Metallic UV absorption lines as diagnostics of the escape of ionising photons from simulated galaxies

Valentin Mauerhofer

N°d'ordre NNT : 2021LYSE1238

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 52
Physique et Astrophysique de Lyon (PHAST)

Discipline : Astrophysique

Soutenue publiquement le 16/11/2021, par :
Valentin Mauerhofer

Les raies d'absorption métalliques UV comme diagnostiques d'échappement des photons ionisants de galaxies simulées

Devant le jury composé de :

Bolmont, Emeline, Professeure Université de Genève	Présidente
Scarlata, Claudia Professeure University of Minnesota	Rapporteure
Wise, John Professeur Georgia Institute of Technology	Rapporteur
Prochaska, Jason Professeur University of California Santa Cruz	Examinateur
Teyssier, Romain Professeur University of Zurich	Examinateur
Richard, Johan Astronome adjoint Université Lyon 1	Examinateur
Verhamme, Anne Professeure Université de Genève	Directrice de thèse
Blaizot, Jérémy Astronome Université Lyon 1	Directeur de thèse

Abstract

With the arrival of the James Webb Space Telescope, it is high time to improve our understanding of high redshift ultraviolet (UV) spectra. Starting from modern high resolution radiation-hydrodynamic simulations, I develop post-processing pipelines to produce accurate mock spectra from virtual galaxies. This allows us to gain insights into the interpretation of the spectra, to assess the realism of the simulations and to make predictions for future observations.

One open question in cosmology is the exact origin of cosmic Reionisation. Around redshift 6, or about one billion years after the Big Bang, the state of the hydrogen in the Intergalactic Medium (IGM) went from neutral to ionised, which constitutes the last phase transition of the Universe. The precise origin of this process is still debated, although the current consensus is that the ionising photons emitted by the young massive stars in early galaxies are responsible for Reionisation. However, most of the ionising photons, also called Lyman continuum (LyC), are absorbed by the Interstellar Medium (ISM) of the galaxies from which they are emitted. Therefore, to better understand which type of galaxy contributed the most to Reionisation, one must know how to measure the escape fraction of ionising photons ($f_{\rm esc}$) from high redshift galaxies. Since the ionising photons are absorbed by the IGM during Reionisation, they cannot be observed. Therefore, the measure of $f_{\rm esc}$ has to be done with indirect tracers.

The down-the-barrel absorption lines of low ionisation state (LIS) ions are promising tracers of $f_{\rm esc}$. They are absorption lines imprinted on the stellar continuum of galaxies by metallic ions such as C^+ or Si^+. Thanks to their wavelengths longer than the wavelength of the Lyman alpha transition of hydrogen, those absorption lines are not erased by the neutral hydrogen in the IGM. Furthermore, since the lines are from low ionisation elements, they trace the neutral medium of the ISM, which itself shapes $f_{\rm esc}$. To study how well we can infer $f_{\rm esc}$ from absorption lines, I develop and exploit a unique post-processing pipeline to produce down-the-barrel LIS absorption line mock observations and to compute the escape fractions in our simulated galaxies. This allows me to compare the observables with $f_{\rm esc}$, and to investigate and quantify correlations and dispersion between them.

After presenting the simulations that I am using in this work and the different steps of my post-processing pipeline, I begin by studying absorption lines themselves. I show that I reproduce realistic lines, which look like observed high redshift absorption lines. Then, I show that many complex physical processes affect the absorption profiles. The infilling effect, due to the scattering of line photons, changes the depth of the lines, and the large-scale velocity gradients of the gas in galaxies modify the lines significantly, due to the Doppler effect. Additionally, inhomogeneous dust screens also play an important role in the shaping of absorption lines. Then, I compare the properties of the lines with the escape fractions of ionising photons in selected directions of observations. Due to the complex processes affecting the lines, and to the imperfect co-location of LIS ions and neutral hydrogen, I find relations with a large scatter, and the absorption lines cannot in general trace the escape fraction of ionising photons.

To investigate further the origin of the dispersion in the relations between escape fraction and absorption lines properties, I first analyse the complex but central role of dust in shaping the UV outputs of galaxies, and the impact of our dust modelling. By comparing observables such as the luminosity function, the β slope and the infrared excess between simulations and observations, I assess how realistic our simulations are. I also test the observational methods used to compute the dust attenuation of high redshift galaxies. By using them on my mock spectra, and comparing with the exact dust attenuation known from the simulation, I deduce that the extinctions measured in observations are usually underestimated. Finally, I explore the link between absorption lines properties and gas kinematics. I compare absorption lines with mass outflow and inflow rates in the Circumgalactic Medium (CGM) of our simulations, to study how well they are correlated.

Résumé

Avec l'arrivée du télescope spatial JWST, il est grand temps d'améliorer notre compréhension des spectres ultraviolets à haut décalage vers le rouge (redshift). À partir de simulations hydrodynamiques à haute résolution, je développe des codes de post-traitement pour produire des spectres simulés précis de galaxies virtuelles. Cela permet de mieux interpréter les spectres observés, d'évaluer le réalisme des simulations et de faire des prédictions pour les observations futures.

Une question ouverte en cosmologie est l'origine exacte de la Réionisation cosmique. Autour du redshift 6, soit environ un milliard d'années après le Big Bang, l'état de l'hydrogène dans le milieu intergalactique (IGM) est passé de neutre à ionisé, ce qui caractérise la dernière transition de phase de l'Univers. L'explication précise de ce processus est encore débattue, bien qu'il semble que les photons ionisants émis par les jeunes étoiles massives dans les premières galaxies soient responsables de la Réionisation. Cependant, la plupart des photons ionisants sont absorbés par le milieu interstellaire (ISM) des galaxies à partir desquelles ils sont émis. Par conséquent, pour mieux comprendre quel type de galaxie a le plus contribué à la Réionisation, il faut savoir mesurer la fraction d'échappement des photons ionisants ($f_{\rm esc}$) des galaxies à hauts redshifts. Les photons ionisants étant absorbés par l'IGM lors de la Réionisation, ils ne peuvent pas être observés depuis la Terre. Par conséquent, la mesure de $f_{\rm esc}$ doit être faite avec des traceurs indirects.

Les raies d'absorption des ions à faible ionisation sont des traceurs prometteurs de $f_{\rm esc}$. Ce sont des raies d'absorption créées par des ions métalliques tels que C^+ ou Si^+, apparaissant sur le continu stellaire des galaxies. Grâce à leur longueur d'onde plus longue que celle de la transition Lyman alpha de l'hydrogène, ces raies d'absorption ne sont pas effacées par l'IGM neutre de l'époque de la Réionisation. En outre, comme les raies viennent d'éléments à faible ionisation, elles tracent le milieu neutre de l'ISM, qui lui-même détermine $f_{\rm esc}$. Pour étudier dans quelle mesure nous pouvons déduire $f_{\rm esc}$ des raies d'absorption, j'utilise mes codes de post-traitement uniques en leur genre pour produire des raies d'absorption et calculer $f_{\rm esc}$ dans nos galaxies simulées. Cela me permet de comparer les observables avec la fraction d'échappement, et d'analyser et quantifier les corrélations entre ces quantités.

Après avoir présenté les simulations que j'utilise dans ce travail et les différentes étapes de mes codes de post-traitement, je commence par étudier les raies d'absorption elles-mêmes. Je montre que je reproduis des raies réalistes, ressemblant à des raies d'absorption de galaxies observées à redshift élevé. Ensuite, je montre que de nombreux processus physiques complexes affectent les profils d'absorption. L'effet de remplissage, dû à la diffusion des photons, modifie la profondeur des raies, et les gradients de vitesse du gaz dans les galaxies modifient considérablement les raies, en raison de l'effet Doppler. De plus, les écrans de poussière non homogènes jouent également un rôle important dans le façonnage des raies d'absorption. Ensuite, je compare les raies avec les fractions d'échappement des photons ionisants dans des directions d'observations choisies. En raison des processus complexes affectant les raies, et des différences entre la distribution des ions et celle de l'hydrogène neutre, je ne trouve que des relations dispersées, et les raies d'absorption ne peuvent en général pas prédire la fraction d'échappement des photons ionisants.

Pour étudier l'origine de la dispersion dans les relations entre les fractions d'échappement et les raies d'absorption, j'analyse d'abord le rôle complexe mais central de la poussière dans la détermination du spectre ultraviolet. En comparant les observables tels que la fonction de luminosité, de la pente β et l'excès infrarouge entre les simulations et les observations, j'évalue à quel point nos simulations sont réalistes. Je teste également les méthodes d'observation utilisées pour calculer l'atténuation par la poussière des galaxies à redshift élevé. En utilisant ces méthodes sur mes spectres simulés, et en comparant les résultats avec l'atténuation par la poussière exacte, connue dans la simulation, je déduis que les mesures sous-estiment la vraie atténuation. Enfin, j'explore le lien entre les propriétés des raies d'absorption et la cinématique du gaz. En particuliers, je compare les raies d'absorption avec les débits de masse de gaz sortants et entrants dans le milieu circumgalactique (CGM) de nos simulations.

Acknowledgements

These four years of PhD have been extremely interesting and the occasion to meet so many people and visit many different places. I am ever grateful to Anne Verhamme, who chose me as her first PhD student and allowed me to experience all of this. She also gave me the opportunity to work jointly with the University of Lyon. Thank you to Jérémy Blaizot for welcoming me there, and guiding me through this thesis, always pushing me to improve things. Both my supervisors have been infinitely helpful and patient, even during the tough times of the pandemic, we continued skyping every week and progressing in our work.

I would like to thank all my colleagues, the atmosphere at both observatories is really nice thanks to you. A special thanks to Josie, David and previously José for being faithful ping-pong partners in Geneva, as well as Adélaïde in Lyon, with whom I also have done nice visits around Lyon, Porto and even Crete. I have had several office mates, one nicer than the other, Andreas, Stéphane, Floriane, Alberto and recently Andrea.

I received a lot of help during those three years. Besides my supervisors, I am grateful to Joki, who always quickly answered all my questions on slack, and helped me install many programs and shortcuts.

Thank you also to Ryan Trainor, who spent a lot of time this year with my supervisors and me to discuss the topics of gas flows and of dust attenuation. His collaboration brought knowledge from the observational side that was very valuable for my work.

This PhD thesis gave me the occasion to visit many countries and places, thank you to Anne, the ERC fund and the observatory of Geneva for making this possible. Above all, the opportunity to go to Chile and to observe during 15 nights in the Swiss Telescope at La Silla resulted in an unforgettable experience.

I am especially grateful to Josie, who spent many hours helping me improve my paper and this thesis. Thank you so much for helping me with English and with the scientific content.

Finally, thank you to my parents, to which I owe everything. Thank you for supporting me, being proud of me, and always giving me delicious food. Merci beaucoup !

This work was funded by the ERC Starting Grant 757258 ”TRIPLE”.

Softwares and Codes used in this thesis

- This thesis was written with LaTeX, https://www.latex-project.org/.
- Python 3.7, https://www.python.org/
- Fortran90, https://www.fortran90.org
- Rascas, Michel-Dansac et al., 2020, http://rascas.univ-lyon1.fr
- Ramses-RT, Rosdahl et al., 2013
- Krome, Grassi et al., 2014, http://kromepackage.org
- Cloudy, Ferland et al., 2017, https://gitlab.nublado.org/cloudy/cloudy/-/wikis/home
- Numpy and Scipy
- Matplotlib, Hunter, 2007
- Astropy, Astropy Collaboration et al., 2013
- PyNeb, Luridiana et al., 2015, http://research.iac.es/proyecto/PyNeb//

Contents

1 Introduction

1.1 A brief introduction to Cosmology

The study of galaxies beyond our Milky Way started in the 1920's, when telescopes became large enough to observe distant nebulae. There were heated debates about their exact distances, to determine whether they were within our Galaxy, which was considered the extent of the whole Universe, or if those nebulae were outside our galaxy and in fact other galaxies in their own rights, or "island universes". Edwin Hubble settled the debate in 1925 (Hubble, 1925) when he determined the distance of the nebula NGC 6822 using Cepheid variable stars. Those are stars whose luminosity varies in time and is proportional to the period of variation, allowing to deduce the distance of the star by its brightness. NGC 6822 contained several of those stars, which Hubble found, and which indicated a distance of 660'000 light years, too far to be inside the Milky Way. On this day the Universe suddenly became much larger.

Soon after, Hubble measured distances to more galaxies. Spectroscopy, the study of the distribution of light as a function of its wavelength, also allows to measure the velocity of objects, since the typical spectral features, absorption and emission lines, change wavelength when there is a difference of velocity between the source and the observer (this is called the Doppler effect, the light is said to be redshifted when the wavelength increases, and blueshifted when it decreases). He concluded that the more distant a galaxy is, the faster it moves away from us, except for our closest neighbour, the Andromeda galaxy, which moves towards us due to gravity. This is the first proof of the expansion of the Universe, which was already theorised by Georges Lemaître (Lemaître, 1927). The proportionality between the distance and the velocity of galaxies is now called the Hubble–Lemaître law.

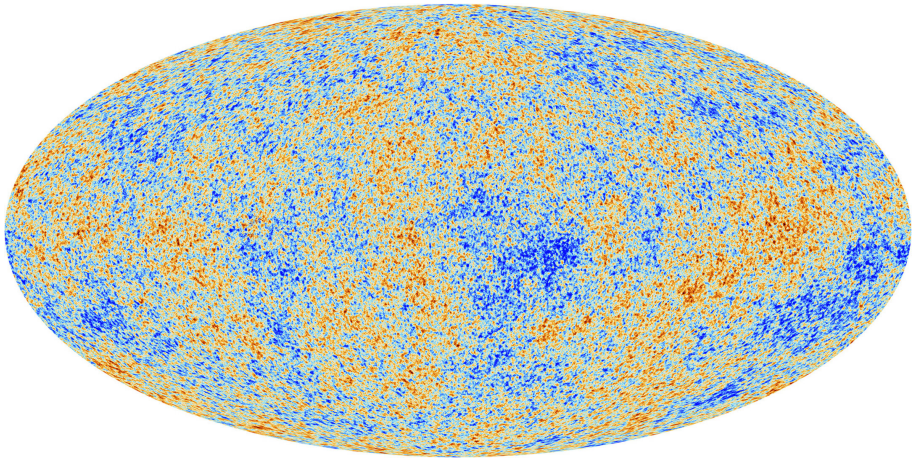

Figure 1.1: Cosmic microwave background observed by the Planck satellite (Planck Collaboration et al., 2020). The colours indicate the temperature fluctuations in the microwave background radiation, around an average of 2.72548 K. There is a difference of 0.00057 K between the blue and the red regions.

The concept of galaxies moving away from us gave rise to the idea that in the past they must have been closer together. Ultimately, this leads to the Big Bang Theory, stating that the Universe is not eternal but was born extremely compact and then expanded into what we see today. Although there were serious doubts for several decades, a surprising observation revealed the presence of a uniform and isotropic microwave background in the whole sky (Penzias & Wilson, 1965), now called the Cosmic Microwave Background (CMB). This fits well with the Big Bang Theory, which predicts that the Universe began as a hot plasma of ionised hydrogen and helium atoms, and that after cooling due its expansion, the plasma recombines, i.e., the free electrons get attached to the nuclei, letting photons that were trapped in the primordial plasma travel freely, which are now redshifted to the microwave range. To this day, the CMB is still our primary source of information in Cosmology. Its details are the best evidence for our current Standard Model of Cosmology, stating that normal matter only makes up 5% of the total energy budget of the Cosmos, while dark matter, interacting via gravity but not via any other forces, makes up 25% and dark energy, a concept created to explain the apparent accelerated expansion of the Universe, makes up the remaining 70%. According to this model, our Universe is around 13.6 billion years old. The most recent observations of the CMB are from the Planck satellite which built a high resolution map shown in Figure 1.1.

Around the same year as the discovery of the CMB, observers began to detect galaxies at extreme distances of more than 2.6 billion light years (e.g. Humason et al., 1956). At those scales, the notion of distance becomes impractical, since the expansion of the Universe changes the basic physical relations between distance, flux, luminosity, size and travel time. In astronomy, we often use the concept of "redshifts" instead of distances. Since far away objects are moving away at a velocity proportional to the distance, and the Doppler effect shifts the wavelength of light of objects receding from us, the amount of which the wavelength is displaced is a direct indicator of the distance. We call this redshift z and the observed wavelength $\lambda_{\rm obs}$ is related to the intrinsic wavelength λ by $\lambda_{\rm obs} = (1+z)\lambda$. In the 1950s, the galaxies with the highest detected redshifts were at $z = 0.2$ (e.g. Hubble, 1953). Nowadays, cosmologists tend to call this "very low redshift" or "the local Universe", even though it is a distance so great that in a sphere with this distance as a radius, there are around ten million galaxies at least as large as the Milky Way, and at least a hundred million smaller galaxies (Conselice et al., 2016). Figure 1.2 gives an impression of the different distances between the furthest point that a man made object reached and the distance of redshift 0.2 galaxies. The small black segments represent the distance between the objects written above them in bold and the long arrows are the corresponding distances between the objects below, illustrated with images. The green lines are separating the different steps. The jump from the distance of Voyager 1 to the distance of the closest star Proxima Centauri is already hard to grasp, but the one from there to the size of our Milky Way is so enormous that it cannot be represented in this kind of figure. In comparison, the distances between galaxies of the same group, or even to neighbouring groups, is not so large. However, redshift 0.2 is more than a thousand times further than our neighbour Andromeda. In this thesis I study galaxies that are even much further away than redshift 0.2.

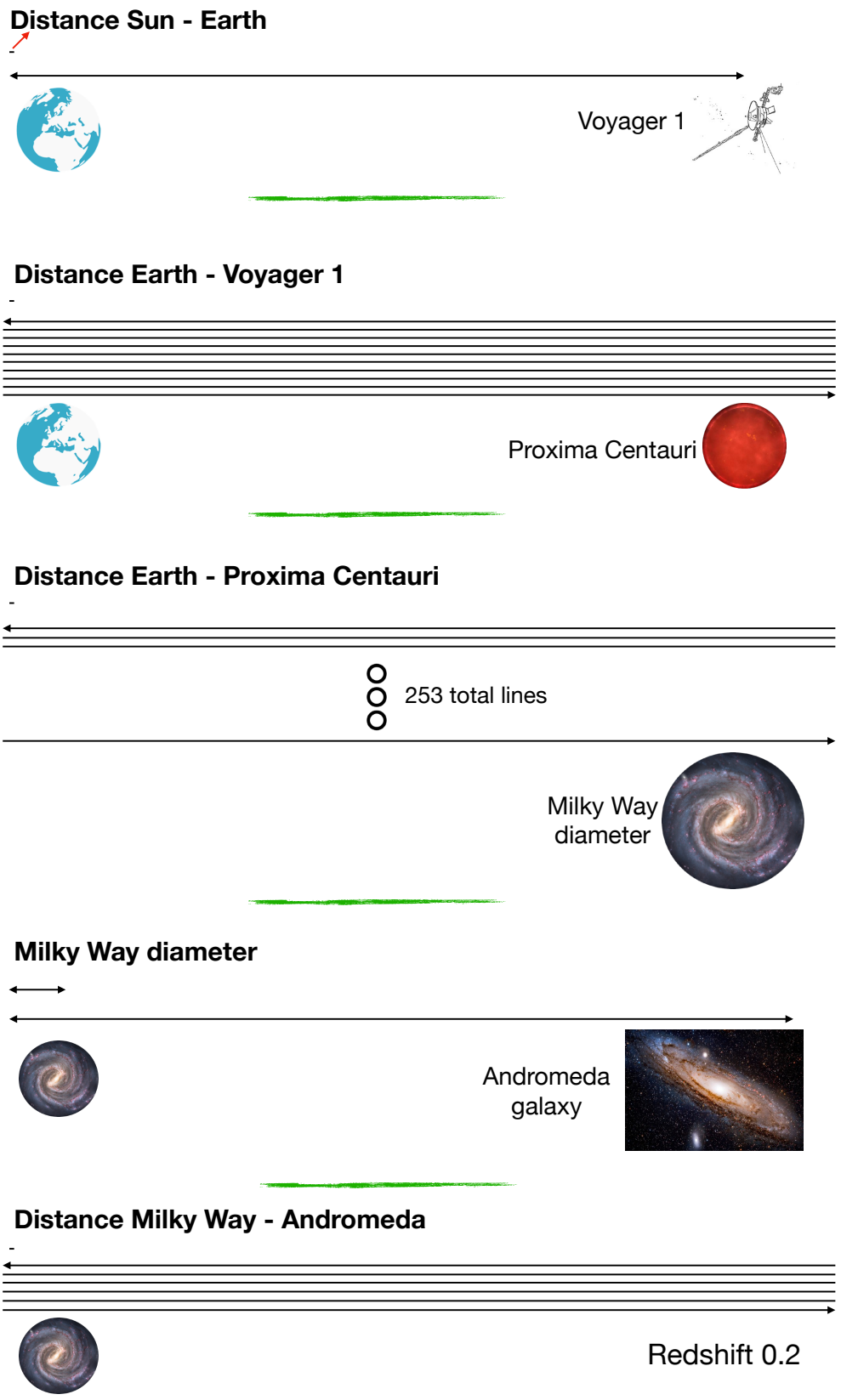

Figure 1.2: Comparison of distances, from the Sun to the Earth, then to the furthest distance that a human object reached: the Voyager 1 probe, to the star Proxima Centauri, to the diameter of the Milky Way galaxy, to the Andromeda galaxy and to redshift 0.2.

