of $< f_{esc} >_{SFR_{10}}(z)$. It may be that $f_{esc,gas}^{ray}$ are too low in CoDa 2.5. However, the comparison with other literature is difficult. Indeed, herein we aimed at producing a set-up that reproduces the observations of Reionization as much as possible. A lot of the work done on f_{esc} has been carried out in smaller more resolved boxes or using zoom-in simulations, where f_{esc} is studied outside of the context of a realistic Reionization simulation [as in Yoo et al., 2020; Ma et al., 2020; Kimm and Cen, 2014; Yajima et al., 2011; Razoumov and Sommer-Larsen, 2010]. Elsewhere, boxes of a comparable (though smaller) size have been used [for instance Katz et al., 2018; Rosdahl et al., 2018], but due to their better resolution, different sub-grid recipes (for instance Katz et al. [2018] shut off star formation for 20 Myr in cells where supernova explosions happen), and their choice of a reduced speed of light, our boxes can be expected to give different $< f_{esc} >_{SFR_{10}}$ results, as well as different predictions for the photon budget of galaxies[3]. At the same time, Katz et al. [2018] only present an average neutral fraction matching the constraints of Fan et al. [2006], and ionisation rates as proof of a realistic Reionization, and do not show matches to the Lyα forest data or the UVLF. The higher values of $< f_{esc} >_{SFR_{10}}$in SPHINX [Rosdahl et al., 2018] could explain its much earlier Reionization. In our case, striving to match Lyα forest results near at the end of the EoR in our set-up, has yielded decreasing values of $< f_{esc} >_{SFR_{10}}$ and of the total escaping galactic ionising emissivity with redshift. This could be encouraging as recent model papers find it impossible to properly match constraints after z=6 if the total escaping galactic ionising emissivity does not decrease with redshift (which could be caused by a decrease in $< f_{esc} >_{SFR_{10}}$) [see Madau, 2017; Puchwein et al., 2019; Finkelstein et al., 2019], and similar results have been reported in simulations by Kulkarni et al. [2019]; Keating et al. [2019].

There are no RHD simulations of comparable scale and physical scope that track metals and dust on the fly in galaxies, or study the effect of dust on Reionization and on f_{esc}(Yoo et al. [2020] use idealised zoom-in simulations that do not cover r_{200}, in a handful of galaxies that have fixed dust to metal ratios). That being said, Wu et al. [2020] do show similar extinction results in larger boxes that were run without radiative transfer. Since constraints on high redshift dust are difficult to produce,

---

[3] On top of this, their $< f_{esc} >_{SFR_{10}}$ result counts photons as having exited galaxies when they are in cells of density less than 180 times the average baryon density, which could cause differences when compared to our approach. In CoDa II we found significant absorption in cells with less than 180 times the average baryon density, albeit with the much larger simulation cells, which could mean that Katz et al. [2018] overestimate the number of photons that reach the IGM.

we must content ourselves with the favourable comparisons between CoDa 2.5 and SAMs. However, even with seemingly realistic dust densities, our results are susceptible to our choice of $\kappa_{\rm dust}$.

4.7 Résumé des résultats

Dans ce dernier chapitre, j'ai exploré les répercussions des ajouts au code vus dans le Chp. 2, ainsi que les modifications de la calibration du code étudiées dans le Chp. 3, sur le budget de photons étudié dans le Chp. 1.

En particulier, j'ai montré que le critère de température pour la formation stellaire avait un grand effet sur le $f_{\rm esc}$ des galaxies. Cet effet se comprend facilement: en interdisant la formation stellaire dans les cellules chaudes, ionisées et transparentes aux photons ionisants, on limite la formation stellaire aux cellules relativement opaques aux mêmes photons, ce qui peut réduire $f_{\rm esc}$. En même temps, la poussière réduit le $f_{\rm esc}$ dans les halos massifs. Aux petites masses, les $f_{\rm esc}$ sont similaires dans la nouvelle simulation et dans CoDa II, mais aux grandes masses la différence est dramatique: dans CoDa II les galaxies de masse $\approx 10^{11}\,{\rm M}_\odot$ avaient des valeurs de $f_{\rm esc}$ autour de 0.1 à z=6, dans la nouvelle simulation elles avoisinent 10^{-3}. Bien que la poussière ait un effet mesurable (une réduction de la luminosité des galaxies avec $M_{\rm halo}\approx 10^{11}\,{\rm M}_\odot$ d'environ 80-90% à z=6), l'effet du gaz neutre est bien plus prononcé (seuls ≈ 1 % des photons parviennent jusqu'au milieu inter-galactique pour les mêmes galaxies à la même époque). Au fil du temps le $f_{\rm esc}$ des galaxies augmente, surtout aux faibles masses ($\lesssim 10^{9}\,{\rm M}_\odot$) où la moyenne passe de 0.4 à z=11, à 0.8-0.9 à z=6. De façon surprenante, l'absorption due à la poussière n'évolue pas en moyenne avec le redshift.