Another tool to study cosmology besides the CMB are Quasi Stellar Objects (quasars), point-like sources of light that are among the most luminous objects in the Universe. They are disks of gas spiraling around supermassive black holes at the center of galaxies, heated to such high temperatures that they emit intense radiation. Only the most matter-absorbing supermassive black holes are quasars. We can see some of them up to distances where the Universe was only 600 million years old when their light was emitted (Pentericci et al., 2002). Becker et al. (2001), and later Fan et al. (2006), discovered that for quasars above a threshold distance the photons below 1215 Å are completely absorbed before reaching us. This is the wavelength of the Lyman alpha transition of hydrogen, the electronic transition between levels 1 and 2 of the hydrogen atom. This is strong evidence that the Intergalactic Medium (IGM) underwent a phase transition at this time. From around 370'000 years to around one billion years after the Big Bang, the hydrogen atoms were neutral, thus absorbing all photons below 1215 Å. After this, the Universe got ionized, all the electrons of hydrogen atoms were kicked out of orbit, leaving a plasma that stopped absorbing ultraviolet photons. This transition is called Reionisation.

1.2 The sources of Reionisation

Madau & Dickinson (2014) find that in order for galaxies to ionise the IGM in the first billion years after the Big Bang, there must have been at least two ionising photons emitted for each hydrogen atom of the Universe. One is not enough because electrons can recombine with protons and reform neutral hydrogen. Careful observations of high redshift galaxies and models of stellar evolution can determine the star-formation rate (SFR) of those galaxies and how many ionising photons are emitted per unit star-formation rate ($\xi_{\rm ion}$). This is not enough, though, to know how many photons reach the IGM. One quantity is missing, which is the escape fraction of ionising photons, $f_{\rm esc}$. Many ionising photons are absorbed by the dense neutral gas inside galaxies, in the ISM. Those photons do not contribute to Reionisation, hence it is crucial to know what fraction escapes into the IGM. The total contribution of all galaxies to the ionising emission is the sum of the product $\mathrm{SFR} \times \xi_{\rm ion} \times f_{\rm esc}$ for all galaxies. Given our knowledge of SFR and $\xi_{\rm ion}$, galaxies must have an average escape fraction of around $10 - 20\%$ to explain the ionising emission needed for the Epoch of Reionisation (e.g. Robertson et al., 2015 and references therein).

The escape fraction is hard to determine, and the uncertainties around this quantity result in open questions surrounding the Reionisation history. Robertson et al. (2015) or Duncan & Conselice (2015) favor a scenario where small galaxies with a UV magnitude between -10 and -15 Mag are the main sources for Reionisation. Recently, Naidu et al. (2020) argued that, on the contrary, models where the majority of ionising photons come from massive galaxies with stellar masses above 10^8 $\mathrm{M}_\odot$ can match with all observational constraints and explain Reionisation. Additionally, the role of Active Galactic Nuclei (AGN, less extreme versions of quasars) is also uncertain. While it was assumed that they contribute significantly to the budget of ionising photons (Volonteri & Gnedin, 2009; Madau & Haardt, 2015), more recent works tend to conclude that their contribution is negligible (Hassan et al., 2018; Parsa et al., 2018; Trebitsch et al., 2020). Nevertheless, there might be a population of faint or obscured AGNs that are not yet detected but contributing to Reionisation (e.g. Giallongo et al., 2019).

To answer those questions we must learn to better constrain the escape fractions of high redshift galaxies. The best way to learn about properties of individual galaxies is to study their spectrum.

1.3 Spectroscopy of star-forming galaxies

The main source of information about galaxies is their spectrum. Since nothing travels faster than light and it is relatively little transformed during its journey through the IGM, it is the best way to better understand galaxies. The spectrum of a galaxy is composed of a continuum inter-

spersed emission and absorption lines. The continuum is mainly emitted by stars, which shine bright at ultraviolet and optical wavelengths. The intrinsic continuum emission of stars is highly modified by the presence of dust grains in the Interstellar Medium (ISM). The distribution of dust determines the slope of the continuum. AGNs also emit continuum radiation, as well as the interstellar gas, through interactions between free electrons and hydrogen and helium atoms, called the nebular continuum. Emission lines are produced either by elements recombining after they were photoionised by energetic photons emitted by stars, or by elements de-exciting after collisional excitations in the ISM.

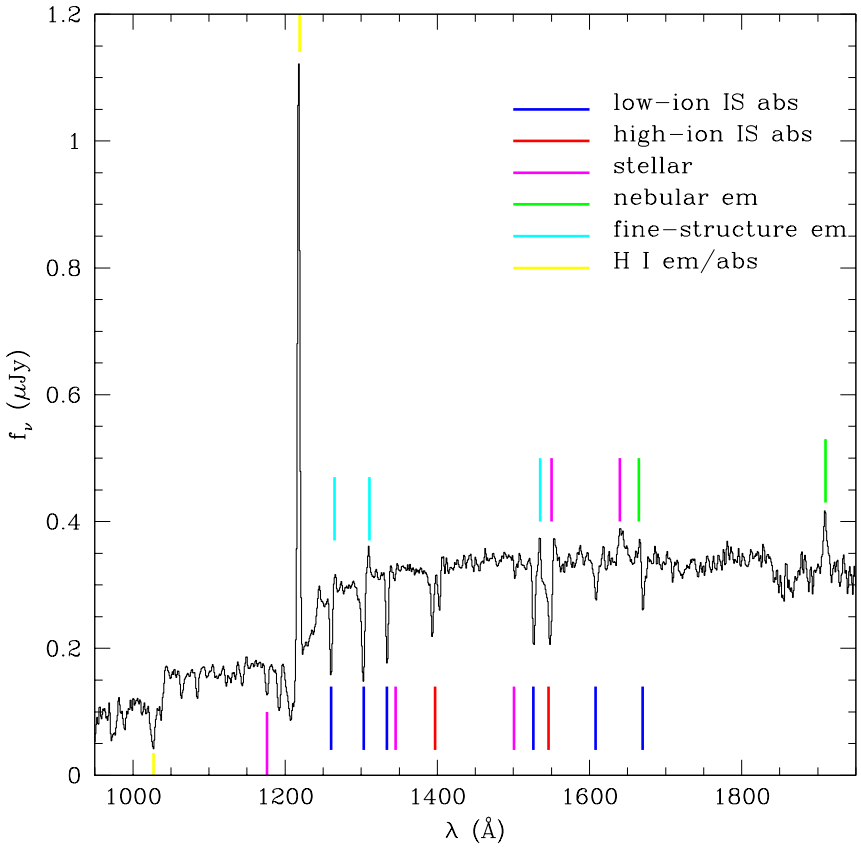

Figure 1.3: Composite UV spectrum of a sample of almost one thousand redshift 3 galaxies, from Shapley et al. (2003). Blue and red segments highlight metallic ISM absorption lines from low and high ionisation species, respectively. Purple segments highlight stellar absorption lines. Green segments highlight nebular emission lines. Cyan segments highlight fluorescent emissions from fine-structure levels (see Section 3.4. Finally, the yellow segment on the left highlights an H^0 absorption line.

As an illustration of a typical UV spectrum of a high redshift galaxy, I show in Figure 1.3 a composite spectrum of almost one thousand $z \sim 3$ galaxies from Shapley et al. (2003). The most prominent line by far is the Lyman alpha line of hydrogen at 1215 Å. Its high luminosity is due to the fact that hydrogen is the most abundant element and that Lyman alpha is the transition between the two lowest energy levels of the atom, which means that a large fraction of radiative cascades following recombinations or collisional excitation events end with this transition (e.g. Dijkstra, 2017). Since it is the brightest line, it is often the only emission line detected in spectra of high redshift galaxies. In this case it can be used to determine the redshift of the galaxy, but its resonant nature could lead to an inaccurate result. Another emission line visible in Figure 1.3 is C III λ1909, which is among the brightest UV nebular emission lines (emitted during radiative cas-

cades following recombination events). It is often used for more accurate redshift determinations, as it is not a resonant line and will not be scattered in neutral hydrogen. Emission lines are great sources of information about galaxies, they are used for example as a measure of star-formation rate, metallicity or dynamical properties. For a recent review on emission lines and the evolution of galaxies, see Kewley et al. (2019).

Numerous other lines can be seen in Figure 1.3, including stellar absorption lines in purple, and ISM absorption lines in blue and red. ISM absorption lines the latter being of particular interest in this thesis.

1.4 Absorption lines

ISM absorption lines of star-forming galaxies, also called "down-the-barrel" absorption lines (in contrast with quasar absorption lines), are formed when stellar light is absorbed by elements in the galaxy between the sources and the observer, creating troughs in the spectrum around the frequencies of electronic transitions of the different species present in the ISM and Circumgalactic Medium (CGM). They are often split into two categories: the low-ionisation state (LIS) lines, like Si II (λ1190, λ1193, λ1260, λ1304), O I λ1302 or C II λ1334, and the high-ionisation state (HIS) lines, like C IV $\lambda\lambda$1548, 1550 or Si IV $\lambda\lambda$1394, 1403. LIS lines are from elements that have ionisation potentials, between around 13.6 eV and 25 eV, while HIS lines have ionisation potentials, between around 45 and 65 eV. Another kind of line is stellar absorption lines, which are produced directly by the stars and not in the ISM, such as N V $\lambda\lambda$1239, 1243.

Absorption lines are challenging to detect, since they are by nature a decreasing of flux around given wavelengths, below the continuum. Therefore, a minimum requirement to observe absorption lines is to have a clear detection of the continuum. The fact that they are in the ultraviolet also complicates their detection. To observe absorption lines with ground based telescopes, galaxies must be at a sufficiently high redshift to shift the rest-frame ultraviolet into the optical wavelengths, which implies that those galaxies are far-away and thus faint. Alternatively, one can also use space-based telescopes to directly detect ultraviolet light, allowing to observe low redshift galaxies. Historically, the first detections of down-the-barrel absorption lines have been done from space, with the International Ultraviolet Explorer (IUE), launched in 1978. Kinney et al. (1993) present an atlas of 143 galaxies observed with the IUE, many of which with detections of absorption lines. The next revolutionary telescope in terms of UV observations was the Hubble Space Telescope (HST). One of the earliest papers presenting UV spectra including absorption lines taken with HST is Leitherer (1994), where they focus on HIS stellar absorption lines to deduce properties of massive stars. Observations of high redshift absorption lines became possible with the advent of large-scale ground based telescopes such as the Keck observatory and the Very Large Telescope (VLT). Early studies with stacks of spectra of $z \sim 3$ galaxies displaying absorption lines are for example Steidel et al. (1996a,b) or Shapley et al. (2003). More recent studies, with the same telescopes but new instruments, are for example Steidel et al. (2018) or Feltre et al. (2018, 2020) with the MUSE integral field spectrograph.

Heckman et al. (2001) is another study which is particularly relevant here, since it made the link between down-the-barrel absorption lines of galaxies and their escape fraction of ionising photons. They found that the ratio of the flux at the bottom of absorption lines to the continuum flux is a measure of the escape fraction, in particular for LIS lines such as C II λ1036. The main assumption for this relation, called the picket-fence model, concerns the geometry of the ISM. It states that a fraction of the stars are covered by an optically thick layer of neutral gas, while the rest is not covered at all. The percentage of covered stars is called the covering fraction. This situation is illustrated in the top part of Figure 1.4, with a covering fraction of 100% on the left and a partial covering on the right. Since C^+ or Si^+ have similar ionisation potential as neutral hydrogen, they occupy the same regions of the ISM. Thus, in the case of a total coverage by neutral gas, both

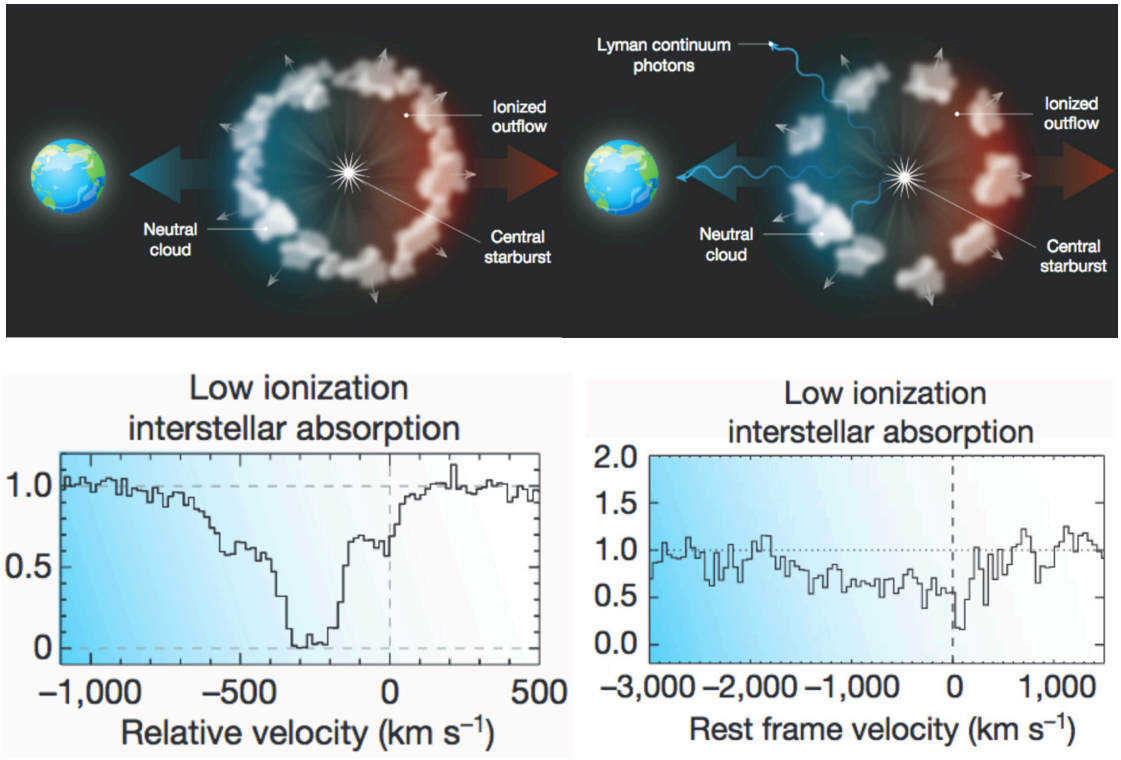

Figure 1.4: Illustration of the picket-fence model, from Erb (2015). The left side depicts a situation where a site of star-formation is covered by a screen of uniform neutral gas, resulting in a wide and saturated absorption line (bottom left panel). On the right side the sources are only partially covered, which allows for some flux to escape the system, resulting in a non-saturated absorption line (bottom right panel).

the ionising photons and the photons of LIS absorption lines cannot escape the system, since the sources are covered by both H^0 and LIS ions. This results in a saturated absorption line, as depicted at the bottom left of Figure 1.4, and the absence of leakage of Lyman continuum. Alternatively, if the covering fraction is less than unity, the LIS absorption line is not saturated, as in the bottom right part of Figure 1.4, and there is leakage of ionising photons. In the simplest form of the picket-fence model, the ratio (called the residual flux) of the flux at the bottom of LIS absorption lines over the flux of the continuum is set equal to one minus the covering fraction, which is itself set equal to one minus the escape fraction of ionising photons. More simply, the residual flux of LIS absorption lines is equal to the escape fraction of ionising photons. Recently, more sophisticated picket-fence models have been developed to include the effect of a non-homogeneous dust screen, for example in Reddy et al. (2016); Gazagnes et al. (2018, 2020) or Chisholm et al. (2018a). This cancels the equality between residual flux and escape fraction, but gives other equations to relate the two quantities.

To test the assumptions of the picket-fence model, such as the dichotomy between optically thick coverage and no coverage at all of the sources, or the collocation of neutral hydrogen and LIS ions, and more generally to study the link between absorption lines and the escape of ionising photons, one possibility is to use modern radiation hydrodynamics simulations.

1.5 Numerical simulations as a complementary approach to understanding observations

In astrophysics, simulations play the role of the laboratory. Indeed, stars and galaxies are much too far away and too large for us to interact with them and study their response. So, unlike most sciences, it is not possible to perform direct experiments in this field. One way to circumvent this limitation is to perform numerical simulations of parts of the Universe. This allows astronomers to test different models, for example by changing the cosmological parameters, the physics of stellar evolution or the models of interaction of different constituents of matter. While it is not in itself an experiment, it provides insights into the mechanisms and physics at play that can help to understand observations of distant objects such as galaxies. In this work I am interested in Reionisation and the formation of galaxies at high redshift, which is why I focus here on cosmological simulations of galaxy formation.

In my opinion, the best way to make the most of simulations, and to get the most insight into actual observations of high redshift galaxies, is to make faithful and accurate mock observations of the simulations. In extragalactic astrophysics, there are what I call the direct observables, such as the redshift, the luminosity, the slope of the UV continuum or properties of spectral lines such as their equivalent width, full width at half maximum, etc. With those observables, one is able to assess properties of the galaxies such as stellar mass, star-formation rate or dust attenuation, which I call indirect observables, as they have to be inferred from comparing observations to models. In many simulation studies, the comparison is made between the stellar mass, dark matter mass or star-formation rate of the simulated galaxies with the corresponding indirect observables for real galaxies (e.g. Katsianis et al., 2021). This can introduce several biases, since the method of deriving those quantities is different in observations and in simulations. Instead, in this work, I post-process simulations to make comparisons of direct observables, following e.g. Barrow et al. (2017, 2018).

However, only complex simulations including several physical ingredients can be used to make reliable mock observations. Indeed, mocks are done by simulating the radiative transfer of photons in simulations, including the interaction with gas and dust, so all those components have to be present in the simulation. Historically, cosmological simulations have begun by including only dark matter, the most abundant massive entity in the Universe. Gravity is the main force governing objects at astronomical scales, so dark matter is indeed the most important ingredient to include. The first decade of the 21st century saw the advent of large scale simulations of the formation of dark matter halos, from around 20 Myr after the Big Bang to nowadays, as in Bode et al. (2001) or Springel et al. (2005). This kind of simulation is essential to understand the formation of large scale structures, but it does not provide the detailed states of stars, gas and dust that are necessary to build mock observations. Simulations that include cosmological hydrogen and helium gas, which collapses to form stars and galaxies, are called hydrodynamics simulations. Hydrodynamics is a highly non-linear problem, which is computationally challenging to solve in the cosmological context. There are two common approaches to solve this problem, Eulerian and Lagrangian algorithms, which have a grid-based approach and a particle-based approach of treating the gas, respectively. Typical examples of such codes are Ramses (Teyssier, 2002) or Enzo (O'Shea et al., 2004) for the Eulerian approach and Gadget (Springel et al., 2001) for the Lagrangian approach. In both cases, stars are modelled by point-like particles of several hundred or thousand solar masses, similarly as for dark matter.

Since hydrodynamic simulations contain stars, one can build mock observations of the stellar continuum, provided that a model for dust attenuation is included. However, to include the absorption lines, or other spectral features like the Lyman alpha emission line, one must know the ionisation state of the gas in the simulated galaxies, which requires the inclusion of the transfer of ionising photons in the simulation. Hydrodynamics alone cannot accurately predict the ionisation state of the gas, since this state is highly influenced by the local radiation field. In this case only

some specific observables such as absorption lines of quasar spectra, created by the CGM of intervening galaxies, can be modelled, since the ionisation state of the CGM is mainly influenced by the ultraviolet background (UVB), which can be added in post-processing (e.g. Hummels et al., 2017). Obtaining other, more complex observables is one motivation to add an additional level of complexity in simulations: the transfer of ionising photons. For most people, the main goal of adding ionising photons to cosmological hydrodynamic simulations is to directly model the Reionisation of the Universe, but I would argue that it is as important to be able to create accurate mock observations, thanks to knowing the ionisation state of the gas. To illustrate all the layers of complexities that are in radiation hydrodynamic (RHD) simulation, I show in Figure 1.5 an image from Trebitsch et al. (2020) displaying the different components of a high resolution RHD simulation. One example of a radiation hydrodynamics simulation studying Reionisation is Ocvirk et al. (2016, 2020), where they simulate a large box size of around $(92\text{cMpc})^3$, ideal to understand the large scale dynamics of Reionisation. However, simulating such a large box forces the resolution to be coarse. For mock observations, very high resolution is needed in the ISM, below around 30 pc, to resolve as well as possible the sites of star-formation, and the complex movement of gas. Only recently such impressive high resolution cosmological simulations have become possible, for example Barrow et al. (2017); Rosdahl et al. (2018) or Trebitsch et al. (2020).

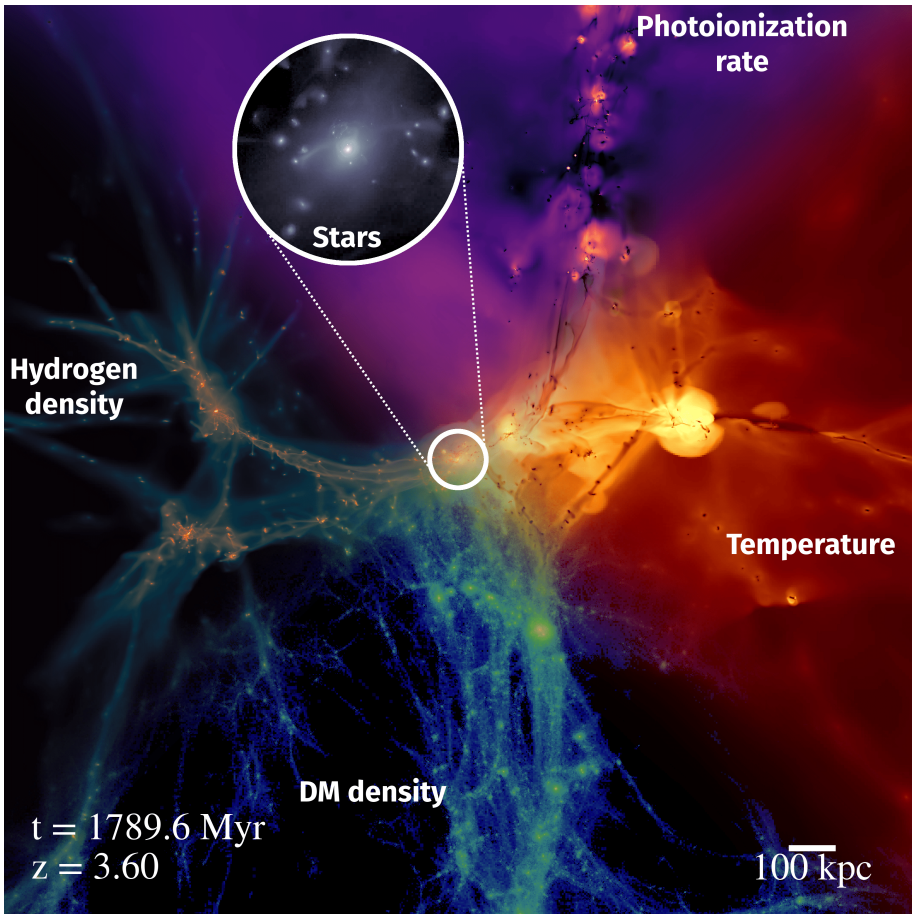

Figure 1.5: Illustration of the components of a modern cosmological radiation hydrodynamics simulation, taken from Trebitsch et al. (2020). Starting at the bottom and going clockwise, one first sees the density of dark matter, which is simulated with a N-body algorithm. Next is the hydrogen density, which evolves according to numerical implementations of hydrodynamics. Then the circle zooms in on the central region of a galaxy cluster, to show its stellar content. On the top of the figure we can see the photoionisation rate, highlighting that this simulation takes into account the transfer of ionising radiation, which is emitted by stars and AGN. Finally, the right side of the image is the gas temperature.

1.6 The aims and structure of this thesis

The first goal of this work is to find ways to indirectly measure the escape fraction of ionising photons from high redshift galaxies, where the IGM prevents all direct measurements. I focused on one promising method, using down-the-barrel absorption lines. To address this, we use new state-of-the-art Ramses-RT simulations, whose high resolution and treatment of the transfer of ionising radiation provides an accurate ionisation state of the gas, with a high enough resolution to self-consistently simulate the emission of Lyman continuum from galaxies. To maximise the usefulness of those simulations, I developed a pipeline of post-processing mainly based on the radiative transfer code Rascas[1] (Michel-Dansac et al., 2020), which is designed to propagate photons in Ramses-RT simulations to make mock observations. This pipeline provides for the first time a robust modelling of down-the-barrel absorption lines, including the scattering of photons and the fluorescent emission. These developments allow us to address several scientific questions:

- Investigate which types of galaxies contribute to Reionisation: By developing an indirect measurement of the escape fraction of ionising photons from high redshift galaxies, where direct detection of Lyman continuum is impossible, we will be able to determine the observable properties that characterise the galaxies emitting the most ionising photons.
- Study the mechanisms of escape of ionising photons: Besides the question of who is responsible for Reionisation, making mock observations and measuring escape fractions in simulations allows us to understand the conditions that are required to allow star-forming galaxies to emit ionising photons. For example, does the Lyman continuum commonly escape through holes in the ISM, devoid of gas, or through optically thin layers of gas without holes?
- Explore the critical role of dust in observations of high redshift galaxies: Dust complicates all measurements in the ultraviolet at high redshift, since it absorbs a significant fraction of photons at those wavelengths. By using simulations and mock observations, I will study how much of the light is absorbed, the consequences this has for different observables, and whether one can recover intrinsic properties of the galaxies using only observable quantities.
- Determine how down-the-barrel absorption lines are formed, and at which distance from the stars: Absorption lines carry information about the velocity and the density of the absorbing gas. Knowing the relative roles of the ISM and the CGM in shaping absorption lines is essential to be able to compute mass outflow rates and understand how galaxies enrich their surroundings in metals.

Mock observations are useful tools to study almost every topic in astrophysics. A significant part of this thesis presents the necessary steps to obtain them, from modern cosmological simulations to extensive post-processing. In Chapter 2 I introduce the simulations that I used for this work, from the code itself to the resulting collections of virtual galaxies. The first part of the chapter is dedicated to the different layers that add up to RHD simulations: the gravity, hydrodynamics, stellar formation and feedback and the radiative transfer of ionising photons. In the second part, I present the two simulations that I used, a zoom simulation following one galaxy evolving to redshift 3, and a cosmological volume containing thousands of galaxies evolving to a redshift of around 5, of which I study the hundred most massive ones. I will explain their design, and the properties of the resulting galaxies. In Chapter 3, I present my methods to post-process simulations to make mock observations in the ultraviolet. I begin with the modelling of the density of metallic ions in the gas, followed by the treatment of dust and the emission from stars. Then I present the main tool to build mock observations, the code Rascas, that transfers the photons in simulations. Chapter 4 contains the main scientific results of this thesis, and was published in early 2021 (Mauerhofer et al., 2021). There I present the resulting absorption lines from my modelling, discuss their realism, and study

[1] http://rascas.univ-lyon1.fr

many physical processes affecting the lines, such as resonant scattering. Then, I compute the escape fraction of ionising photons in different directions of observations, and compare them with the absorption lines as seen from the same directions. I also explain and discuss the resulting correlations and non-correlations. In the last part of this chapter, I increase the sample of Mauerhofer et al. (2021) by studying the escape fractions of the massive galaxies of Sphinx20. In Chapter 5, I extend the discussion of the realism of the simulated galaxies, focusing on the dust properties and the ultraviolet continuum. I discuss the critical impact that dust has on observables, by absorbing the light of bright young stars in dense star-forming regions, and whether observational methods for correcting dust attenuation accurately recover intrinsic properties of galaxies. I study gas flows in the CGM of our galaxies in Chapter 6, and explore their relations with down-the-barrel absorption lines. Finally, Chapter 7 concludes this thesis and describes the future directions I plan to take to pursue and extend this work.

2 RAMSES-RT simulations

In this chapter I present the simulation code RAMSES. In the first section I explain the functioning of RAMSES, from the AMR structure to gravity, hydrodynamics, star-formation and feedback. I also detail RAMSES-RT, where the radiative transfer of ionising photons is implemented to compute the accurate out-of-equilibrium photo-chemistry of the gas. In a second section, I present the simulations that I used in this thesis.

2.1 Presentation of RAMSES-RT

In this first section I describe the RAMSES-RT code, focusing on the routines used to simulate the formation of galaxies.

2.1.1 Initial conditions

Just after Recombination, at the start of the Dark Ages, the Universe evolved in a relatively simple manner. The evolution of the dark matter overdensity field at this time can be solved analytically with linear perturbation theory (e.g., Bernardeau et al., 2002). After a couple million years, structures started to grow larger, overdensities became large and the linear approach breaks down. It is shortly before this time that cosmological simulations start, allowing to model all the nonlinear physics involved in structure formation. The initial conditions for our RAMSES-RT simulations are generated with the code MUSIC (MUlti-Scale Initial Conditions, Hahn & Abel, 2013). Starting from given cosmological parameters, the size of the simulation box and the resolution for dark matter particles, this code computes the position and velocity of dark matter particles, as well as the gas density field, formed by hydrogen, helium, and trace amounts of heavy elements.

2.1.2 The basic structure of RAMSES

Like in all galaxy formation codes, dark matter and stars are treated as point-like particles, whose positions evolve following equations of motion. For the gas, RAMSES is an Eulerian code, more precisely using Adaptive Mesh Refinement (AMR), meaning that it splits space into cells and sub-cells of different sizes. Those cells are organised into an *octree*, following the work of Khokhlov (1998). The basic element of this octree is the "oct", which is defined by a level l and a position. An oct contains pointers to its six direct neighbours, and to the "son" oct, as illustrated in Figure 2.1. The son has a level l+1. A cell that is not refined is called a "leaf cell", and it does not contain an oct. The full simulation box is the oct of level 0, which is divided into 8 parts of level 1, which are also divided, etc. Simulations have two parameters $l_{\rm min}$ and $l_{\rm max}$, which set the minimum and maximum level of refinement, i.e. the maximum and minimum size of leaf cells.

The level of refinement in the simulation depends on different criteria given by the user. In general, the cells are more refined in regions of high density. The code has statements like "if

[1] https://en.wikipedia.org/wiki/Octree

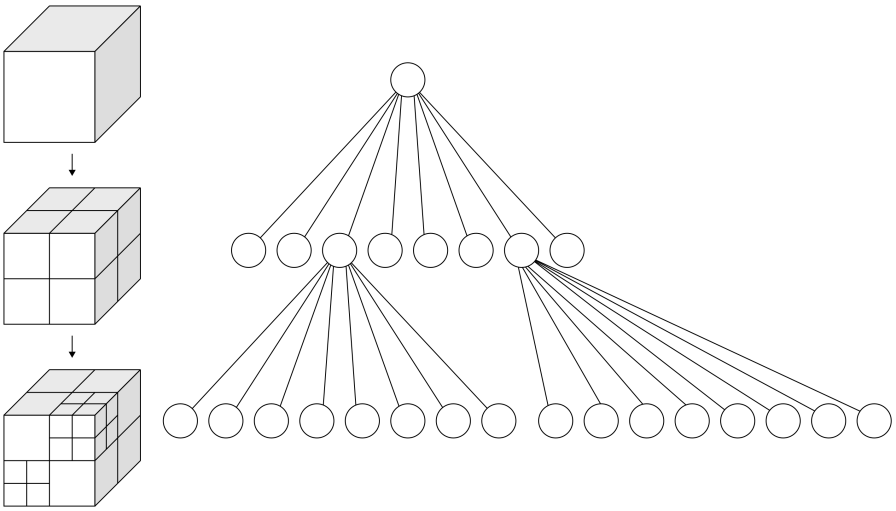

Figure 2.1: Illustration of an octree[1].

the density of this cell is larger than a threshold (including all kinds of matter), refine the cell". One can also choose to refine where the *gradient* of density is steep, or in a region of a given size at the center of a simulation, for example. A particularity of Ramses is that neighbouring cells cannot have more than one level of difference, to ensure a smooth transition between highly refined regions and more coarsely resolved ones.

In Ramses not only the spatial refinement is adaptative, but also the timestep Δt. In order to have numerical stability, a hydrodynamics scheme must satisfy the Courant condition (Courant et al., 1928):

$$\Delta t < C_{\mathrm{CFL}} \frac{\Delta x}{c_s}, \tag{2.1}$$

where C_{CFL} is a constant number chosen to be close to 0.7 for our simulations, Δx is the size of a cell and c_s the sound speed. At the most refined level l_{max}, the regions in the simulations where Δt is the smallest imposes the timestep to all the cells at this level. Then, for each lower level, the timestep is divided by two: $(\Delta t)_{l-1} = 1/2(\Delta t)_l$. The full Ramses code is organised as a recursive loop that evolves each level after another.

2.1.3 Gravity

The gravitational potential is produced by the sum of all massive components of the Universe, dark matter, gas and stars, and in turn affects the movements of those components. This potential Φ is computed via the classical formula

$$\Phi(\boldsymbol{x}, t) = -G \int \frac{\rho(\boldsymbol{x}, t)}{|\boldsymbol{x} - \boldsymbol{r}|} \mathrm{d}^3 \boldsymbol{r}, \tag{2.2}$$

where $\vec{x}$ is the position, t is the time, G is the gravitational constant and ρ is the total mass density. Dark matter and star particles move in this gravitational potential following Newton's second law, which can be expressed as

$$\frac{\mathrm{d}^2 \boldsymbol{x_p}}{\mathrm{d}t^2} = -\boldsymbol{\nabla} \Phi, \tag{2.3}$$

where $\boldsymbol{x_p}$ is the position of the particle.

Dark matter is discretised into particles of mass that typically range from 10^3 to 10^5 $\mathrm{M}_\odot$. However, in AMR simulations one needs to know the value of the gravitational potential in each cell, hence the mass of dark matter particles has to be attributed to each cell (as well as the mass of

stellar particles). Our simulations use the so-called "Cloud-in-Cell" algorithm (Hockney & Eastwood, 1981), which attributes the mass of particles to neighbouring cells. There are several ways of doing this, and we use one of the basic ones where we define a cube, of the same size as the local cells, around the particle, determine what fraction of the cube intersects with neighbouring cells, and distribute this fraction of the mass of the particle to those cells. Once the potential is known, Equation 2.3 can be solved by going into Fourier space, where the solution is simpler. then we go back to real space, and inverse the Cloud-in-Cell interpolation. For a more detailed description of the procedure I refer the reader to the thesis of Dr. Maxime Trebitsch[2].

2.1.4 Hydrodynamics solver

The inclusion of the gas component, essential for galaxy formation, brings a new set of equations to solve. The hydrodynamics of the gas is described by a set of equations containing the mass density ρ, the gas velocity $\boldsymbol{u}$, the energy density $\mathcal{E}$ and the pressure p. Starting from the Boltzmann equation, which is the most general equation describing a thermodynamic system statistically, one can derive equations for those variables by applying the method of moments and using simplifications, such as assuming the absence of friction. The moment 0 is the continuity equation:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0. \tag{2.4}$$

The first moment gives the famous Euler equation

$$\frac{\partial (\rho \boldsymbol{u})}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) + \boldsymbol{\nabla} p = \rho \boldsymbol{\nabla} \Phi, \tag{2.5}$$

with a gravity component $\rho \boldsymbol{\nabla} \Phi$, since we are in an astrophysical context. Additionally, we have the equation for the energy density:

$$\frac{\partial \mathcal{E}}{\partial t} + (\mathcal{E} + p) \boldsymbol{\nabla} \cdot \boldsymbol{u} = -\rho \boldsymbol{u} \cdot \boldsymbol{\nabla} \Phi + \Lambda(\rho, \epsilon), \tag{2.6}$$

which is also affected by the gravitational field, and by the cooling function $\Lambda(\rho, \epsilon)$, which plays a crucial role in astrophysical simulations and which I describe in the next subsection. ϵ is the thermal energy density of the gas, such that the total energy density $\mathcal{E}$ is divided into a kinetic and a thermal part:

$$\mathcal{E} = \frac{1}{2} \rho u^2 + \epsilon. \tag{2.7}$$

Finally, the system of equation is closed by an equation of state, as it is always the case in a moment based treatment:

$$p = (\gamma - 1) \epsilon. \tag{2.8}$$

In this context the ideal gas approximation applies well, giving $\gamma = \frac{5}{3}$, and an expression for the temperature T:

$$k_B T = (\gamma - 1) \mu m_H \epsilon, \tag{2.9}$$

where k_B is the Boltzmann constant, m_H is the mass of hydrogen and μ is the mean molecular weight of the gas.

It is challenging to solve those equations numerically. The first step is to regroup these equations in a global expression

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \mathcal{F}(\boldsymbol{U}) = \boldsymbol{S}(\boldsymbol{U}), \tag{2.10}$$

where $\boldsymbol{U}$ is the vector containing the variables, $\mathcal{F}(\boldsymbol{U})$ is called the hyperbolic flux and $\boldsymbol{S}(\boldsymbol{U})$ is the source term (containing the gravitational potential and the cooling function). This is a hyperbolic

[2] https://www.theses.fr/2016LYSE1119

set of equations, for which a multitude of research papers have developed resolution algorithms. Ramses uses the Godunov method (Godunov et al., 1959), which is a finite volume method. In this method, the vector $\boldsymbol{U}$ is discretised, its value $\boldsymbol{U}_i^n$ at time t_n in cell i is a volume average:

$$\boldsymbol{U}_i^n = \frac{1}{V_i} \int_{V_i} \boldsymbol{U}(\boldsymbol{x}, t_n)\, \mathrm{d}V.$$

Solving the equation of evolution consists of getting $\boldsymbol{U}_i^{n+1}$ knowing $\boldsymbol{U}_i^n$ in each cell. In Ramses this is done in two steps. First, one ignores the source term $\boldsymbol{S}(\boldsymbol{U})$ in Equation 2.10. This results in the formal solution

$$\boldsymbol{U}_i^{n+1} = \boldsymbol{U}_i^n - \frac{1}{V_i} \int_{S_i} \int_{t_n}^{t_{n+1}} \mathcal{F}(\boldsymbol{x}, t)\, \mathrm{d}t,$$

where an integral over the surface of the cell appears. The problem now amounts to getting the value of the hyperbolic flux at the border of the cells. This is equivalent to solving the *Riemann problem* (Toro, 1999).

To add the source term, one introduces an intermediary value $\tilde{\boldsymbol{U}}_i^{n+1}$ between $\boldsymbol{U}_i^n$ and $\boldsymbol{U}_i^{n+1}$. First $\tilde{\boldsymbol{U}}_i^{n+1}$ is determined by doing the computation without sources, then $\boldsymbol{U}_i^{n+1}$ is obtained by

$$\frac{\boldsymbol{U}_i^{n+1} - \tilde{\boldsymbol{U}}_i^{n+1}}{\Delta t} = \boldsymbol{S}(\boldsymbol{U}).$$

I refer the reader to Toro et al. (1994); Toro (1999); Teyssier (2002), or Teyssier (2015) for further information on the Riemann problem and the hydrodynamics solver.

2.1.5 The cooling function

There is one quantity in the last section that I have so far omitted and that requires more explanation, the cooling function $\Lambda(\rho, \epsilon)$. Even disregarding movement, friction, gravity etc, the internal energy of an astrophysical gas changes due to collisions. Free electrons collide with atoms, ions and molecules, which excites their bounded electrons, thus transferring the energy from the gas to the radiation emitted during de-excitation. These processes then reduce the internal energy of the gas. Without this cooling the gas would maintain too much energy and never collapse to form stars.