A cause de ces différences, la contribution relative des galaxies massives et lumineuses au budget de photons ionisants est plus faible que dans CoDa II. Dans la nouvelle simulation, environ 80 % des photons ionisants qui parviennent au milieu intergalactique sont produits par des galaxies de masses $2\times10^{8}<M_{\rm halo}<6\times10^{9}\,{\rm M}_\odot$ à z=7 (contre $6\times10^{8}<M_{\rm halo}<3\times10^{10}\,{\rm M}_\odot$ dans CoDa II à z=7). De manière générale, plus de photons ionisants parviennent au milieu inter-galactique que dans CoDa II. Comme dans CoDa II, l'intervalle en masse et en magnitude qui domine le budget de photons ionisants se décale vers les grandes masses et luminosités avec le temps. Ce décalage est particulièrement prononcé dans la nouvelle simulation entre $z = 6$ et $z = 5$, quand la suppression de la formation stellaire dans les galaxies de faible masse devient très forte (entre ces deux époques la contribution des galaxies de $10^{8}\,{\rm M}_\odot$ au budget diminue de plus de 2 dex, contre 0.2 dex dans CoDa II).

L'un des résultats les plus intéressants vient de l'étude des émissivités ionisantes totales de la nouvelle simulation. En effet, alors que dans CoDa II elles augmentaient au fil du temps, dans la nouvelle simulation elles diminuent après $z = 6$. En fait, cette décroissance semble être poussée par une augmentation de l'absorption de photons ionisants par le gaz d'Hydrogène neutre dans les galaxies. En effet, alors que le nombre total de photons émis augmente entre $z = 6$ et $z = 5$, l'émissivité totale ne considérant que la poussière stagne, et l'émissivité totale ne considérant que le gaz diminue. De par cette diminution, la nouvelle simulation est bien plus comparable aux résultats simulés de Kulkarni et al. [2019]; Keating et al. [2019], et aux modèles de Puchwein et al. [2019]; Madau [2017]. Que nos résultats d'émissivité totale soient plus proches de ceux de Kulkarni et al. [2019]; Keating et al. [2019], qui eux aussi ont calibré leurs simulations de façon à reproduire les observations de l'épaisseur optique de la forêt Lyman alpha est un signe encourageant. De plus, nos résultats sont compatibles avec le modèle de Madau [2017], qui met en avant la nécessité d'avoir une émissivité totale décroissante après $z = 6$ afin de reproduire l'évolution de l'ionisation de l'Hydrogène des contraintes observationnelles.

De même, quand on considère le f_{esc} moyen pondéré par le taux de formation stellaire: alors que dans CoDa II [et dans Kimm and Cen, 2014; Katz et al., 2018; Rosdahl et al., 2018] cette valeur évoluait peu avec le redshift, dans la nouvelle simulation elle diminue avec le temps ce qui nous rapproche des tendances utilisées dans les modèles de Finkelstein et al. [2019]; Puchwein et al. [2019]; Ferrara and Loeb [2013] . Finkelstein et al. [2019] montrent que des modèles avec un f_{esc} moyen constant, qui n'évolue ni en masse ni en redshift, peinent à reproduire les contraintes sur l'émissivité totale, ce qui semble faire écho à nos résultats.

BIBLIOGRAPHY

Bouwens R.J., et al.. Reionization After Planck: The Derived Growth of the Cosmic Ionizing Emissivity Now Matches the Growth of the Galaxy UV Luminosity Density. *The Astrophysical Journal*, 811(2):140, 2015.

Fan X., et al.. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132:117–136, 2006.

Ferrara A. and Loeb A. Escape fraction of the ionizing radiation from starburst galaxies at high redshifts. *Monthly Notices of the Royal Astronomical Society*, 431(3):2826, 2013.

Finkelstein S.L., et al.. Conditions for Reionizing the Universe with a Low Galaxy Ionizing Photon Escape Fraction. *The Astrophysical Journal*, 879:36, 2019.

Katz H., et al.. A Census of the LyC Photons that Form the UV Background During Reionization. *arXiv:1802.01586 [astro-ph]*, 2018. ArXiv: 1802.01586.

Keating L.C., et al.. Long troughs in the Lyman-α forest below redshift 6 due to islands of neutral hydrogen. *arXiv e-prints*, page arXiv:1905.12640, 2019.

Kimm T. and Cen R. Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars. *The Astrophysical Journal*, 788:121, 2014.