The gas is cooled down by atoms emitting photons after an excitation by a collision or during recombination events. The bremsstrahlung emission of electrons slowing down and Compton interactions between electrons and photons also contribute to the cooling. The presence of radiation can also heat up the gas, but for now let us exclude a radiation field. All those processes shape the cooling function, which describes the sum of the gain and loss of energy in the gas per unit volume and time. In Figure 2.2 I show the cooling function at different metallicities, taken from Maio et al. (2007). Metals play an important role in cooling, as we can see in the figure, the cooling is much more efficient for higher metallicities, due to the presence of more atomic transitions that allow the different ions to excite and emit photons, thus losing energy. At the highest temperatures in Figure 2.2, all atoms are completely ionised, and the most efficient cooling mechanism by far is the bremsstrahlung. Between $10^4 K$ and $10^6 K$ atomic lines dominate cooling. The last peak, at around $20'000 K$ is the temperature at which hydrogen atoms emit Lyα emission, which removes energy from the gas quite efficiently. Finally, at the lowest temperatures, the cooling is inefficient, and transitions in molecules are the only source of photon emission and cooling.

The situation is different in the presence of an external radiation field. Apart from the heating that it provides, it also ionises some species. Then even if the temperature would be such that atomic lines could cool down the gas, this does not happen because the atoms are ionised.

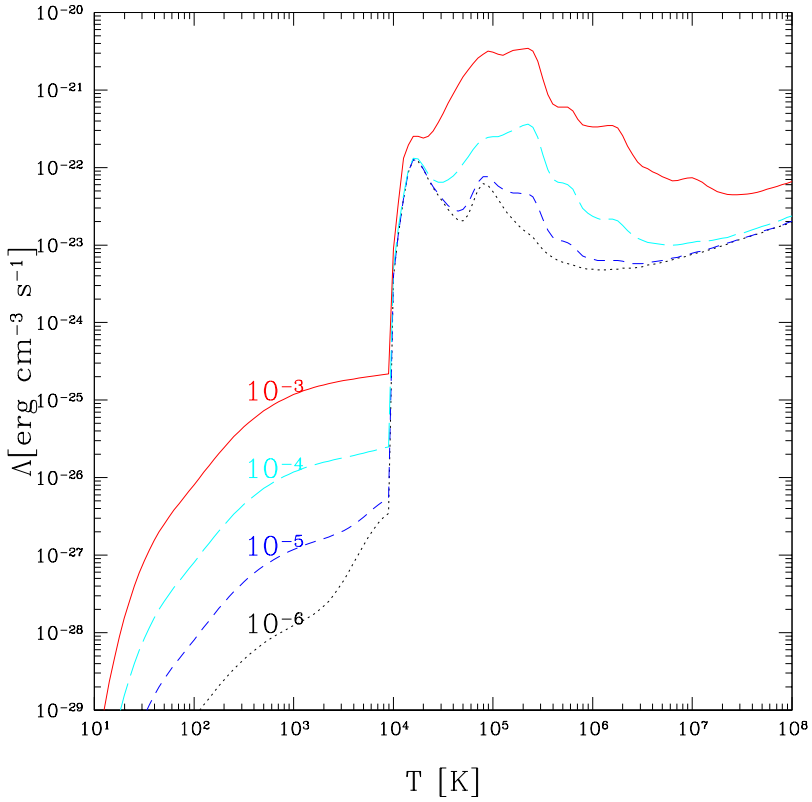

Figure 2.2: Cooling function as a function of temperature, assuming collisional ionisation equilibrium (no radiation field). The numbers on the curves of different colors indicate the metallicity, as a ratio of mass of metals over total baryonic mass. Figure taken from Maio et al. (2007).

2.1.6 Star formation

As it is, Ramses would successfully simulate the cosmological evolution of dark matter, the collapse of structure and the gas would follow the gravitational potentials to form massive baryonic structures. However, the gas would collapse to very large densities, at which we know that stars begin to form. The equations of hydrodynamics are not designed to take that into account, so one has to implement recipes to form stars. Star formation is a highly non-linear and complex problem. It is not feasible to self-consistently model this process in a galaxy formation code, because of the vast range of spatial scales between the formation of individual stars and the dynamics of a galaxy as a whole (see Figure 1.2). Star formation is also highly affected by complex physical processes such as magnetic fields or dust properties that are often not treated. Hence, all galaxy formation codes rely on so-called subgrid models to include star-formation. Those are phenomenological models that transform some of the baryonic mass of the gas to point-like particles that represent a population of stars. Although there are many different ways of doing so, I present only the algorithm that is used in the simulations that I use in this work, citing Rosdahl et al. (2018).

Stellar particles are formed in gas cells that follow the conditions described hereafter:

1. The hydrogen density in the cell satisfies $n_H > 10\,\mathrm{cm}^{-3}$.
2. The cell size is larger than the turbulent Jean length:

$$\Delta x > \lambda_{J,\mathrm{turb}} \equiv \frac{\pi\sigma_{\mathrm{gas}}^2 + \sqrt{36\pi c_s^2 G\Delta x^2\rho + \pi^2\sigma_{\mathrm{gas}}^4}}{6G\rho\Delta x}, \tag{2.11}$$

where c_s is the speed of sound, ρ is the gas density and σ_{gas} is the velocity dispersion among the 26 neighbouring cells sharing vertices with the star-forming cell.

3. The gas is locally convergent and at a local density maximum compared to the six closest neighbour cells.

In such cells, stars are formed in a stochastic manner, following a Poisson distribution, such that the average star-formation rate is

$$\overline{\mathrm{SFR}} = \frac{\epsilon_\star \rho}{t_{ff}}, \tag{2.12}$$

where $\epsilon_\star$ is the star-formation efficiency and $t_{ff} = \sqrt{\frac{3\pi}{32G\rho}}$ is the free-fall time of the star-forming cloud. Our simulations use a novel way of treating the star-formation efficiency: it varies from cell to cell based on the local properties of the gas, as explained in Kimm et al. (2017) or Trebitsch et al. (2017). Stellar particles have a constant metallicity, equal to the metallicity of the gas cells in which they are formed. As a last note, star-formation in a given cell is limited so that it converts a maximum of 90% of the gas mass into stars.

2.1.7 Feedback

If left like this, the code would produce small compact galaxies that quickly convert all their gas into stars, which would result in a stellar mass function significantly different from observations. What is missing is an energy input to counteract the gravity and slow down the collapse of gas into stars. While at first glance single stars might seem to be unable to affect a galaxy as a whole, it turns out that young massive stars exploding into supernovae release a tremendous amount of energy, and the addition of all those explosions over time provides an energy source that is far from negligible and that actively shapes the evolution of galaxies (e.g. Dekel & Silk, 1986). We call this energy the "feedback". Other types of stellar feedback include stellar winds (e.g. Geen et al., 2015), stellar radiation (Rosdahl & Teyssier, 2015, e.g.), cosmic rays (e.g. Girichidis et al., 2016) or even Lyman alpha pressure (e.g. Dijkstra & Loeb, 2008; Smith et al., 2017). Simulations including dense regions of the Universe, which form massive galaxies, also treat the feedback coming from the supermassive black holes at the center of galaxies, whose accretion disks can also provide copious amounts of energy. Here I focus on the feedback from supernovae, while in Section 2.1.8 I also discuss radiative feedback. Our simulations do not contain the other form of energy input, since we focus on relatively small galaxies where AGN do not play a major role, and other forms of feedback have proven to be subdominant (e.g. Kimm et al., 2018; Geen et al., 2021; Hopkins et al., 2021).

The feedback model of our simulations mainly comes from Kimm et al. (2017). We model the energy injected by type II supernovae. Each of these supernovae injects a total energy of 10^{51} erg in the simulation. The particularity of our model is that it distributes the energy in the form of momentum rather than thermal energy, which is done mainly to reduce the problem of overcooling (e.g. Oh & Benson, 2003). Each neighbouring cell of the cell where the stellar particle explodes receives a share of the energy, depending on the geometrical configuration. Details on the computation of the momentum "kick" is given in Rosdahl et al. (2018).

What is left to determine is the rate of SN explosions. The stellar particles represent a stellar population following the initial mass function (IMF) from Kroupa (2001). Integrating this IMF up to 100 $M_\odot$ yields a total of one SN explosion per 100 $M_\odot$ of stars formed. However, to reproduce a realistic stellar mass function, it has been determined that this rate should be multiplied by four (Rosdahl et al., 2018). This is a way to account for other feedback processes that are not modelled directly, for overcooling or for the possible presence of stars with masses larger than 100 $M_\odot$.

Finally, we assume that SN recycle 20% of the mass of the stellar particles into the ISM (Kroupa, 2001). Of this recycled mass, 7.5% is converted into metals, increasing the metallicity of the surrounding cells, based on yields in Starburst99 models (Leitherer et al., 2014). This metallicity is advected as a passive scalar, spreading through gas cells.

2.1.8 Radiative transfer

In order to simulate the Reionisation of the Universe by the first sources of light, and to better model the ISM and its chemistry, one last piece is necessary in our simulations: the transfer of ionising radiation. High energy photons have a strong impact on the gas due to heating and ionisation effects. This changes the chemistry of the gas and affects the observable properties of the galaxies, hence it is important to include those effects in order to build reliable mock observations.

However, the physics of radiation is challenging to implement numerically. The radiation field is described by the specific intensity $I_\nu(\boldsymbol{x}, \boldsymbol{n}, t)$ [erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}], which depends on seven variables: the three dimensional space, the direction of propagation described by two variables, the time and the frequency of the light. It is impossible to model the specific intensity in those seven dimensions in simulations, even if coarsely discretised. Additionally, the speed of light is much larger than the fastest velocities of hydrodynamical quantities, which means that the Courant condition (Equation 2.1) imposes an extremely short timestep, slowing down the simulations significantly.

There are two main implementations that allow us to overcome those difficulties. The first one is using "ray tracing". This method consists in launching rays from the stellar particles to the gas cells and computing the optical depth of photons. This is computationally expensive since the numerical cost is proportional to the number of sources and of resolution elements. There are of course many developments to go past these limitations (e.g. Wise et al., 2012; Wünsch et al., 2021). However, it is still challenging to apply ray tracing in large scale cosmological simulations, especially with the high resolution needed for resolving the escape of ionising photons from galaxies. The other method, that is used in Ramses-RT, is called the "moment method" (Rosdahl et al., 2013). The principle is to drastically reduce the dimensionality of the radiation field I_ν. We can derive the following "radiative transfer equation" based on phenomenological considerations:

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \boldsymbol{n}\cdot\boldsymbol{\nabla} I_\nu = \eta_\nu - \kappa_\nu I_\nu, \tag{2.13}$$

where η_ν is the radiative energy density input per unit time and solid angle, for example due to photons emitted by atoms, and κ_ν is an absorption term, which determines what fraction of light is absorbed by the medium at frequency ν. Let us consider a simplified case, in one dimension with no time dependence and no emission:

$$\frac{\partial I_\nu}{\partial x} = -\kappa_\nu I_\nu.$$

The solution here is:

$$I_\nu(x) = I_\nu(0) e^{-\kappa_\nu x}.$$

We introduce the quantity $\tau_\nu \equiv \kappa_\nu x$, called the "optical depth". When a medium has $\tau_\nu \ll 1$, the absorption is negligible, only a fraction τ_ν of photons are absorbed, and we call the medium "optically thin". On the contrary, if $\tau_\nu > 1$, the medium is called "optically thick", and only a small fraction $e^{-\tau_\nu}$ of photons are not absorbed.

Let us come back to the moment method, to implement radiation in simulations. The quantities that we use are the so-called moments of I_ν, normalised by photon energy, namely, the photon density N_ν, the photon flux density $\boldsymbol{F}_\nu$ and the normalized radiation pressure tensor $\bar{P}_\nu$:

$$N_\nu = \frac{1}{ch\nu}\oint I_\nu \,\mathrm{d}\Omega, \tag{2.14}$$

$$\boldsymbol{F}_\nu = \frac{1}{h\nu}\oint \boldsymbol{n}\, I_\nu \,\mathrm{d}\Omega, \tag{2.15}$$

$$\bar{P}_\nu = \frac{1}{ch\nu}\oint \boldsymbol{n}\otimes\boldsymbol{n}\, I_\nu \,\mathrm{d}\Omega. \tag{2.16}$$

Those new variables are less costly to simulate since we drop the information of directionality when integrating over $d\Omega$. Applying Equation 2.13, one finds new equations for the moments:

$$\frac{\partial N_\nu}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{F}_\nu = -\kappa_\nu c N_\nu + \dot{N}_{\text{source}}, \tag{2.17}$$

$$\frac{\partial \boldsymbol{F}_\nu}{\partial t} + c^2 \boldsymbol{\nabla} \bar{P}_\nu = -\kappa_\nu c \boldsymbol{F}_\nu, \tag{2.18}$$

where $\dot{N}_{\text{source}}$ comes from the emissivity term η_ν, and is equal to the number of photons emitted by the stars in the simulation (and the gas recombination and AGN if present, which is not the case in our simulations). The absorption term κ_ν can also be specified here. Since Ramses-RT deals with the transfer of ionisation radiation, absorption of photons in our simulations is only due to the ionisation of hydrogen and helium, and a small contribution from dust. Hence, we have

$$\kappa_\nu = n_{\text{H I}}\, \sigma_\nu^{\text{H I}} + n_{\text{He I}}\, \sigma_\nu^{\text{He I}} + n_{\text{H II}}\, \sigma_\nu^{\text{He II}} + \kappa_{\text{dust}}, \tag{2.19}$$

where n_i is the number density of element i, σ_ν^i is the photoionisation cross-section of element i as a function of frequency, that we take from Verner et al. (1996).

When deriving hierarchical equations with the moment method, the system has to be closed by an additional equation, called the closure. The code Ramses-RT uses the M1 closure (Levermore, 1984), which is:

$$\bar{P}_\nu = \left(\frac{1-\chi}{2} \mathbb{I} + \frac{3\chi - 1}{2} \boldsymbol{n} \otimes \boldsymbol{n} \right) N_\nu, \tag{2.20}$$

where $\mathbb{I}$ is the identity matrix and

$$\chi = \frac{3}{c h \nu N_\nu} + \frac{4 |\boldsymbol{F}_\nu|^2}{c^3 h \nu N_\nu^3}.$$

This method has the advantage of making the problem hyperbolic, so that the same hyperbolic solver can be used for hydrodynamics and for the radiative transfer. More details and explanation about this method are in Aubert & Teyssier (2008) or Rosdahl & Teyssier (2015).

The system of equations is simplified after the discretisation of the frequency of photons into three energy groups. Since the simulation is designed to follow the ionisation of hydrogen and helium, the photons are divided in a group with energy between 13.6 and 24.59 eV, that ionises H^0, a group with energy between 24.54 and 54.42 eV, which ionises H^0 and He^0 and a third group with energy above 54.42 eV, which ionises H^0, He^0 and He^+. All quantities that depend on frequency are integrated between those energies. The resulting equations are similar to Equations 2.17,2.18 and 2.20.

The system is resolved in different steps, using the so-called operator splitting approach. First, photons are injected via the source term $\dot{N}_{\text{source}}$. In our simulations the only sources are the stellar particles, and the number of photons per particle is computed with SED libraries, based on the metallicity, age and mass of the particle, which I explain in Chapter 3. Then the code performs the hydrodynamical step, as described in Section 2.1.4. The third step is to do the radiation transport while ignoring all emission and absorption, that is to say putting the emissivity and the photoionisation cross-section to zero. This amounts to solving a *homogeneous* hyperbolic set of equations, which the hydro solver is designed to do. Last but not least, the code performs the non-equilibrium chemistry in every cell. As explained in Rosdahl & Teyssier (2015), this means updating the ionisation fractions of hydrogen and helium, as a result of the new ionisation radiation that was added in the cells, updating the photon density and flux field since some photons were absorbed during ionisation, and updating the temperature of the cell based on the cooling and heating mechanisms at play.

I mentioned above that the speed of light was so much faster than the hydrodynamic variables that the Courant condition would impose a timestep too small to be used in practice. The solution

adopted in Ramses-RT is to artificially reduce the speed of light. This approximation remains accurate as long as the speed of propagation of ionisation fronts, which depends on gas density and temperatures, is lower than the chosen reduced speed of light. For dense star-forming clouds, this speed of propagation is extremely low, allowing for a significantly reduced speed of light. On the contrary, ionisation fronts in the diffuse IGM propagate nearly at the speed of light. This is why the Sphinx simulation (that I present below), which simulates a large-scale box but with a very high resolution, needed a "variable speed of light" approach, which is explained in Katz et al. (2017) and Rosdahl et al. (2018).

2.2 Resulting collection of simulated galaxies

In this work I use galaxies from two distinct simulations performed with the code Ramses-RT, described in the last section. The first one is a zoom simulation focusing on one galaxy, while the second is a large cosmological simulation of 20 cMpc in size.

2.2.1 Zoom simulation

A zoom simulation uses initial conditions such that one dark matter halo, at the center of the box, is well resolved, while the surrounding cosmological environment has a coarser resolution. This is more accurate than an idealised simulation of an isolated galaxy, while still being much less computationally intensive than a full cosmological simulation. The zoom I am using was done by Prof. Jérémy Blaizot. We begin with a large scale simulation of dark matter only, then choose a halo with the desired properties (mass, in a group or isolated, etc) and then we rerun the simulation forcing the dark matter particles and gas cells to be well resolved in the vicinity of the halo, while keeping it coarser in the surroundings. The simulation I use follows a relatively isolated galaxy with a mass making it a typical Lyman alpha emitter, with a mass of 2.3×10^9 $M_\odot$ at redshift 3. It includes the transfer of ionising photons, modelling the emission from stellar particles based on the stellar library of Spectral Energy Distribution (SED) Bpass[3] (Eldridge et al., 2008; Stanway et al., 2016). SED libraries allow us to determine the emissivities of each stellar particle based on its mass, age and metallicity. BPASS has the particularity to include the effect of binary stars, which results in a more intense emission of ionising photons.

I plot in Figure 2.3 the column density of neutral hydrogen and the stellar surface density for the last snapshot of the simulation, at redshift 3. In the left panel, we see streams of gas being accreted from the IGM, while supernova feedback creates holes with low density in the ISM. The right panel shows that the stars are concentrated in the middle of the halo, and that small satellites and low-density stellar streams galaxies are orbiting the main galaxy.

[3] https://bpass.auckland.ac.nz/index.html

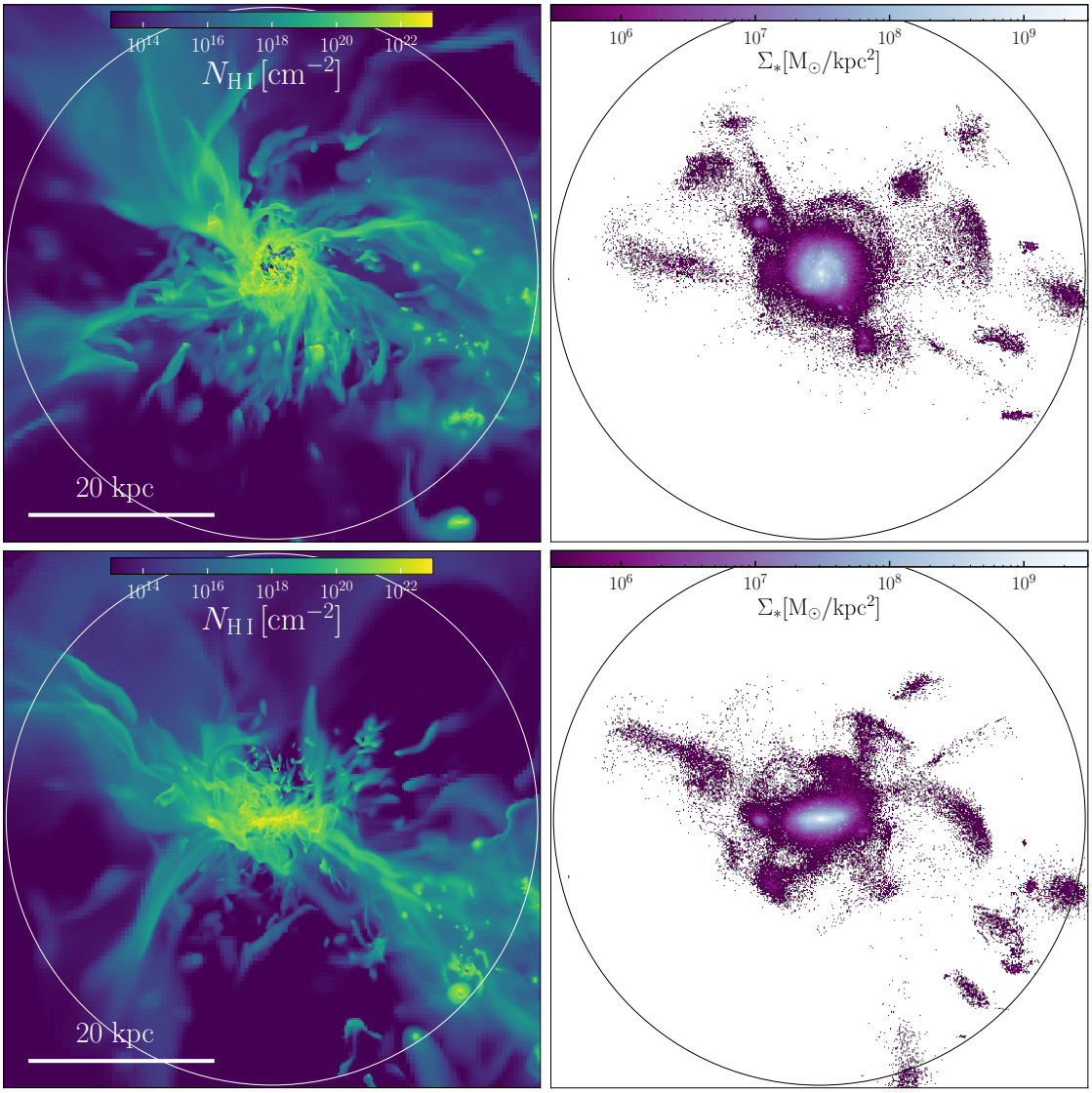

Figure 2.3: Images of the last snapshot of the zoom simulation, at redshift 3. Upper left panel: column density of neutral hydrogen, in atoms per cm^3, seen from the z-axis of the simulation box. Upper right panel: stellar surface density, in M$_\odot$/kpc^2, seen from the z-axis. The two lower panels are the same but seen from the x-axis. The white circles indicate the virial radius of the halo. The white segments represent a size of 20 kpc.

Several properties of this galaxy, evolving from redshift 8 to 3, are displayed in Figure 2.4. The top panel shows the evolution of the mass of dark matter, gas and stars. This galaxy is similar to the one presented in Mitchell et al. (2021). They both have a large stellar mass to halo mass ratio, of around 0.04 around redshift 3, while Behroozi & Silk (2015) and Girelli et al. (2020) predict that at this redshift and halo mass, the ratio should around ten times lower. It appears that the simulation is forming too many stars, which suggests that the star-formation and feedback modelling is not yet fully realistic. The middle panel displays the metallicity of the gas and stars, in the ISM for the solid lines and the CGM for the dotted lines. The stellar metallicity at redshift 3 is about 0.4 Z$_\odot$, which can be converted to 12 + log(O/H) $\sim$ 8.3. This is slightly larger than the $\sim$ 8.0 predicted by Steidel et al. (2014) or Sanders et al. (2015), who analyse spectra of high redshift galaxies to determine their mass metallicity relation. The gas metallicity in the ISM is higher than the stellar metallicity, which is easily understood when considering that a stellar particle has a fixed metallicity equal to the one of the gas cell in which it is born, while the surrounding gas continues

to be enriched in metals by the supernovae explosions. In the CGM however, the metallicity of the stars is higher than the one of the gas, because of pristine gas falling in from the IGM and lowering the metal mass fraction of the gas. The lower panel shows the star-formation rates averaged over the last 10 Myr and the last 100 Myr. It increases steadily until redshift 4, and then stabilises, with small oscillations.

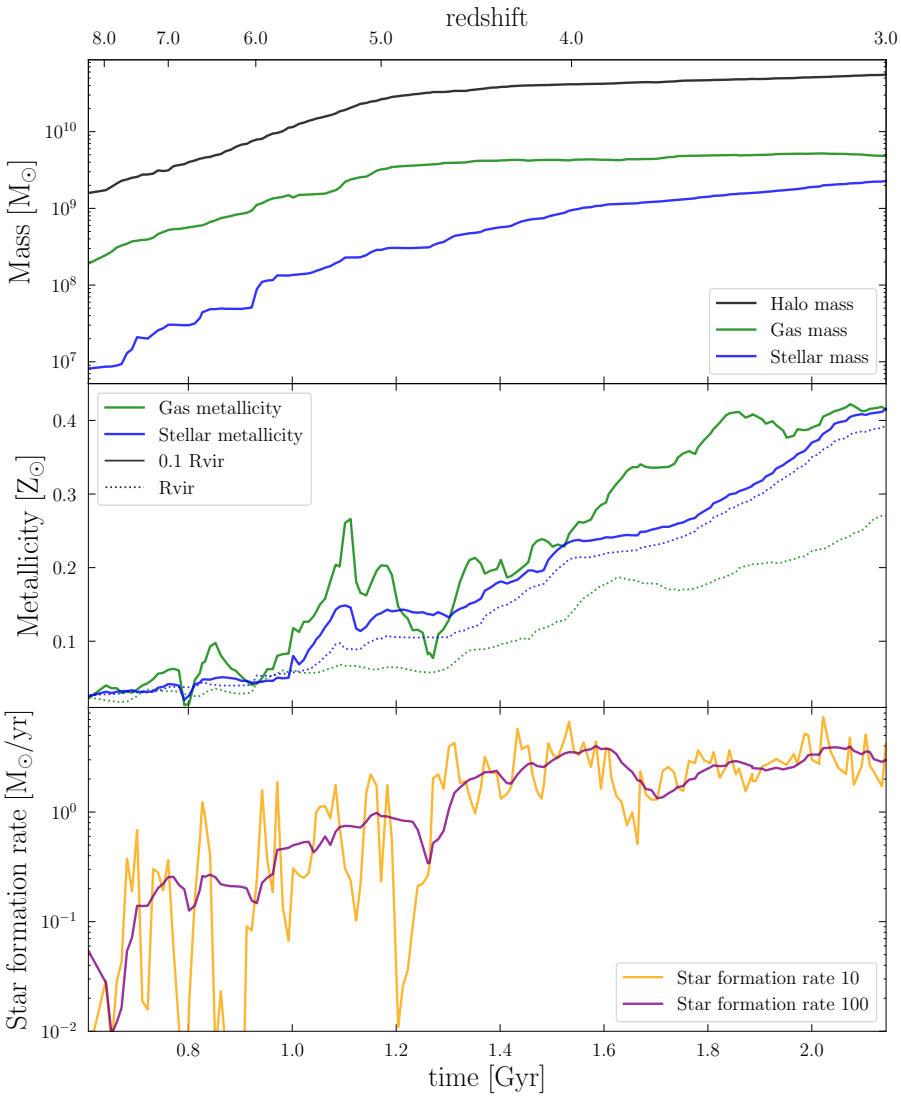

Figure 2.4: Properties of the central galaxy of the zoom simulation described in this section, evolving from redshift 8 to 3. The top panel shows the evolution of the total mass, the gas mass and the stellar mass. The middle panel shows the evolution of the gas metallicity in green and of the stellar metallicity in blue. The solid lines are metallicities in the ISM, in a sphere of radius ten times smaller than the virial radius. The dotted lines are metallicities in the full virial radius of the halo. The bottom panel displays the evolution of the star formation rate, averaged over the last 10 Myr in orange and over the last 100 Myr in purple.

This simulation is state-of-the-art in terms of resolution and physics. In Figure 2.5 I show the distribution of cell size in the ISM and the CGM for the last snapshot at redshift 3. The highest resolution is 14 pc, which is unusual for a cosmological simulation including radiative transfer of ionising photons. In the ISM, the blue curve, 85% of the mass is in cells with this maximum resolution. When weighing by volume it is lower, with only about 4% of the volume resolved at 14 pc, but almost 80% resolved better than 56 pc. In the CGM, the green curve, 16% of the

mass is at the maximum resolution, and 80% in cells with sizes smaller than 450 pc. This high resolution allows us to consistently simulate the escape of ionising radiation from the galaxy, and to reproduce realistic observables, as I show in Chapters 4 and 5.

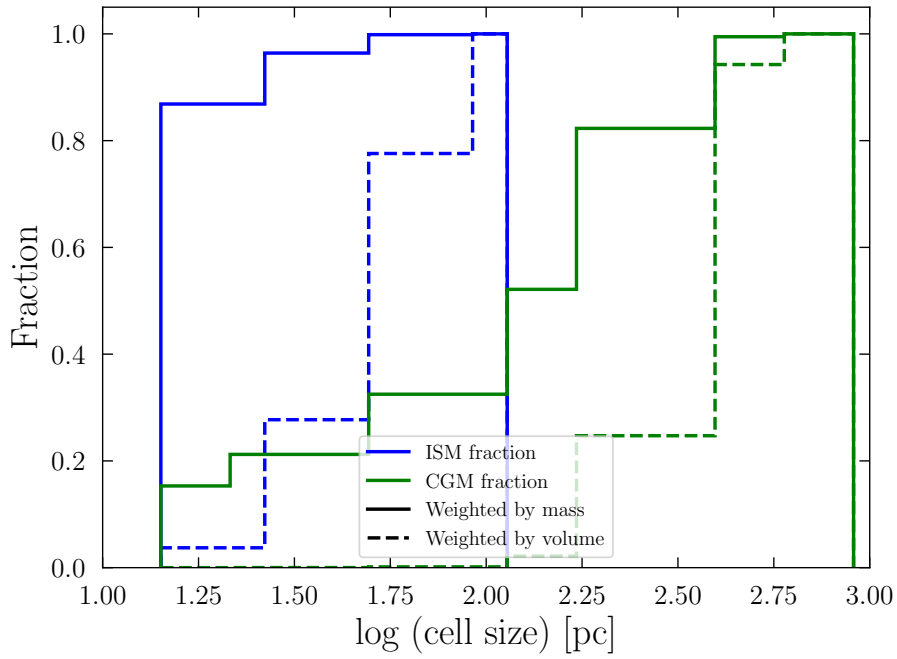

Figure 2.5: Cumulative distributions of cell resolution in the last snapshot of the zoom simulation, at redshift 3.0. The solid curves show the fraction of gas mass that is resolved better than a given cell size, and the dashed curves show the fraction of volume. Blue curves are computed in a sphere with a radius of 10% of the virial radius and the green curves are computed in a shell with inner and outer radii equal to a tenth and to the full virial radius.

It is not clear from Figure 2.3 alone whether the simulated galaxy has a disk structure. To answer this question, I compute the angular momenta of the stars and of the gas, to see their evolution in time. I plot in Figure 2.6 the z-component of the angular momentum of stars in the virial radius, and of gas in spheres of radius 0.1 and 0.33 times the virial radius. The angular momentum vectors are normalised, such that a value of 1.0 or -1.0 means that the axis of rotation of the stars (gas) is parallel to the z-axis of the grid of the simulation, and a value of 0 means that it is perpendicular. We see from Figure 2.6 that at early times, the curves are oscillating significantly, indicating that the stars and gas do not rotate around an axis of symmetry, but have a chaotic movement. Then the curves stabilise, which shows that a symmetry is arising, and stars and gas form a disk. The outer gas stabilises slightly earlier than the stars. It is intriguing that the z-component of the angular momentum stabilises at a value of -1. This means that the axis of rotation of the galaxy is parallel to the z-axis of the simulation box. This is a common behaviour in AMR simulations (e.g. Hopkins, 2015 and references therein).

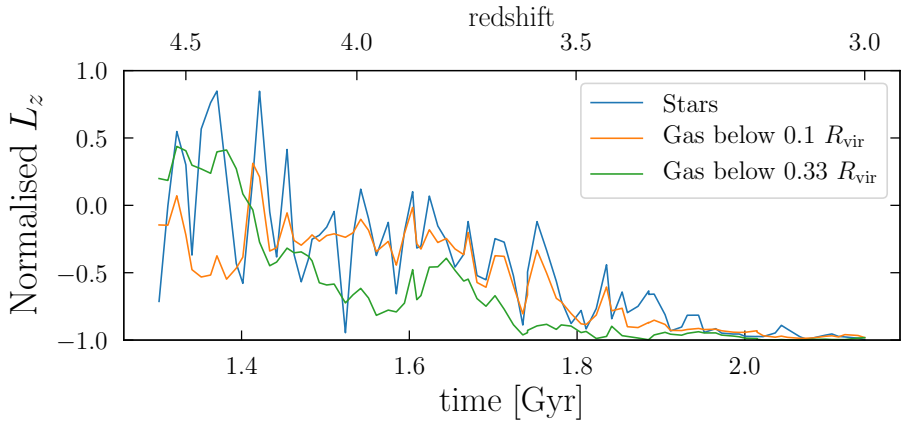

Figure 2.6: Evolution of the z-component of the angular momentum normalised between -1 and 1, from redshift 4.6 to 3. The z-axis corresponds to one of the axes of the simulation box. The blue line is the angular momentum of all the stellar particles in the virial radius of the galaxy. The orange and green lines represent the gas in a sphere of radius equal to a tenth of the virial radius, and a third of the virial radius, respectively. The data is from the zoom simulation.

2.2.2 Sphinx20

The Sphinx suite of RAMSES-RT simulations[4] is a project to understand the Reionisation of the Universe in a volume with an average density (not a void nor a galaxy cluster). The original simulation (Rosdahl et al., 2018) had a volume of $(10\,\mathrm{cMpc})^3$, while the new generation Sphinx20 simulation has a volume of $(20\,\mathrm{cMpc})^3$ (Rosdahl et al. in prep.). It has an exquisite resolution, with the smallest cells being only 10 pc at redshift 6. This resolution and box size allow Sphinx20 to be a unique state-of-the art simulation of Reionisation. In Figure 2.7 I show a comparison of box sizes and spatial resolution for different simulations of Reionisation. Sphinx is by far the one with the highest resolution, and the last version has a volume large enough to include a halo of the size of the Milky-Way.

[4] https://sphinx.univ-lyon1.fr/

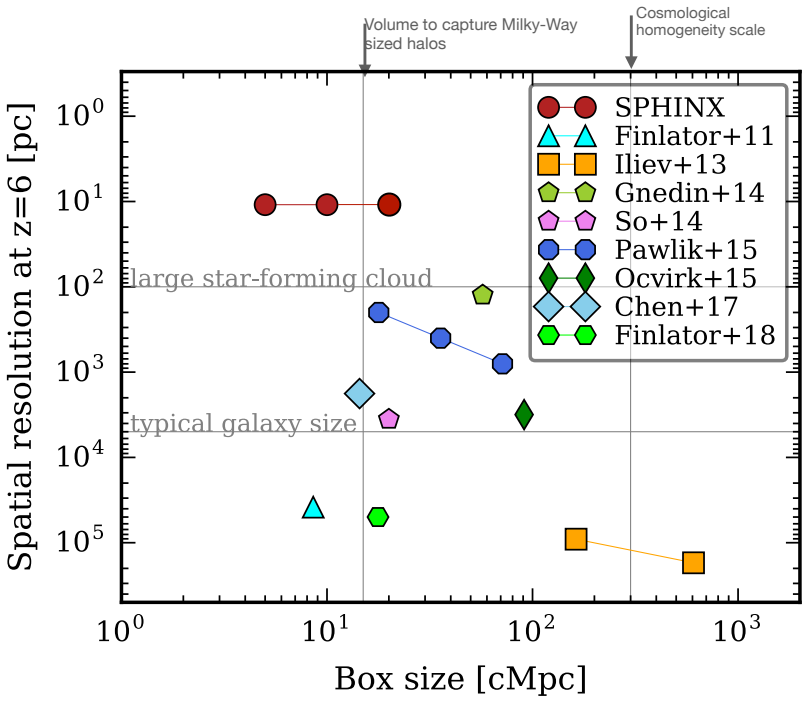

Figure 2.7: Comparison of simulation sizes and resolution. The three red circles represent three versions of the Sphinx simulations (Rosdahl et al., 2018), called Sphinx5, Sphinx10 and Sphinx20. Other simulations are represented with different colours and symbols (see legend). Studies containing several simulations with different parameters are shown with the same symbols connected by a segment. The vertical grey lines indicate the minimum size a cosmological volume must have to include a galaxy as large as the Milky Way (15 Mpc) and the scale at which the Universe is homogeneous (300 Mpc). The two horizontal grey lines indicate the resolution that is necessary to resolve galaxies (5000 pc) and large star-forming clouds (100 pc). Figure provided by Joakim Rosdahl.

One of the main results of the Sphinx simulations is that the history of Reionisation highly depends on the predictions of the strength of the ionising luminosity of stellar populations. Different stellar libraries result in a different timing of the Reionisation, as can be seen in Figure 2.8. The green dotted line is a version of Sphinx using the Galaxev[5] SED library (Bruzual & Charlot, 2003), which does not include the effect of binary stars. As a result, the Universe does not reionise, in tension with observations. On the contrary, the version of Sphinx using the SED library BPASS 2.0 reionises the Universe too early. The solid red line and dotted yellow line are from Rosdahl et al. in prep., using the newer version BPASS2.2.1, where they assume a lower fraction of stars in binary systems. This results in an intermediate speed of Reionisation, in better agreement with observations, although slightly too slow.

[5] http://www.bruzual.org/galaxev/galaxev.php

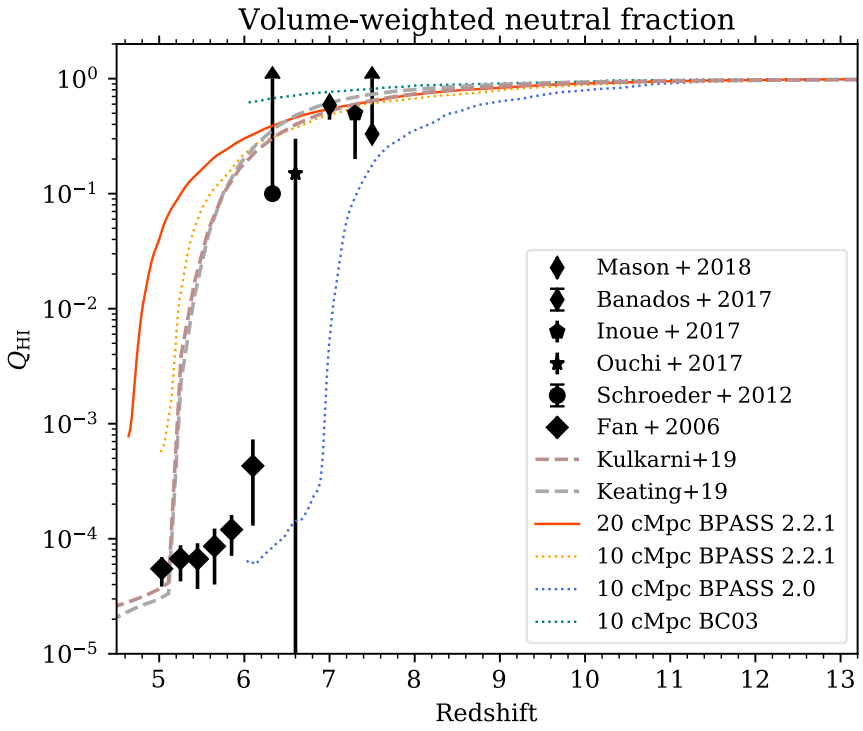

Figure 2.8: Hydrogen neutral fraction as a function of time in the first 1.3 Gyr of the Universe. The solid red line and dotted yellow line are obtained from two new simulations (Rosdahl et al. in prep.) with volumes 20 and 10 cMpc, and which use the SED library BPASS2.2.1. The blue and green dotted lines are from Rosdahl et al. (2018), with a volume of 10 cMpc and the SED library BPASS2.0. and Galaxev (Bruzual & Charlot, 2003), respectively. The brown and grey dashed lines are from other simulations (see legend). The different black symbols indicate observational constraints of the hydrogen neutral fraction (references in legend). Figure provided by Joakim Rosdahl.

The Sphinx20 simulation contains thousands of galaxies. The stellar mass - halo mass relation is shown in Figure 2.9. As for the zoom simulation, the stellar mass for a given halo size is slightly larger than expected in e.g. Girelli et al. (2020), although there is almost no data at such high redshifts. In this thesis I will analyse the 100 most massive halos of Sphinx20 at redshift 5.11, which have stellar masses in the range 10^8 – $10^{10.3}$ $M_\odot$.