Kulkarni G., et al.. Large Ly α opacity fluctuations and low CMB τ in models of late reionization with large islands of neutral hydrogen extending to $z < 5.5$. *Monthly Notices of the Royal Astronomical Society: Letters*, 485(1):L24–L28, 2019.

Lewis J.S.W., et al.. Galactic ionising photon budget during the Epoch of Reionisation in the Cosmic Dawn II simulation. *Monthly Notices of the Royal Astronomical Society*, 2020.

Ma X., et al.. No missing photons for reionization: moderate ionizing photon escape fractions from the FIRE-2 simulations. *arXiv:2003.05945 [astro-ph]*, 2020. ArXiv: 2003.05945 version: 1.

Madau P. Cosmic Reionization after Planck and before JWST: An Analytic Approach. *The Astrophysical Journal*, 851:50, 2017.

Puchwein E., et al.. Consistent modelling of the meta-galactic UV background and the thermal/ionization history of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 2019.

Razoumov A.O. and Sommer-Larsen J. Ionizing Radiation from z = 4–10 Galaxies. *ApJ*, 710(2):1239, 2010.

Rosdahl J., et al.. The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Monthly Notices of the Royal Astronomical Society*, 479:994–1016, 2018.

Rémy-Ruyer A., et al.. Gas-to-dust mass ratios in local galaxies over a 2 dex metallicity range. *Astronomy and Astrophysics*, 563:A31, 2014.

Wu X., et al.. Photometric properties of reionization-epoch galaxies in the Simba simulations. *Monthly Notices of the Royal Astronomical Society*, page staa1044, 2020. ArXiv: 1911.06330.

Yajima H., et al.. Escape fraction of ionizing photons from high-redshift galaxies in cosmological SPH simulations. *Mon Not R Astron Soc*, 412(1):411–422, 2011.

Yoo T., et al.. On the origin of low escape fractions of ionizing radiation from massive star-forming galaxies at high redshift. *arXiv e-prints*, 2001:arXiv:2001.05508, 2020.

CONCLUSIONS AND PERSPECTIVES

Motivation

Over the next decade or so, new instruments will allow astronomers to peer further and deeper into the Epoch of Reionization than ever before. In fact, 21 cm observations of the IGM during Reionization, will represent a new, unique probe into Reionization. As of such, numerical simulations will remain a crucial tool with which to guide the interpretation of this new data, and advance our physical understanding of the complex coupled physics that govern the EoR. In particular, simulations can help us understand how Reionization unfolds. One centrally important question to this understanding, and that interested us in this thesis is : Which galaxies contributed to Reionization? (Or more specifically : What is the ionising photon budget of galaxies?) Indeed, the nature and distribution of the dominant ionising sources informs the geometry of Reionization, the sizes of ionised bubbles in the IGM, and the speed with which they grow and merge. To answer this question is highly non-trivial as it requires simulations that capture all the important physics that mostly cannot be resolved, across the massive scales necessary to avoid biases due to sampling variance, whilst satisfying constraints on the evolution of IGM ionisation and on the density of sources. Due to the daunting computational costs of such an endeavour, simulations in the literature often fail to match some of the observed constraints on Reionization, to posses the necessary scale, or to consistently feature some essential physics.

CoDa II

To explore this question, I analysed the CoDa II simulation, a massive fully coupled RHD simulation of the EoR, that agrees with some of the major constraints that are known (in particular on the UVLF). I discovered that in CoDa II, Reionization is driven by haloes of intermediate masses ($6 \times 10^8\,\mathrm{M}_\odot$ to $3 \times 10^{10}\,\mathrm{M}_\odot$). Lower mass haloes are too faint due to their inefficient star formation, whereas higher mass haloes are too rare and have low LyC escape fractions. In fact, I showed that $\mathrm{f}_{\mathrm{esc}}$ decreases with mass in CoDa II. It turns out that higher mass haloes are better able to recombine ionised gas in their densest cells despite higher star formation. This trend of $\mathrm{f}_{\mathrm{esc}}$ with mass has also been reported elsewhere [for instance in Yajima et al., 2011; Kimm and Cen, 2014; Katz et al., 2018], but never with the statistical significance of the CoDa II dataset, and often without a realistic Reionization occurring. Though the average $\mathrm{f}_{\mathrm{esc}}$ is a function of halo mass, the values of $\mathrm{f}_{\mathrm{esc}}$ are highly scattered around this average. I showed that the main driver of this scatter is differences in the distribution of neutral halo gas (the higher the maximum neutral gas density, the lower the $\mathrm{f}_{\mathrm{esc}}$ at fixed mass). As Reionization progresses, bright galaxies become more important as they increase in number, and as star formation in faint galaxies is suppressed.