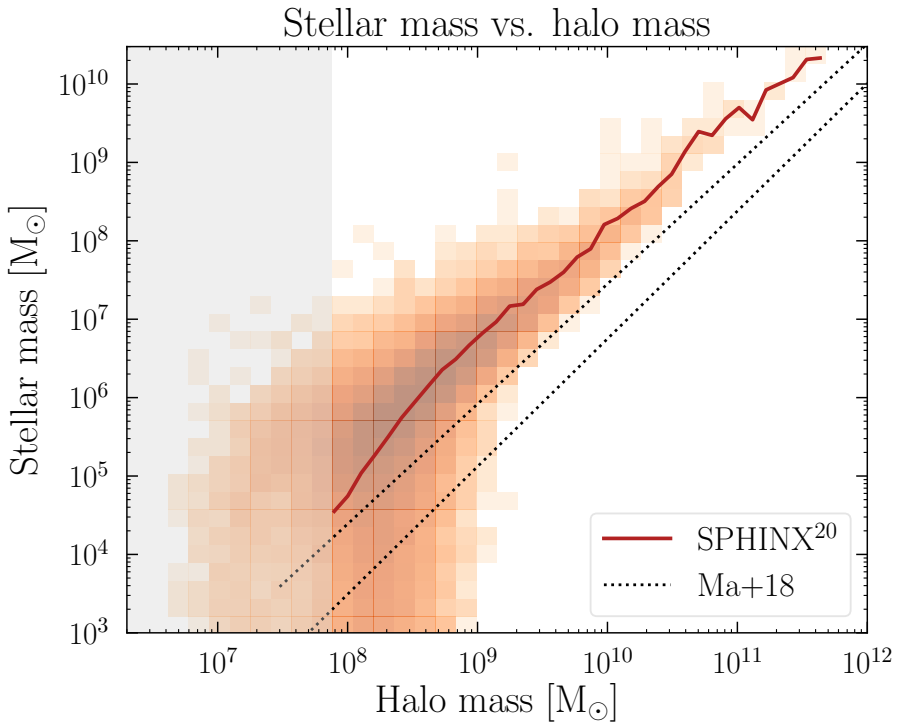

Figure 2.9: Stellar mass - halo mass relation for the galaxies of the Sphinx20 simulation at redshift 5.11. The red line shows the mean of the data and the dotted lines are from Ma et al. (2018). The color-scale represents the number of galaxies with given stellar and halo mass, with darker regions having a larger number. The grey shaded area is where halos have fewer than 300 dark matter particles.

3 The path from simulation to observables

In this chapter I present how to bridge the gap between simulations and observations. The post-processing of simulations is done in three main steps, which correspond to the three main parts of the Monte-Carlo post-processing code Rascas, which I am using in this work. The first part is the determination of the gas properties in the simulation, with which the light will interact. In Section 3.1, I explain how I compute the density of different elements in each gas cell of the simulations. In Section 3.2, I show how to extend the radiation field at energies lower than 13.6 eV, which is important to compute the ionisation equilibrium of metallic elements. The computation of this ionisation equilibrium is discussed in Section 3.3. The next section about the first part of Rascas is Section 3.4, where I explain the role of the fine-structure levels of the ground state of ions, and how I compute the level populations. Finally, I present our dust modelling in Section 3.5.

The second part of Rascas is to model the emission of continuum from the stellar particles. This is presented in Section 3.6. Finally, the third part of Rascas is the transfer of the photons in simulations, resulting in mock observations, which I explain in Section 3.7. I also define several properties of absorption lines and show examples of mock observations in Section 3.8.

3.1 Abundance ratios and dust depletion

To begin our journey from the outputs of simulations to reliable and accurate mock observations of galaxies, including spectral lines of metallic species, we have to begin by inferring the density of those elements in each cell of the simulation. Ramses-RT treats the chemistry of hydrogen and helium on-the-fly, including out-of-equilibrium ionisation, and follows the metallicity in every cell, as described in the previous chapter. However, it does not compute how this metallicity is divided into different "heavy elements" (everything outside hydrogen and helium) and how ionised those elements are. In this section, we address the first issue: how much of each element is there in our metallicity variable?

When facing such a complex problem as stellar nucleosynthesis in the high-redshift Universe, one sometimes has to resort to the simplest possible solution. Since little is known about the ratio of heavy elements at the epoch that interests us, above redshift 3, and since it is too computationally intensive to add models in our simulations which include different supernova yields for each element, we adopt solar abundance ratios. For a cell of hydrogen number density n_H and metallicity Z, the number density of an element E is

$$n_E = A_E^{\odot} \frac{Z}{Z_{\odot}} n_H, \tag{3.1}$$

where $A_E^{\odot} = n_E^{\odot}/n_H^{\odot}$ is the ratio of n_E over n_H in the Sun, and $Z_{\odot}$ is the solar metallicity. The most recent reference for those solar values is Asplund et al. (2021), where $Z_{\odot} = 0.0139$, and from which some values of $A_E^{\odot}$ are shown in Table 3.1.

Table 3.1: Solar abundances of the nine most abundant metals in the Universe, in increasing order of atomic mass. The numbers are ratios of element densities over hydrogen density, to be multiplied by 10^{-6}. Data taken from Asplund et al. (2021).

element E	C	N	O	Ne	Mg	Al	Si	S	Fe
$A_E^\odot \equiv n_E^\odot / n_H^\odot \, [\times 10^{-6}]$	288	67.6	490	115	35.5	2.69	32.4	13.2	28.8

As a way to check if this approximation is realistic, we look at the predicted yields of metals from stellar populations in Figure 3.1, taken from Few et al. (2012). Looking at the upper panel, some elements have no evolution after the initial increase that lasts until $\sim$ 10 Myr, such as oxygen, neon, magnesium or silicon. Other elements continue to grow on long time scales, like carbon, nitrogen or iron. Elements that do not evolve after 10 Myr likely have a similar abundance ratios at high redshift as they do now, whereas the ones that evolve might have been quite different at the epoch we are interested in compared to nowadays in the Sun. However, this is a rough estimation, and Figure 3.1 only shows results for solar metallicities and zero metallicity. Regardless of those numerous uncertainties, I show in Mauerhofer et al. (2021), which is in Chapter 4, that this abundance has no large impact on observed metallic absorption lines, even when the density is changed by a factor of two.

As an additional difficulty, metallic elements in galaxies are present in different physical states, and only the gaseous component interacts with stellar light to produce absorption lines. A fraction of each element is instead depleted on dust grains (e.g. De Cia et al., 2016; De Cia, 2018), which removes it from the gaseous phase which we are interested in. Once again, it is a challenging problem to know how much of each element is trapped into dust. Recent progress has made use of zinc spectral features (Jenkins, 2009; De Cia, 2018), which is an element that is almost not present in dust, to infer the dust depletion factors of other elements, either in our Galaxy using ISM-absorption lines of single stars, or in high-redshift galaxies using absorption lines imprinted on the continuum of background quasars.

Since this adds up with the uncertainties of the abundance ratios, we choose to ignore dust depletion factors. As I stated above, I show in my first paper that the exact density of an element is not important to model down-the-barrel absorption lines. What matters most is the fraction of stars that are covered by optically thick gas and the velocity dispersion of the gas.

3.2 Extending the spectral range of the radiation field

Before tackling the question of ionisation fractions of different elements, I present here one additional necessary step of the post-processing. In the previous chapter I described in depth the properties of the simulation of the typical Lyman-alpha emitter galaxy that I use in this work. I explained how the radiation is split into three energy groups, whose energies are designed to trace the ionisation of the three species that interact with ionising photons: H^0, He^0 and He^+. This means that the simulation does not consider photons with energies below 13.6 eV.

However, many (neutral) metals have ionising energies lower than this limit. Photoionisation is an important process to take into account when computing the densities of ions in galaxies. Therefore I compute the radiation strength at lower energies in our simulations, which allows to accurately predict the balance between some neutral elements and their first ionised state, for example C^0 and C^+ or Si^0 and Si^+. Figure 3.2 shows the ionisation energies and wavelengths of several elements that we are interested in, overplotted on a typical UV spectrum of a simulated

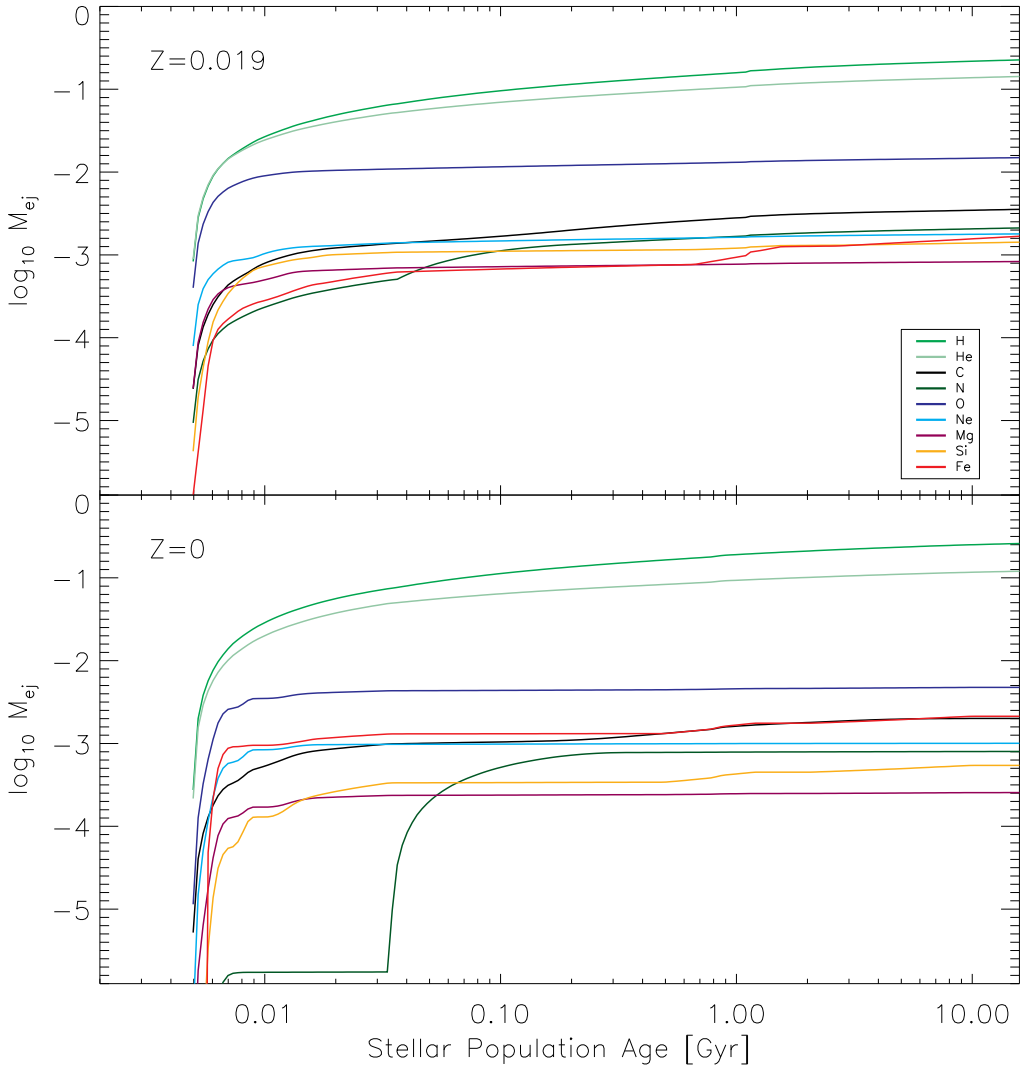

Figure 3.1: Citing Few et al. (2012): *Ejection rate of dominant elemental isotopes, per unit stellar mass, as a function of age for a Kroupa (2001) IMF. The upper and lower panels correspond to solar and Population III metallicity simple stellar populations respectively.*

galaxy. For neutral carbon, silicon, iron and magnesium one needs to know the radiation down to around 8 eV to be able to compute their photoionisation rate.

I explore different ways of achieving this in the next sections.

3.2.1 Assuming a homogeneous "sub-ionising" radiation field

As a first approximation, one can say that the metal-ionising, but non-hydrogen ionising, radiation field (which I designate as "sub-ionising") is equal to the cosmic UV background (UVB), emitted by the ensemble of galaxies and AGN at a given redshift. This model is of course omitting the contribution of stars in the galaxies, which emit intense UV radiation in the range of interest, 6-13.6 eV. Which of the two is the strongest, the UVB or the stellar light, depends on the environment. Typically, the CGM is far enough from stars to be mainly affected by the UVB, while the ISM is strongly impacted by stellar radiation. However, dust can shield both contributions, UVB and stars, resulting in a nontrivial configuration.

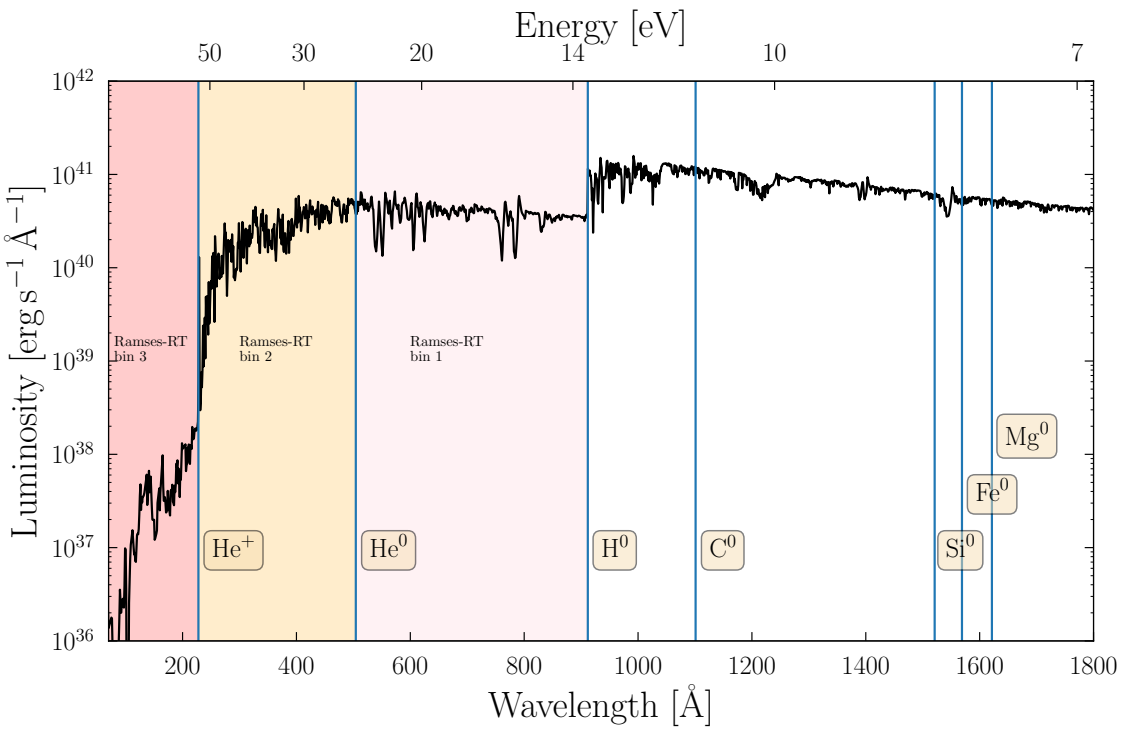

Figure 3.2: Intrinsic spectrum of the galaxy from the zoom simulation presented in Chapter 2, at the last snapshot ($z = 3$). It is the sum of the spectra of each stellar particle, which were obtained with BPASS. The vertical blue lines indicate ionisation wavelengths (and energies) of different species.

In Figure 3.3 I compare the strength of the UVB at redshift 3 from Haardt & Madau (2012) without dust attenuation, with the flux from a bright young stellar population at a distance of 10 parsec, with different dust attenuation factors. As I explain in Section 3.5, the opacity of dust in our models is proportional to the density of neutral hydrogen, so I index the strength of dust in Figure 3.3 using $n_{\rm H\,I}$. We see that without dust the stellar spectrum is almost two orders of magnitude more intense than the UVB. However, they become similar for a density of neutral hydrogen of around 175 cm^{-3}, which is easily reached in the ISM. Above 500 cm^{-3}, the stellar spectrum is extremely attenuated by dust.

In practice, I choose to set a uniform UVB of ten times the value of Haardt & Madau (2012), to account for a global contribution of stars. However, I do not put any sub-ionising radiation in cells that have $n_{\rm H\,I} > 100\,{\rm cm}^{-3}$, because in those cells dust alone would shield the UVB. This is a first approximation, but I verified that our mock observations do not change if I increase or decrease the value of the UVB by a factor of ten or even 100. My interpretation is that the regions that are the most affected by the change of details in our modelling of the sub-ionising radiation field are the dense and cold regions, which could be sensitive to small differences in the models, but are so dense in dust that they screen efficiently both continuum and line radiation. Hence, these regions are not visible in the final mocks. Less dense regions are not sensitive to the details of the model, they are photo-ionised by the sub-ionising radiation anyway. Thus our mock observations seem to be robust to my modelling of sub-ionising radiation. I come back to the intriguing dust properties of the simulations in Chapter 5.

3.2.2 Using Ramses-RT as a post-processing tool: restarting the simulation

My second approach to compute the sub-ionising radiation field in post-processing is to use Ramses-RT itself, as the code has all the tools to propagate radiation in the simulation. The proce-

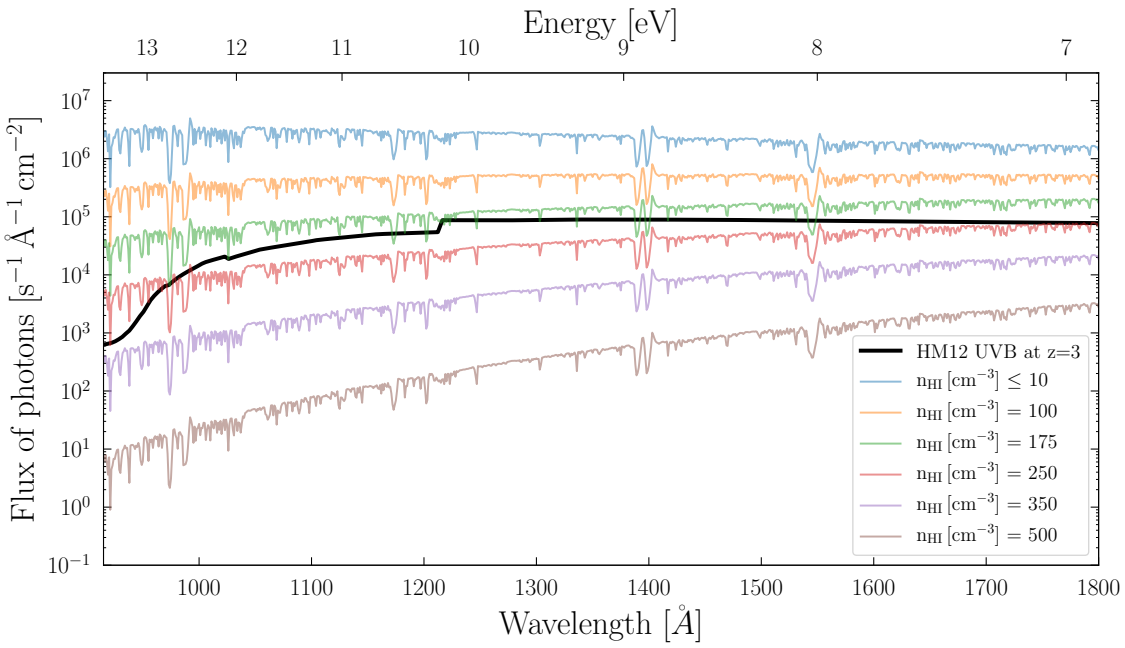

Figure 3.3: The black line represents the SED of the UV background taken from Haardt & Madau (2012) at redshift 3. The colored lines are spectra of a stellar population of 1000 solar masses at a distance of 10 pc, with a metallicity of 0.006 $Z_\odot$ and an age of 10 Myr, for different dust attenuations. The optical depth of dust is proportional to $n_{\mathrm{H\,I}}$ (Section 3.5).

dure is to modify the code to stop the gravity and hydrodynamics part and keep only the radiative field, since we want to keep the snapshot as it is, and only add a sub-ionising energy bin to the radiation. A similar approach has been used in Katz et al. (2021).

The advantage of this method is that the result is much more realistic than assuming an homogeneous sub-ionising radiation field. It is almost as if the simulation contained the sub-ionising bin itself, except that in post-processing the galaxy is frozen in time so that it is equivalent to having an infinite speed of light for the sub-ionising radiation. The major downside is that the design of Ramses forces us to use the same number of CPU cores for the restart as the number used during the actual simulation. This prevents the post-processing of simulations on small, fast and cheap clusters, since it requires access to large scale computer centers. The restart of a single snapshot is in general short, using only the time for photons to reach the outskirts of the halo from which they are emitted, but the CPU time can become large if one wants to post-process all the outputs of a simulation.

The post-processing of one snapshot of the zoom simulation yielded similar conclusions as in Section 3.2.1: Cells in low density environments have all their C^0, Mg^0 and Si^0 ionised, and dense cells do not contribute to the mock observation because of dust. Therefore, even if the new radiation bin affects dense cells differently than a UVB, the result is not observable.

3.2.3 Computing the contribution of each stellar particle in each cell with rays

A third method, more accurate than the first one, and not requiring hundreds of CPU-cores like the second one, is to compute the contribution of each stellar particle to the radiation field everywhere in the simulation, by tracing rays, that is to say compute the optical depth (of dust in our case, since sub-ionising radiation is only affected by dust) from the star to a given point by adding the contribution of every cell in between. The most complete way would be to trace rays from every star to every cell, but it is not possible since it would require to trace trillions of rays. However one can design a way to regroup sources (stars) and absorbants (cells), which exponentially decreases the

number of rays. Although I did not apply this method here, it could be computationally efficient, since they are used in other RHD simulation, on-the-fly at each time step (e.g. Grond et al., 2019). However I expect that observed absorption lines would not be affected by this complex treatment of the sub-ionising radiation because of dust, as in the previous two sections. This is why I use a homogeneous radiation field for the sub-ionising energies, as described in Section 3.2.1, in my fiducial post-processing pipeline.

3.3 Computing the ionisation fraction of any chemical element

Once the density of an element is known and the radiation field is computed at all relevant energies, we have all the information necessary to compute the ionisation fractions of this element, which are required to deduce the density of a particular ion. In this section I give an overview of the reactions that determine the ionisation fractions of a species in an interstellar plasma, along with the corresponding reaction rates from the literature.

3.3.1 Collisional ionisation Equilibrium

The so-called collisional ionisation equilibrium (CIE) is a theoretical situation where the ionisation state of a gas is determined solely by collisional excitation and by recombinations, ignoring other effects such as photoionisation or cosmic rays. Although it is in general incomplete, this model is useful to understand the first principles of the ionisation of a plasma. For this section I use Dopita & Sutherland (2003) as a reference.

Collisional ionisation

The way ions get ionised in this model is through collisional ionisation. In the following text I use the generic symbol "A" to designate any ion, in the ionisation state "i+". Collisional ionisation is a reaction of the form

$$\mathrm{A}^{i+} + \mathrm{e}^- \longrightarrow \mathrm{A}^{(i+1)+} + 2\,\mathrm{e}^-.$$

The electron that is used for the collision loses the energy corresponding to the ionisation energy of the ion, hence this process removes energy from the electron gas. As in all such collisional processes, the rate of the reaction, $R\,[\mathrm{cm}^{-3}\,\mathrm{s}^{-1}]$, is proportional to the product of the densities of all reactants and has the following form:

$$R^{A,i}_{\mathrm{coll}} = n_e n_{A,i} \int_{E_{\mathrm{min}}}^{\infty} \sigma_{\mathrm{coll}}(E) E f(E) \mathrm{d}E, \tag{3.2}$$

where n_e and $n_{A,i}$ are respectively the densities of electrons and ions, in cm^{-3}, E_{min} is the ionisation energy of A^{i+}, $\sigma_{\mathrm{coll}}(E)$ the cross-section of the reaction as a function of energy and $f(E)$ the distribution of the energy of electrons, that I assume to be Maxwellian. The cross-section is by far the most complicated part of this computation, and many years of work have been necessary to estimate it (for almost all ions), through both quantum physics developments and laboratory experiments. The complex results are then fitted with convenient analytical functions, fine-tuned to be applicable to any ion. In practice, the whole integral of Equation 3.2 is fitted by a reaction rate as a function of temperature $\alpha^{A,i}_{\mathrm{coll}}(T)$:

$$\int_{E_{\mathrm{min}}}^{\infty} \sigma_{\mathrm{coll}}(E) E f(E) \mathrm{d}E = \alpha^{A,i}_{\mathrm{coll}}(T)\,[\mathrm{cm}^3\,\mathrm{s}^{-1}].$$

One example of such a fit is from Voronov (1997), which I use in this work. It takes the form

$$\alpha^{A,i}_{\mathrm{coll}}(T) = A \frac{1 + P\sqrt{U}}{X + U} U^K e^{-U}, \tag{3.3}$$

where $U \equiv E_{\rm min}/(k_b T)$, and $A[{\rm cm}^3\ {\rm s}^{-1}]$, P, X and K are the fit parameters, given in Voronov (1997) for each ion of elements up to Nickel. In this paper, the authors explain that their fit functions are close to recommended rates coming from experimental measurements, with a deviation less than 10%. This is better than the estimated accuracy of the experimental data, which is 40-60 %.

Recombination

In CIE, the ionisation processes are counterbalanced by recombinations. Those are reactions of the form

$$\mathrm{A}^{(i+1)+} + \mathrm{e}^- \longrightarrow \mathrm{A}^{i+} + \gamma,$$

where a photon γ is emitted, which carries away the extra energy. Exactly the same kind of developments have been made as for collisional ionisation, leading to fitting functions that apply to almost all ions. The recombination rate is thus of the form

$$R_{\rm rec}^{A,i+1}(T)\,[{\rm cm}^{-3}\ {\rm s}^{-1}] = n_e n_{A,i+1}\,\alpha_{\rm rec}^{A,i}(T). \tag{3.4}$$

This process is actually slightly more complex than collisional ionisation, as it is the sum of two components, radiative and dielectronic recombination. While radiative recombination is simply an ion capturing a free electron, resulting in a new ion in its ground state, dielectronic recombination is different in that after the capture of the free electron, two electrons of the ions are out of their ground state, leading to radiative cascades. This process has a different cross-section than radiative recombination, and often dominates the recombination rate. More details are in Dopita & Sutherland (2003).

In this thesis I use the rates from Badnell (2006), which are also used in Cloudy. The radiative recombination term is as follows:

$$\alpha_{RR}(T) = \frac{A}{\sqrt{\frac{T}{T_0}}\left(1+\sqrt{\frac{T}{T_0}}\right)^{1-B(T)}\left(1+\sqrt{\frac{T}{T_1}}\right)^{1+B(T)}}, \tag{3.5}$$

where $B(T) = B + C\exp\left(-\frac{T_2}{T}\right)$, and A, B, C, T_0, T_1 and T_2 are the fit parameters for the different ions[1]. For most ions with charge two or more, we have $C = 0$, simplifying the expression. The dielectronic recombination rate is

$$\alpha_{DR}(T) = T^{-3/2}\sum_{i=1}^{9} c_i \exp\left(-\frac{E_i}{T}\right), \tag{3.6}$$

containing 18 fit parameters c_i and E_i[2]. However, the fewer electrons an ion has (i.e. the more ionised it is), the more of those parameters are equal to zero. According to Badnell (2006), those fitting functions accurately reproduce experimental data, with a deviation of less than 1%. For simplicity, let us call

$$\alpha_{\rm rec} = \alpha_{RR} + \alpha_{DR}.$$

Application of collisional ionisation equilibrium for H and He only: Ramses-RT, Krome and Cloudy

At a given temperature T, the system is in CIE if the time derivatives of all ion densities are zero, which means that the corresponding ionisation and recombination rates are all equal. For a model with only hydrogen and helium, this gives equations such as

$$0 = \frac{{\rm d}n_{\rm H\,I}}{{\rm d}t} = -n_{\rm H\,I} n_e\,\alpha_{\rm coll}^{\rm H\,I}(T) + n_{\rm H\,II} n_e\,\alpha_{\rm rec}^{\rm H\,I}(T), \tag{3.7}$$

[1] http://amdpp.phys.strath.ac.uk/tamoc/RR/
[2] http://amdpp.phys.strath.ac.uk/tamoc/DR/

and similar equations for He I and He II. The three equations are linked together via the electron density n_e. The system has six unknown variables, $n_{\mathrm{H\,I}}$, $n_{\mathrm{H\,II}}$, $n_{\mathrm{He\,I}}$, $n_{\mathrm{He\,II}}$, $n_{\mathrm{He\,III}}$ and n_e. The three equations of the type of Equation 3.7 are completed by the relations $n_{\mathrm{H\,I}} + n_{\mathrm{H\,II}} = n_H$, $n_{\mathrm{He\,I}} + n_{\mathrm{He\,II}} + n_{\mathrm{He\,III}} = n_{He}$ and $n_e = n_{\mathrm{H\,II}} + n_{\mathrm{He\,II}} + 2n_{\mathrm{He\,III}}$, where the total density of hydrogen and helium are given. Several codes offer the possibility to compute the evolution of those densities. Here are two examples of codes that do it, among many other things. Cloudy[3] (Ferland et al., 2017) is a photoionisation code that is designed to compute the transmitted spectrum of a one dimensional slab of gas and dust, or of a spherically symmetric system, illuminated by some input spectrum, at equilibrium. It also computes the ionisation equilibrium of the gas, as well as many other gas properties. Alternatively, Krome[4] (Grassi et al., 2014) is a software developed to compute the non-equilibrium evolution of a chemical network given any number of reactions and rates as a function of temperature. It does not allow to compute as many gas properties as Cloudy, but it is flexible, fast, and can be used in simulations to compute the chemical evolution at each timestep. Finally, Ramses-RT itself can be used to compute the ionisation equilibrium of a system of hydrogen and helium.

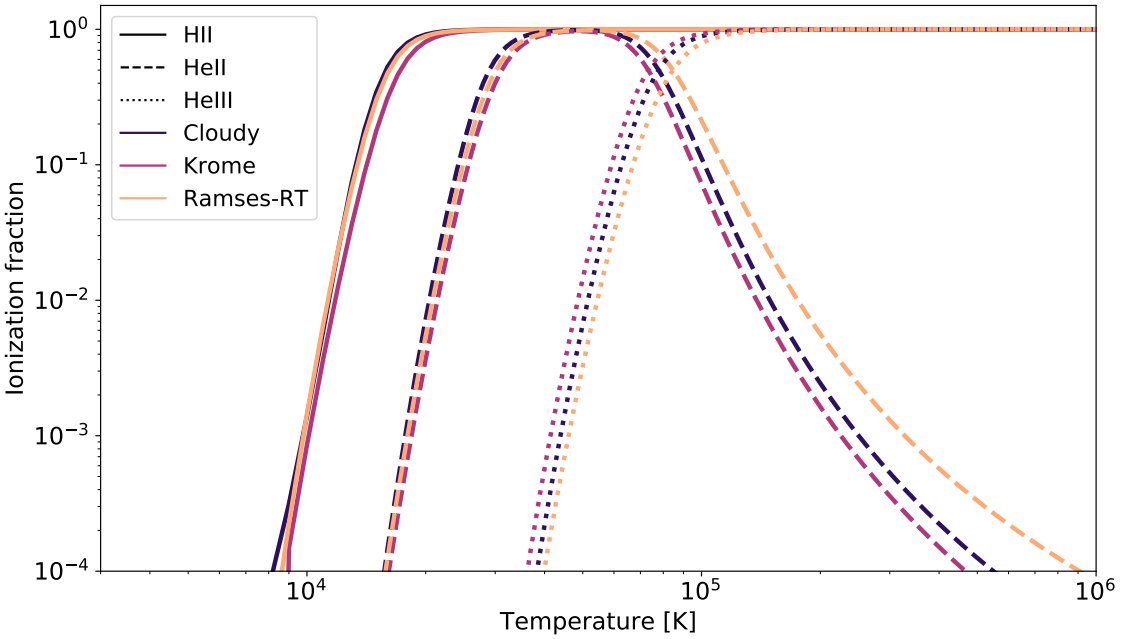

Figure 3.4: Hydrogen and helium ionisation as a function of temperature, in a collisional-ionisation equilibrium setup in Cloudy, Krome and Ramses-RT.

To study the variation of ionisation fractions when using different sources for the recombination and collisional ionisation rates $\alpha(T)$, I show in Figure 3.4 the ionisation fractions of a system containing only hydrogen and helium, computed by Cloudy, Krome and Ramses-RT. For Krome, we used the default rates that are given in the code. For example, the recombination rates for hydrogen are taken from Sasaki & Takahara (1993). The Cloudy recombination rates are from Badnell (2006), and the collisional ionisation rates come from an update of Voronov (1997) results, given in Dere (2007). The collisional ionisation rates used in Ramses-RT are from Cen (1992), while the recombination rates are from Hui & Gnedin (1997). Figure 3.4 shows that the curves of ionisation as a function of temperature match well even though they use different reaction rates.

[3] https://nublado.org
[4] http://kromepackage.org

Metals: Cloudy or Krome?

The treatment of a system in collisional ionisation equilibrium with heavy elements is similar to the case with only hydrogen and helium, only with more equations. When it comes to metals, Ramses-RT was not built to trace their chemical (ionisation) evolution, so we turn to Cloudy and Krome, which I compare in this section. More precisely, I begin by showing that the built-in rates in Krome do not match the results of Cloudy well, and then show how the modular nature of Krome allowed me to update its reaction rates until a good match was found. My motivation to use Krome even though it seemed to be different than the state-of-the-art code Cloudy is that Krome is orders of magnitude faster than Cloudy for the relatively simple task of computing ionisation fractions. I compare the speed of the two codes in Section 3.3.3.

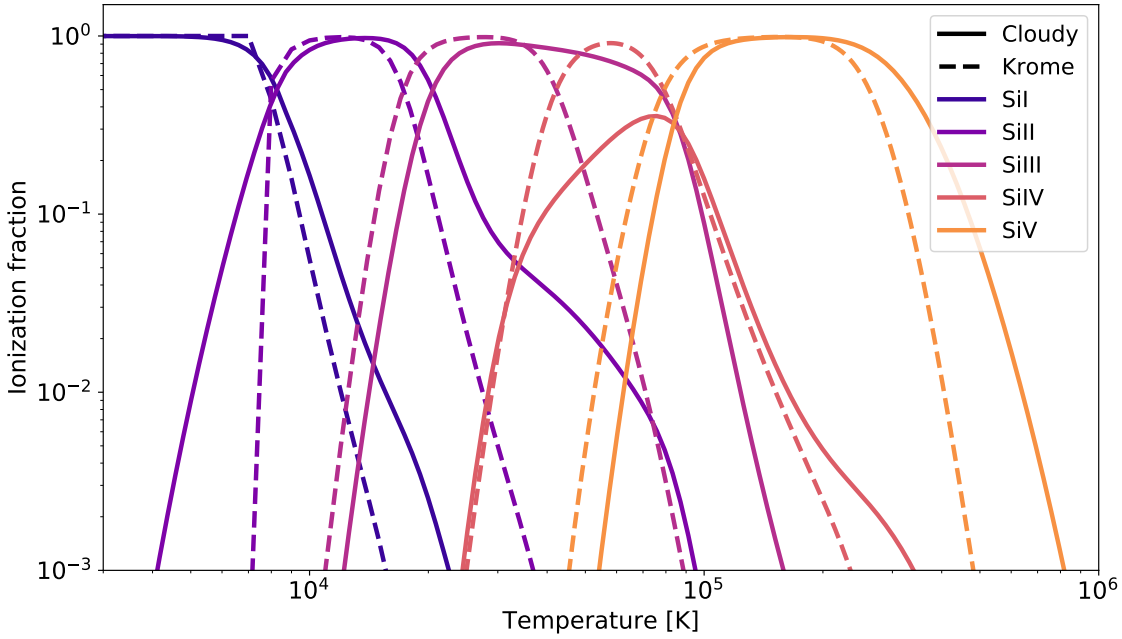

Figure 3.5: silicon ionisation fractions as a function of temperature. This is using the default rates of Krome, which lack dielectronic recombination rates, and do not match well with Cloudy.

I focus on the elements that I also adopted in Mauerhofer et al. (2021), silicon and carbon. Figure 3.5 shows the ionisation fractions of silicon, with the default rates of Krome. We see that there are major discrepancies with Cloudy. I found that this is because Krome used only radiative recombination rates and not dielectronic ones.

However, Krome uses chemical networks that can be fully customised by the user, so I added the dielectronic component of the recombination rate, following Badnell (2006). The new results are shown in Figure 3.6. There is a substantial improvement, but the match is still not perfect. It is now useful to look at another element, to see if I reach the same conclusion. Figure 3.7 is the same as Figure 3.6 but for carbon ionisation fractions. Here the match between Krome and Cloudy is perfect. How can this be explained?

I found the reason for this to be that silicon is more prone to charge transfer reactions, that were omitted from Equation 3.7, as for example

$$\mathrm{Si}^+ + \mathrm{H}^0 \longrightarrow \mathrm{Si} + \mathrm{H}^+,$$

or similar reactions with higher ionisation levels of silicon, or with helium instead of hydrogen. In general such reactions do not affect the ionisation equilibrium, except for specific cases like silicon. Exactly which of these reactions can occur in practice, and at which rate, is not an easy task to determine from the literature. It is hard to find sources, and programs such as Cloudy and

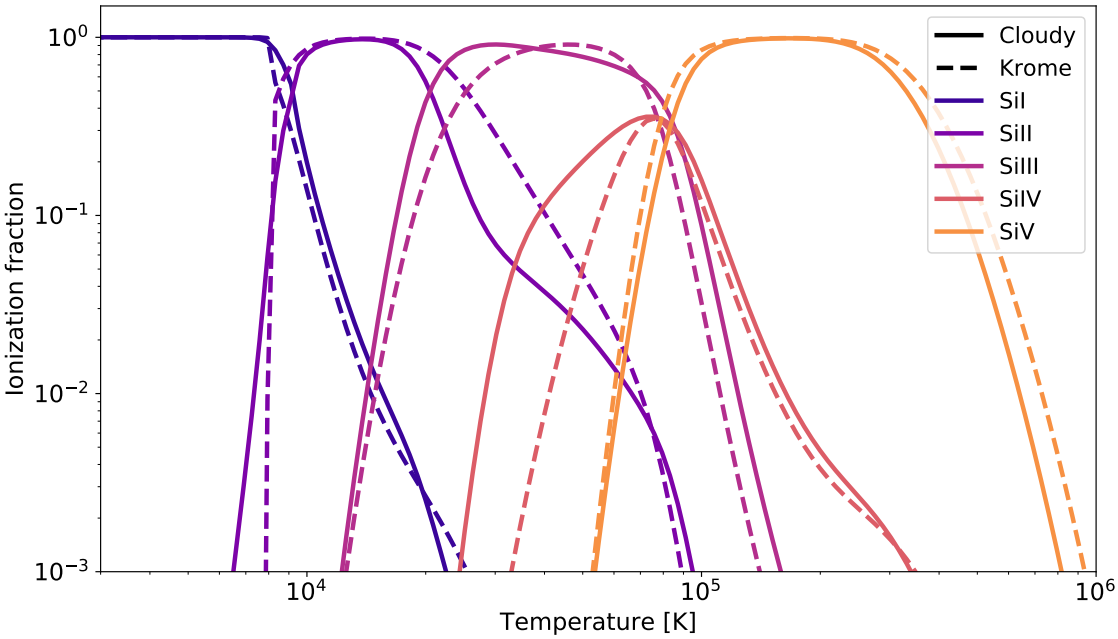

Figure 3.6: Silicon ionisation fractions as a function of temperature. This is using recombination rates from Badnell (2006) instead of the rates given by default in Krome.

Mappings (Sutherland et al., 2018) use different rates.By copying the rates present in Mappings into a Krome chemical network, I find the results of Figure 3.8. The curves are now matching well, although not as well as for carbon (which is not affected by charge transfer reactions). However, I do not include any charge transfer reactions in my pipeline to compute the ionisation fractions of metals in Ramses-RT simulations. Indeed, including reactions containing hydrogen or helium would alter the density of those elements, which is not desirable since the simulation treats them accurately, with out-of-equilibrium chemistry. From Figure 3.6, we see that omitting charge transfer reactions implies that I slightly overestimate the density of Si^+ at temperatures around 3×10^4 K. However, this does not impact the mock observations.