However, the agreement between CoDa II and Reioinzation constraints is far from perfect. Indeed, the post overlap J21 is larger than observed constraints, and mock Lyα observations show that the average neutral density in the IGM is too low after overlap. In fact, Madau [2017]; Finkelstein et al. [2019]; Kulkarni et al. [2019]; Keating et al. [2019] find that it's necessary for the total ionising photon budget to decrease with redshift, and to adopt late Reionization scenarios (with overlap happening at $z \leq 5.5$), to match constraints on the evolving neutrality of IGM gas near the end of the EoR and later. At the same time, there are physics missing from CoDa II, that may play an important role in Reionization, and potentially reduce the number of excess photons/ionisations. In order to pave the way for CoDa III, and to seek to improve on the results from CoDa II, I sought to build on the physical models in RAMSES-CUDATON, and to perform new simulations with a late Reionization calibration.

CoDa 2.5

First, I adapted a version of the RAMSES-RT [Rosdahl et al., 2013] metal cooling to the simulation code, as well as a model for helium cooling. Both cooling mechanisms play an important role by cooling the high density gas, and significantly increase star formation in the test runs that I performed. Further, I sought to implement changes that might serve to reduce the excess of photons after overlap. To do so, I modified the stellar emissivity laws in RAMSES-CUDATON to take into account stellar population ageing and chemical enrichment (both reduce ionising luminosity). This slightly reduces the average number of ionising photons emitted per stellar mass, and delays Reionization. At the same time, I coupled a physical dust model (Dubois 2020, in prep) to the RAMSES-CUDATON code, and implemented on the fly LyC dust absorption. I found a calibration of the dust model that is able to reasonably match both the few observations of high redshift dust masses in galaxies, and predictions from SAMs, justifying my approach. With this believable setup, dust has a measurable impact on the escape of ionising photons to the IGM, which it can reduce by more than 90% in the most massive galaxies, and therefore having a measurable impact on their photon budget. In this setup, dust is present early enough and dense enough to produce significant extinction of the bright ($\mathrm{M_{AB1600}}$<-19) end of the UVLF, even at high redshifts [as in Wu et al., 2020].

Second, I sought to make Reionization occur later by altering the CoDa II setup, whilst reducing the density of ionising photons toward the end of the EoR. I found that by including a temperature criterion for star formation, I could produce simulations that were far closer to the constraints on the opacity of the Lyα forest towards the end of Reionization. Not only that, but in the CoDa 2.5 simulation, I found a much better agreement with observations of the ionisation rate Γ, whilst retaining a realistic UVLF. Interestingly, this also meant that the total ionising emissivity decreases with redshift after z=6, in line with the expectations of Madau [2017]; Finkelstein et al. [2019]; Kulkarni et al. [2019]; Keating et al. [2019]. This decrease seems to be driven mostly by star formation suppression in low mass haloes. With this setup, the photon budget is similar to that of CoDa II, but I found that faint objects have a larger relative contribution than in CoDa II, and bright objects have a lower relative contribution. It turns out that this is driven by a much steeper average slope of $\mathrm{f_{esc}}$ as a function of mass than in CoDa II. With massive galaxies ($\gtrsim 10^{11}\,\mathrm{M_\odot}$) having $\mathrm{f_{esc}}$ of less than 10^{-3}. This is both due to the addition of dust absorption and to the temperature criterion that favours lower $\mathrm{f_{esc}}$ by restricting star formation in hot, ionised and LyC transparent cells (In turn demonstrating the powerful effect that sub-grid recipes can have on the results of EoR simulations).

The road to CoDa III

However, there remains much to be done. For instance, the calibration may yet be improved, ideally to match the Lyα forest for z>5.2 (where CoDa 2.5 over-predicts the neutrality of the IGM) and, in bigger boxes, the distribution of Lyα forest opacities. At the same time, the agreement with the observed UVLF could yet be improved : in CoDa 2.5 the simulated UVLF at $z \leq 6$ shows a significant excess between $M_{AB1600} \approx$ -17 to $M_{AB1600} \approx$ -20 that does not appear in observations (I am currently investigating the origin of this feature). Elsewhere, important testing of the dust model and its post processing needs to be done. For example, we have yet to examine the effect of different assumptions for the dust grain cross section per mass, on $f^{ray}_{esc,dust}$, extinction of the UVLF, or the slope of the UV continuum in galaxies (β). In the case case of β, it may be that such changes could improve the agreement with the average observations from Finkelstein et al. [2012a]. Regarding the dust model, we have yet to perform resolution tests, which are crucial since CoDa III shall have a higher resolution than in CoDa II and CoDa 2.5.

CoDa III and beyond

Assuming that we can replicate and improve our match with the Lyα forest data, the UVLF, the ionising emissivity Γ, and consolidate the work done on dust in CoDa 2.5, then CoDa III will provide a very interesting dataset. As it will have improved resolution, and a better agreement with constraints on the IGM during Reionization. Not only will this represent a further opportunity to study the photon budget and escape fraction, but will also give us an interesting window into the effects of our new setup (late Reionization, star formation temperature criterion, metal and helium cooling, dust, evolving stellar emissivities) on the results from CoDa II. For instance, as CoDa III will also use constrained initial conditions, we will be able to compare predictions for the Reionization of the Local Group between an early Reionization scenario and a late Reionization scenario.

Beyond CoDa III, there are many future advances that we can look forward to. On the simulation side of things, gains in both hardware and software (mostly) should eventually allow us to bridge the gap between very large scale simulations like CoDa II, and more resolved but small boxes like SPHINX [Rosdahl et al., 2018], marking a huge turning point for simulators of Reionization. In terms of observations, we can expect both improvements of existing constraints, and new innovative constraining observations. Particularly, the JWST will extend the flux limit of high redshift galaxies by around two magnitudes. At the same time, it will provide new constraints on the properties of high redshift galaxies. For instance, it will likely measure their gas-phase metallicity, a crucial aspect in understanding supernova feedback and chemical enrichment during the EoR. Meanwhile, 21 cm astronomy promises to revolutionise EoR science. Indeed, experiments like EDGES are already calling into question our understanding of the state of cosmic gas prior to Reionization, and future experiments will give us access to the tomography of the neutral IGM during Reionization. These new radio experiments will require extensive interpretation, and will give further incentive to model and perform simulations.

CONCLUSIONS ET PERSPECTIVES

Motivation

Au cours des prochaines décénies, de nouveaux instruments apparaîtront, permettant aux astronomes de sonder l'époque de la Réionisation plus profondément qu'auparavant. Notamment, une grande nouveauté sera la possibilité de directement observer le signal à 21 cm de l'hydrogène neutre pendant la Réionisation. La richesse et complexité nouvelle de ces données futures promettent le maintien des simulations numériques comme un outil crucial dans l'interprétation des données, et dans la compréhension de la physique couplée qui régit la Réionisation. En particulier, les simulations peuvent nous faire entrevoir le déroulement de la Réionisation. L'une des questions fondamentales associées, et qui m'a intéressée pendant cette thèse est : quelles sont les galaxies ayant contribué à réioniser l'Univers ? (Ou plus spécifiquement : quel est le budget de photons ionisants des galaxies ?). En effet, la nature et la distribution des sources ionisantes dominantes dicte la géométrie de la Réionisation, la taille des régions ionisées dans le milieu inter-galactique, et la vitesse à laquelle celles-ci croissent et fusionnent. Répondre à ce questionnement est non-trivial, vu que cela nécessite des simulations où sont modélisées toutes les physiques importantes (dont la plupart ne peuvent pas être résolues), sur des échelles cosmologiques nécessaires pour éviter des biais dus à la variance d'échantillonnage, tout en satisfaisant des contraintes sur l'évolution de l'ionisation du milieu intergalactique et sur la densité des sources ionisantes. A cause du coût numérique prohibitif d'une telle simulation, la littérature existante a souvent échoué à reproduire certaines contraintes, étudié des volumes trop petits, ou omis certains processus physiques importants.

CoDa II

Pour explorer des réponses à cette question, j'ai analysé la simulation CoDa II. CoDa II est une simulation de très grande échelle de la Réionisation, dans lequel le rayonnement ionisant et le gaz d'Hydrogène sont couplés. CoDa II reproduit certaines des contraintes principales sur la Réionisation (en particulier sur la fonction de luminosité ultra-violette des galaxies), témoignant que la Réionisation se déroule de façon réaliste dans CoDa II. Mon analyse m'a révelé que dans CoDa II la Réionisation est principalement alimentée en photons par les galaxies appartenant à une gamme intermédaire de masses ($6 \times 10^{8}\,\mathrm{M}_{\odot}$ to $3 \times 10^{10}\,\mathrm{M}_{\odot}$). En effet, les galaxies de plus faibles masses sont trop peu lumineuses car la formation stellaire y est peu efficace, tandis que les galaxies plus massives sont trop rares et comprennent plus de gas dense et neutre qui absorbe une grande partie de la lumière ionisante qui y est produite. Plus précisément, j'ai montré qu'en moyenne plus une galaxie est massive, plus une part importante de la lumière ionisante qu'elle produit est absorbé localement par le gaz d'Hydrogène neutre qu'elle abrite. Cette tendance découle du fait que la recombinaison des atomes d'Hydrogène ionisés est plus efficace dans le gaz très dense des galaxies massives, et ce en dépit de la formation stellaire plus forte qui s'y opère. La décroissance de f_{esc} avec la masse a aussi été signalée dans plusieurs travaux [tels que Yajima et al., 2011; Kimm and Cen, 2014; Katz et al., 2018], mais jamais avec la richesse statistique qu'offre le jeu de données CoDa II, et souvent dans des simulations où la Réionisation se déroule de façon moins réaliste. Bien que cette tendance soit très claire en moyenne, les données sont fortement dispersées. J'ai mis en évidence une corrélation entre cette dispersion et la distribution de gaz dans les galaxies (pour des galaxies de même masses, plus le maximum de densité de gaz est grand, moins la fraction de lumière ionisante produite qui parvient à s'échapper de la galaxie est grande). Finalement, j'ai montré qu'au fur et à mesure de l'avancée de la Réionisation, les galaxies massives et lumineuses deviennent de plus en plus importantes pour la Réionisation. Ce gain d'importance découle simplement de la croissance de leur population, et de la suppression de la formation stellaire dans les galaxies de plus faibles masses et luminosités.

Cependant, l'accord entre CoDa II et les contraintes sur la Réionisation est loin d'être parfait. Par exemple, le J21 après la Réionisation est trop élevé vis à vis des contraintes observationnelles, et des observations factices de la forêt Lyα ont montré qu'en moyenne la fraction de gaz neutre après la Réionisation était trop faible. En fait Madau [2017]; Finkelstein et al. [2019]; Kulkarni et al. [2019]; Keating et al. [2019] mettent en avant scénarios de Réioinsation tardive, qu'ils montrent être nécessaires afin de bien reproduire les contraintes sur l'évolution de la neutralité du gaz d'Hydrogène inter-galactique. En même temps, des processus physiques importants pour la Réionisation, comme le refroidissement des gaz dû à la présence d'Hélium et de métaux, ou l'absorption du photons ionisants par la poussière font défaut dans CoDa II. Tous ces éléments expliquent peut-être le nombre excessif de photons ionisants et d'ionisation dans CoDa II. Alors, pour préparer CoDa III, et pour chercher à améliorer les résultats tirés de CoDa II, j'ai travaillé à l'inclusion de nouveaux modèles physiques dans le code de simulation RAMSES-CUDATON, ainsi qu'à la réalisation de nouvelles simulations avec des Réionisation tardives.

CoDa 2.5

Premièrement, j'ai inclus à RAMSES-CUDATON une version adaptée du refroidissement dû aux métaux figurant dans [Rosdahl et al., 2013], ainsi qu'un modèle simple rendant compte du refroidissement dû à l'Helium. J'ai montré que les deux mécanismes de refroidissement avaient un rôle important dans les simulations, car en permettant d'avantage de refroidissement dans le gas à très grande densité, elles engendrent plus de formation stellaire. De plus, j'ai cherché à implémenter des changements qui pourraient pallier à l'excès de photons ionisants à la fin de la Réionisation dans CoDa II. Pour ce faire, j'ai modifié les lois gouvernant l'émissivité ionisante des populations stellaires dans RAMSES-CUDATON, pour qu'elles prennent en compte l'enrichissement en métaux et le vieillissement des populations. Cette modification réduit légèrement le nombre moyen de photons ionisants émis par masse stellaire, et retarde la Réionisation. En même temps, j'ai couplé un modèle physiquement motivé de la poussière au code, et ai implémenté l'absorption des photons ionisants par la poussière. J'ai abouti à un étalonnage du modèle de poussière qui reproduit raisonnablement bien les masses de poussière observées à haut redshift, ainsi que des prédictions de modèles semi-analytiques, ce qui justifie mon approche. Avec cet étalonnage crédible, j'ai montré que la poussière a un effet important sur l'échappement des photons ionisants depuis les galaxies et vers le milieu inter-galactique, pouvant la dimunuer de plus de 90 % dans les galaxies les plus massives qui ont le plus de poussière. Par ailleurs, la poussière est présente de manière suffisamment dense pour produire une extinction conséquente de la fonction de luminosité ultra-violette des galaxies brillantes (M_{AB1600}<-19), même à des redshifts élevés [comme ce qui a été mis en avant par Wu et al., 2020].

Deuxièmement, j'ai cherché à rendre la Réionisation tardive dans nos simulations, tout en diminuant la densité de photons ionisants à la fin de la Réionisation. J'ai trouvé qu'en incluant un critère en température pour la formation stellaire, je pouvais produire des simulations qui reproduisaient bien mieux les contraintes sur la forêt Lyα vers la fin de la Réionisation. Mais encore, dans la simulation CoDa 2.5, j'ai montré que le taux d'ionisation Γ était plus fidèle aux observations, tout en préservant une fonction de luminosité ultra-violette réaliste. De manière intéressante, ces changements produisent une émissivité totale ionisante qui décroît avec le temps après z=6, ce qui nous rapproche des attentes de Madau [2017]; Finkelstein et al. [2019]; Kulkarni et al. [2019]; Keating et al. [2019]. Cette tendance décroissante avec le redshift semble être déterminée par la suppression de la formation stellaire dans les galaxies de faibles masses. Les nouveaux processus physiques et la nouvelle calibration, donnent lieu à un budget galactique de photons ionisants semblable à celui de CoDa II, mais où les galaxies peu massives et peu brillantes jouent un rôle plus important. D'une part celà est dû à la fraction des photons qui parviennent à s'échapper vers le milieu intergalactique, qui décroit plus rapidement avec la masse des galaxies que dans CoDa II. Cette différence naît du critère de température pour la formation stellaire qui restreint la formation stellaire aux cellules chaudes, ionisées et transparentes aux photons ionisants. D'autre part, la présence de poussières réduit aussi la contribution des galaxies massives et brillantes. En effet dans CoDa 2.5, une fraction de moins de 10^{-3} des photons ionisants produits dans les galaxies de plus de $10^{11}\,M_{\odot}$ parvient à s'échapper jusqu'au milieu inter-galactique.

Vers CoDa III

Tout ceci étant dit, il reste des préparations à faire avant CoDa III. Par exemple, la calibration peut encore être améliorée du point de vue de l'accord avec les contraintes sur la forêt Lyα à z>5.2 (dans CoDa 2.5, le milieu intergalactique est trop neutre avant z=5.2), et de l'accord avec la distibution d'opacités Lyα. En même temps, le réalisme de la fonction de luminosité ultra-violette des galaxies pourrait être améliorée : en effet, à $z \leq 6$ il y a une bosse près de $M_{AB1600} \approx$-19, qui est absente des observations et des données CoDa II (j'explore actuellement l'origine de cette différence). Par ailleurs, il-y-a encore des tests importants du modèle de poussière et de son analyse par post-traitement à effectuer. Notamment, nous n'avons pas encore examiné l'effet des hypothèses quant à la section efficace par masse des grains de poussière sur l'absorption des photons ionisants, l'extinction, ou le rougissement de la pente du spectre ultra-violet des galaxies (β). Dans le cas de β, cela pourrait améliorer l'accord avec les observations tirées de Finkelstein et al. [2012a]. De plus, nous n'avons pas encore sufisemment contrôlé l'évolution du comportement du modèle de poussière avec la résolution, ce qui est crucial car CoDa III bénéficiera d'une plus grande résolution que CoDa 2.5 et que CoDa II.

Au-delà de CoDa III

Si l'on considère qu'il nous sera possible de reproduire et d'améliorer notre accord avec les données issues de la forêt Lyα , la fonction de luminosité ultra-violette des galaxies, l'émissivité ionisante Γ, tout en itérant sur le travail effectué sur les poussières dans CoDa 2.5, alors le jeu de données produit par CoDa III sera très riche. En effet, l'inclusion de nouveaux modèles physiques et l'amélioration de l'accord avec les observations feront qu'il sera très intéressant d'étudier le budget de photons ionisants des galaxies, ainsi que la fraction d'échappement des photons ionisants des galaxies dans CoDa III. Au-delà, la comparaison avec les résultats CoDa II fournira des informations importantes sur l'effet des changements opérés sur le code et sa configuration. Mais encore, cela permettra par exemple d'étudier la Réionisation du Groupe Local dans le contexte d'une Réionisation tardive.

Plus loin encore, nous pouvons nous attendre à bien des avancées. Du côté des simulations, des progrès logiciels (surtout) et matériels devraient un jour nous permettre de faire le lien entre des petites boîtes très bien résolues comme SPHINX [Rosdahl et al., 2018], et des très grandes boîtes comme CoDa II, ce qui promet un grand bond en avant dans le réalisme des simulations de la Réionisation. Du côté des observations, nous attendons à la fois des améliorations de contraintes existantes, ainsi que de nouvelles constraintes observationnelles innovantes. En particulier, le télescope spatial James Webb étendra la limite d'observation de galaxies à haut redshift d'environ deux en magnitude. En même temps, il fournira de nouveaux renseignements sur les propriétés des galaxies à haut redshift. Par exemple, il sera probablement capable de fournir des contraintes sur la métallicité du gaz de ces galaxies, crucial pour comprendre la rétroaction des supernovas et l'enrichissement chimique pendant la Réioinsation. Pendant ce temps, l'astronomie 21 cm promet une révolution de la science de la Réionisation. En effet, des expériences comme EDGES ont déjà remis en question notre compréhension de l'état du gas cosmique avant la Réioinsation, et des expériences futures nous donneront accès à des tomographies du gaz neutre en train de réioniser dans le milieu intergalactique. L'exploitation des données issues de ces nouvelles expériences radio requerra beaucoup d'interprétation, motivant encore plus la réalisation de modèles et de simulations.

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Acronyms and diminutive terms

AGN	Active Galactic Nuclei
AMR	Adaptive Mesh Refinement
BB	Big Bang
BH	Black Hole
CGM	Circum-Galactic Matter
CoDa	Cosmic Dawn
CPU	Central Processing Unit
DM	Dark Matter
DMF	Dust Mass Function
DTG	Dust To Gas ratio
DTM	Dust To Metal ratio
EoR	Epoch of Reionization
GMC	Gaint Molecular Cloud
GPU	Graphics Processing Unit
GRB	Gamma Ray Burst
HST	Hubble Space Telescope
IGM	Inter-Galactic Matter
ISM	Inter-stellar Medium
IMF	Initial Mass Function
JWST	James Webb Space Telescope
LMC	Large Magellenic Cloud
Lyα	Lyman α
LyC	Lyman Continuum
MBH	Massive Black Hole
MW	Milky Way

RHD	Radiative Hydrodynamics
RT	Radiative Transfer
SAM	Semi-Analytical Model
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SFR	Star Formation Rate
SFRD	Star Formation Rate Density
SMC	Small Magellenic Cloud
SN	Supernova
SKA	Square Kilometer Array
UV	Ultra-Violet
UVBG	Ultra-Violet Background
UVLF	Ultra-Violet Luminosity Function
QSO	Quasi-Stellar-Object or Quasar

Joseph LEWIS

Qui a réionisé l’Univers ?

Résumé

La lumière ionisante des premières étoiles a alimenté la Réionisation du milieu intergalactique. Dû au lien entre l’avancée de la Réioinsation et la formation des premières structures, l’étude des galaxies qui réionisent l’Univers a une motivation cosmologique. J’ai cherché à comprendre quelles sont ces galaxies grâce à des simulations numériques. Dans la simulation Cosmic Dawn II, des galaxies de masses intermédiaires de 6x10^8 à 3x10^{10} masses solaires propulsent la Réionisation. L’importance de ces galaxies découle de leur nombre, formation stellaire et fractions d’échappement importants. Pour aller plus loin, j’ai implémenté de nouvelles physiques dans le code RAMSES-CUDATON, telles que la poussière, et le refroidissement dû aux métaux et à l’Hélium. J’ai aussi modifié la configuration du code pour mieux reproduire les contraintes sur la fraction neutre d’Hydrogène. Il en résulte une Réionisation tardive, qui dépend de galaxies de plus faible masses de 10^8 à 2x10^{10} masses solaires. La poussière participe à réduire le rôle des galaxies massives dans la Réionisation, et pourrait affecter les observations des galaxies les plus brillantes pendant la Réionisation. Ces résultats ont des implications importantes pour les observations de galaxies et du gaz réionisant. Les nouvelles physiques et calibration seront utilisées pour Cosmic Dawn III.

Mots clés : Réionisation, galaxies, budget des photons ionisants galactique, haut redshift, fraction d’échappement, poussière

Summary

Ionising light from the first stars drove the Epoch of Reionisation (EoR) of the intergalactic medium. As the progression of the EoR is linked to the formation of the first structures, understanding which galactic sources drive it is cosmologically interesting. I investigated the EoR by studying its galactic sources across time in numerical simulations. I explored which galaxies propel the EoR in Cosmic Dawn II, finding the culprits to be an intermediate mass range of galaxies between 6x10^8 to 3x10^{10} solar masses. Their importance is due to its relatively high abundance, high star formation rate, and high escape fractions. To take this answer further, I sought to implement new physics into the code RAMSES-CUDATON (Most importantly, dust, Helium and metal cooling). I also altered the simulation setup to better match constraints on the neutral fraction of Hydrogen. I found that this new configuration supports late Reionisation driven by lower mass galaxies of 10^8 to 2x10^{10} solar masses. Dust plays a rôle in reducing the contribution to Reionization of massive galaxies, and could affect observations due to extinction in the brightest galaxies. These results have important implications for observations of reionising galaxies and gas. The new physics and setup with be used in Cosmic Dawn III.

Key words : Reionisation, galaxies, ionising galactic photon budget, high redshift, escape fraction, dust