3.3.2 Photoionisation

Until now we have focused on the collisional ionisation equilibrium, but in reality there are other processes that affect the ionisation state of a plasma. The main one is photoionisation; the ionisation of species by photons with energies above their ionising energy. Simulations that I used during my PhD have the advantage of following the (hydrogen and helium) ionising radiation field on-the-fly, which can then be used to compute the photoionisation rates of species with ionising energies above 13.6 eV. However, as I have explained in Section 3.2, in Cloudy they lack the radiation at lower energies, which can still ionise many neutral heavy elements. The way I estimated this lower-energy radiation field is explained in Section 3.2.

Computing photoionisation rates

Given that we have the radiation field above and below 13.6 eV, I now present how to compute the photoionisation rates $\Gamma\,[s^{-1}]$ from the radiation field divided in n bins (typically 4 bins, 6-13.6 eV, 13.6-24.59 eV, 24.59-54.42 eV and 54.42-∞.). This is also explained in Appendix B of Mauerhofer et al. (2021), see Chapter 4.

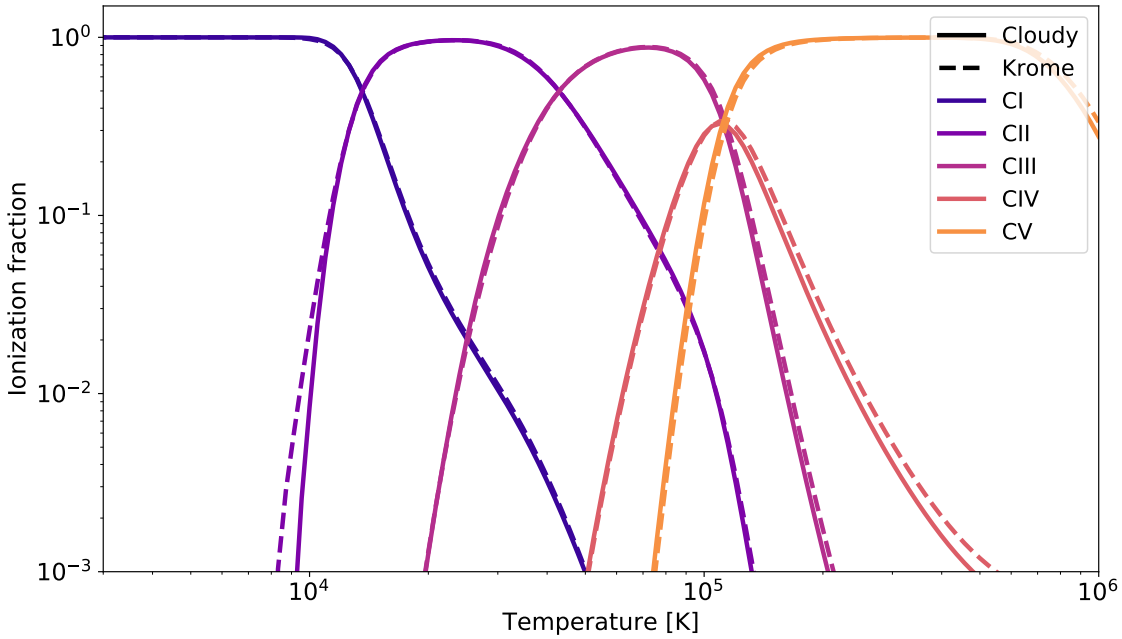

Figure 3.7: Same as Figure 3.6 but showing carbon ionisation fractions.

In general, the ionisation rate of a species X is equal to

$$\Gamma_X(\vec{x}) = \int_{E_{\min}}^{\infty} F_E(\vec{x}) \sigma_E^X \, \mathrm{d}E, \tag{3.8}$$

where $\vec{x}$ is the position, $E_{\min}$ is the ionising energy of X, E is the energy of a photon in eV, $F_E(\vec{x})$ is the flux of photons with energy E at position $\vec{x}$, in units of $\mathrm{s}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$, and σ_E^X is X's photoionisation cross-section at energy E. When applying this to a Ramses-RT simulation, the position $\vec{x}$ is discretised into gas cells, and the photon flux is separated into only four wavelength (or energy) bins, as mentioned above. Thus the cross-section also has to be discretised spectrally, in four bins, $\bar{\sigma}_i^X$, which will yield a new equation for Γ_X:

$$\Gamma_X^{\mathrm{cell}}[\mathrm{s}^{-1}] = \sum_{i=1}^{N_{\mathrm{bins}}} F_i^{\mathrm{cell}} \cdot \bar{\sigma}_i^X, \tag{3.9}$$

where "cell" is a given cell of the simulation, N_{bins} is the number of energy bins in the simulation and F_i^{cell} is the flux of photons in the cell, in units of $\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$, for the i^{th} bin of radiation.

To discretise the cross-section spectrally, one could simply define the average cross-section in each bin i,

$$\bar{\sigma}_i^X = \frac{\int_{E_i}^{E_{i+1}} \sigma_E^X \, \mathrm{d}E}{E_{i+1} - E_i}. \tag{3.10}$$

However, a more accurate solution is to take into account the shape of the radiation field of the simulation, by doing a weighted-average. Let us call this average radiation field F_E^{avg}. The cross-section then becomes

$$\bar{\sigma}_i^X = \frac{\int_{E_i}^{E_{i+1}} \sigma_E^X F_E^{\mathrm{avg}} \, \mathrm{d}E}{\int_{E_i}^{E_{i+1}} F_E^{\mathrm{avg}} \, \mathrm{d}E}. \tag{3.11}$$

More precisely, I adopt the method used in Ramses-RT for the photoionisation cross-section of hydrogen and helium, which is the following. First, we assign a photoionisation cross-section to each stellar particle in the galaxy in which we want to compute photoionisation rates. To do

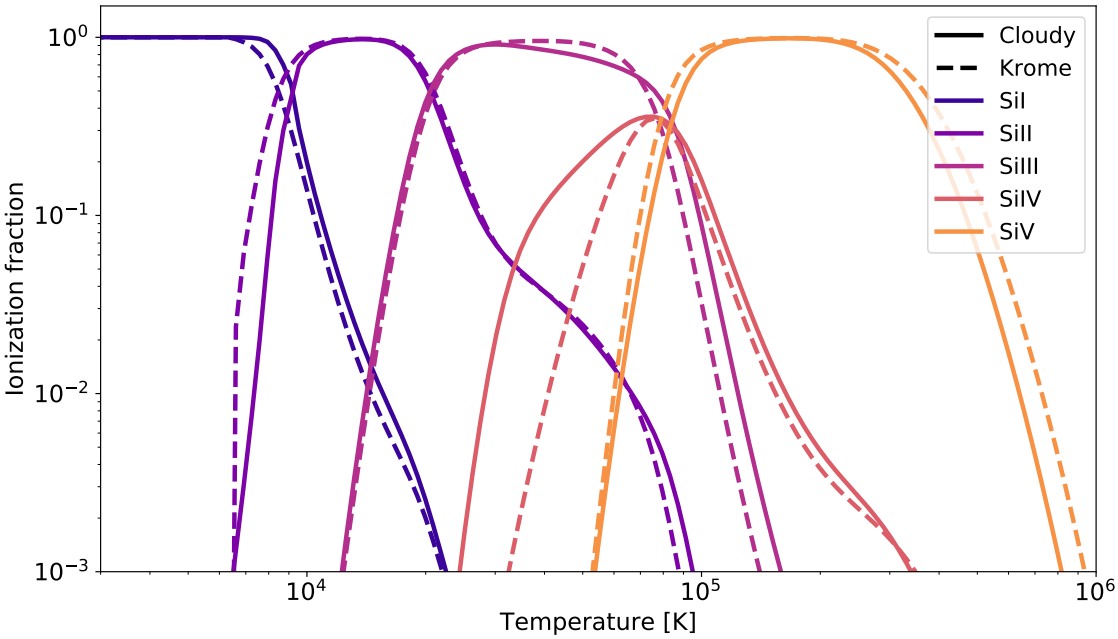

Figure 3.8: Same as Figure 3.6 but including charge transfer reactions, with rates taken from Sutherland et al. (2018).

so, we use the spectrum of the stellar particle (given by a spectral library, see Section 3.6), and apply it into Equation 3.11, which yields $\sigma^X_{i,\mathrm{star}}$. We then compute an average cross-section for the galaxy by taking the mean of all $\sigma^X_{i,\mathrm{star}}$, weighing by the luminosity of each star, in the energy bin considered:

$$\bar{\sigma}^X_i = \frac{\sum_{j_{\mathrm{star}}=1}^{N_{\mathrm{stars}}} \bar{\sigma}^X_{i,j_{\mathrm{star}}} \cdot L_i^{j_{\mathrm{star}}}}{\sum_{j_{\mathrm{star}}=1}^{N_{\mathrm{stars}}} L_i^{j_{\mathrm{star}}}}. \tag{3.12}$$

We can make sure that this way of taking an SED-weighted average cross-section is more accurate than taking the non-weighted average (as in Equation 3.10) by imagining a situation where the stars are emitting photons at a single energy E_0. In this case, our approximation yields $\Gamma^{\mathrm{cell}}_X = F_{E_0}\,[\mathrm{s}^{-1}\,\mathrm{cm}^{-2}] \cdot \sigma^X_{E_0}$, which is actually exactly equal to the real photoionisation rate. Using Equation 3.10 would not be exact in this extreme case. This reassures us that taking the shape of the ambient radiation field into account really increases the accuracy of the photoionisation rate.

Consequences for the ionisation equilibrium

Once the photoionisation rate is known, we can rewrite the equation for ionisation equilibrium, which is not anymore only "collisional" as in Equation 3.7:

$$0 = \frac{\mathrm{d}n_{\mathrm{H\,I}}}{\mathrm{d}t} = -n_{\mathrm{H\,I}} n_e\, \alpha^{\mathrm{H\,I}}_{\mathrm{coll}}(T) + n_{\mathrm{H\,II}} n_e\, \alpha^{\mathrm{H\,I}}_{\mathrm{rec}}(T) - n_{\mathrm{H\,I}}\, \Gamma_{\mathrm{H\,I}}, \tag{3.13}$$

and the same equations with any other ions present in the system. There is a major difference compared to CIE, which is that the photoionisation term is only proportional to the first power of density, while all other terms are to the second power. In CIE one can normalise all equations by the density, which means that the solutions are independent of the density. In CIE, only the temperature defines the ionisation equilibrium. However this is not anymore the case in Equation 3.13. As a consequence, the density is indeed important for the determination of the ionisation equilibrium in the presence of a radiation field. Dense gas is less affected by radiation, and more dominated by collisions and recombinations, while diffuse gas is significantly altered by an ionising radiation field. To clarify, not only dense gas can shield itself from radiation (photons do not

easily penetrate inside dense clouds), but even when there are photons in dense gas, their effect is often weaker than collisions. The transition between dense and diffuse gas in this context is highly dependent on the rates of recombination and ionisation of the considered ions.

3.3.3 Merging of Krome and the IO of Ramses-RT

I highlight here the main advantage of using Krome instead of Cloudy to compute the ionisation fractions in all cells of a Ramses-RT simulation: speed. Cloudy is ideal for getting many gas properties at the same time, not only ionisation fractions, but also line emissivities, level populations, cooling rates, etc. It can easily be used on non-RT simulations using interpolation tables. Typically, one would run Cloudy models over a grid of temperatures, densities and metallicities, and then interpolate for the actual values in the gas cells of simulations, assuming a uniform background for the radiation field. However, once a radiation field is added in the simulation, the number of parameters becomes too large to use tables. Running Cloudy cell by cell is also costly, from a few seconds to over 30 seconds per cell. It would easily require several hundred thousand CPU hours. It has however been done: Katz et al. (2019) used neural networks to extend the Cloudy results from a subset of $\sim$ tens of millions of cells, to the rest of the cells.

However when computing the ionisation fractions alone, the method I developed provides satisfying results in a short time. A few benchmark numbers for my Krome code: On the zoom simulation described in Chapter 2, it took about 5600 seconds with 32 cores (31 doing the job, the workers, and one telling the workers what to compute, the master), to compute the densities of C^0-C^{3+}, Mg^0-Mg^+, Si^0-Si^{3+} and Fe^0-Fe^+ over around 80 millions cells. This means that it takes approximately 2.5 CPU milliseconds per cell to get the density of twelve species, including the initialisation time of the code (reading the stars to compute the SED-weighted average cross-section, opening the AMR files, etc.). This time is even reduced when computing the densities of a smaller number of ions. On the Sphinx20 simulation, at snapshot 360 (redshift 5.11), which contains 1.84 billion cells, it took 1.82 CPU milliseconds per cell to compute the densities of four ions, Si^0-Si^+ and Fe^0-Fe^+. Finally, when using another queue on the "Lesta" cluster at the Observatory of Geneva, the code was even faster, with an average of 0.5 CPU milliseconds per cell, to compute the densities of four ions. To summarise, Krome is around 10'000 times faster than Cloudy to compute ionisation fractions.

However, there are a few caveats of this code, especially concerning the convergence to equilibrium. As stated before, the computation of the ionisation fractions is done in post-processing and so we are forced to get equilibrium ionisation fractions (except for hydrogen and helium, for which we keep the values from the simulation, out-of-equilibrium). For the metals, we ask Krome to converge to equilibrium, which sometimes fails. I use the "krome_equilibrium" subroutine of the code, that I modified to adapt the timestep to the gas density of each cell and to print diagnostics when something goes wrong. This routine starts from a fixed timestep Δt of one thousand years and evolves the densities (for example of carbon ions) during this time, and repeats the process until the densities are stable. Δt is multiplied at each cycle by a number slightly above 1. After performing several tests, this was the best solution. Choosing Δt of the age of the Universe right away was causing more bugs. However, there are still problems for cells with the lowest densities. Luckily those cells do not impact the observables significantly. However, the code sometimes fails to find an equilibrium, either because of a cyclic change of densities, thus not converging, or because of numerical bugs (sometimes matter is not conserved). In the future, this problem could be tackled by developing more ingenious algorithms, which was beyond the scope of this thesis. Finally, I noticed more problems when increasing the number of ions, which is why I recommend keeping the number of ions below 12-15, to avoid crashes, and to gain time. After some tests it does not change any observable properties to just compute one or a few ions at a time, which is also faster.

3.4 The fine-structure levels and fluorescence

A common feature of many metallic ions is that their ground-state is split into several spin levels. Typical examples are C^+ and Si^+, which have two levels, that we call the true ground state and the fine-structure level, while O^0 has three levels and Fe^+ has nine. There are also ions with a single level, the true ground state, such as Mg^+, Al^+ and high ionisation ions such as C^{3+} or Si^{3+}. The number of spin levels depends on the atomic configuration of the ground state. For example, an s1 configuration has only one level, a p1 has two or p4 has three. The main effect of those spin-levels is what we call *fluorescent emission*. This is a crucial feature of metallic absorption lines, and has to be taken into account for the creation of mock spectra. We explain the implementation of fluorescent lines in Section 3.7.

Here are the mechanisms of resonance and fluorescence: When an ion with a fine-structure level absorbs a photon at the wavelength of the absorption line, it can deexcite either to the true ground state or to a fine-structure level. The former usually has a higher probability, around 85%, and leads to resonant scattering. The latter however leads to the emission of a photon at a wavelength longer than the absorption line, which prevents it from being absorbed again by an ion in the ground-state, thus leaving resonance. Those photons will then be seen as an excess emission on the red side of the line. This is fluorescent emission. The structure of ions is illustrated in Figure 3.9, with the example of Si^+.

It turns out that fine-structure levels can affect the line transfer in an additional, more complex way than just a fluorescent emission. During the analysis of the zoom simulation, a paper was published (Jaskot et al., 2019) in which they state that a non-negligible fraction of ions (mainly C^+ and Si^+) in the ISM of galaxies are not in their true ground-state but are rather populating their fine-structure levels (level 2 in Figure 3.9), and that this affects the observed profile, by creating an absorption feature at the wavelength of the fluorescence. This is especially the case in denser parts of the ISM, where collisions with electrons excite the ions in this spin level, which is long lived. This led me to add an additional layer of complexity to my absorption line modelling, by computing the fraction of ions populating their fine-structure levels, and including all the possible channels of the ions, that are all shown and explained in Figure 3.9 for Si^+. Note that absorption lines at the wavelength of fluorescence had already been observed before, for example in quasar absorption lines (e.g., Wolfe et al., 2003), down-the-barrel observations (Lebouteiller et al., 2013) or GRB afterglow absorption lines (e.g., Gatkine et al., 2019).

Similarly as discussed for ionisation fractions in Section 3.3, Cloudy is the most powerful tool for computing the level populations, taking into account the radiation field and many other parameters, but is extremely slow and works as a black box. Since Krome is not adapted to deal with fine structure populations, I used the python package PyNeb[5] (Luridiana et al., 2015). Even though it does not take as many physical processes into account as Cloudy (for example there is no radiation field involved), it is efficient for a first approximation. The level populations in PyNeb depend on two parameters, the electron temperature and density, which are already known in all cells of the simulations. Since only two variables are involved, it is practical to use interpolation tables (which is not what I did for the first paper, but developed later).

[5] https://pypi.org/project/PyNeb/

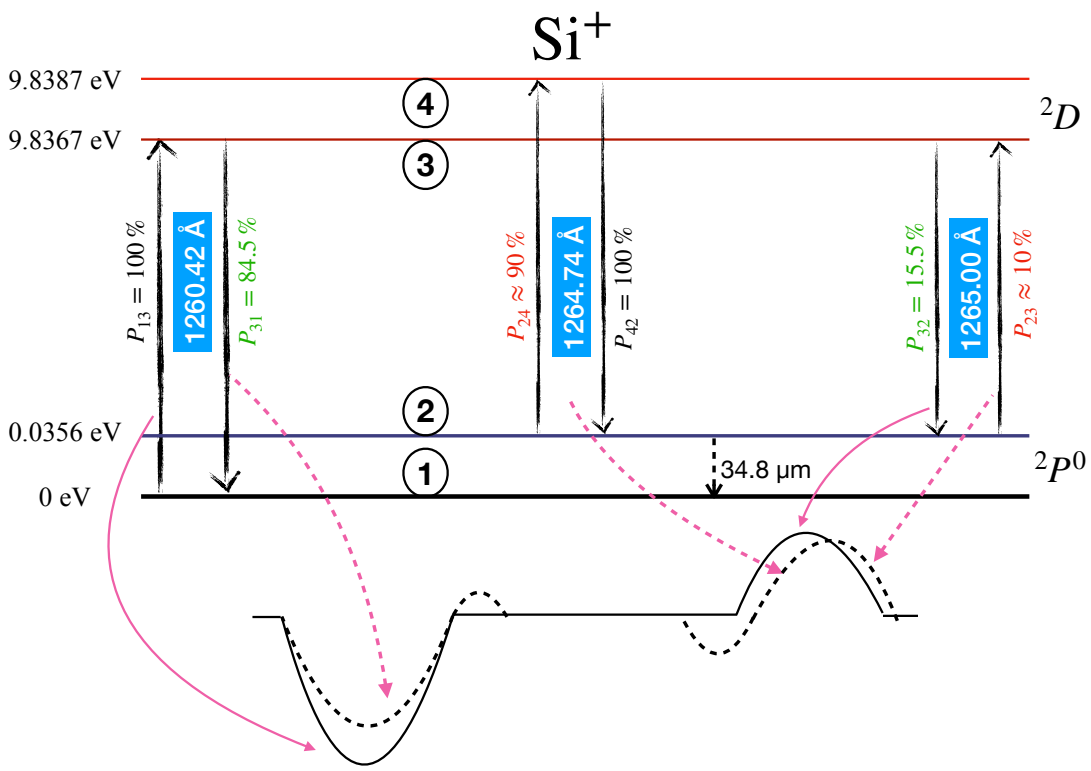

Figure 3.9: Structure of the Si^+ ion around the Si II λ1260 line. *Arrows at 1260.42* Å: the only channel of interaction starting from the true ground-state is the one that arrives at level 3. This absorption of a photon creates the main feature of Si II λ1260, the absorption trough on the bottom left (solid line). With a probability of 84.5%, the ion deexcites back to the ground-state, reemitting a photon at 1260.42 Å. This creates a resonant effect, leading to an *infilling* of the absorption, sometimes with a small P-cygni feature (emission bump on the right-side of the absorption, dashed line). *Arrows at 1265.00* Å: however, after an absorption from level 1 to 3, the ion has also a probability 15.5% to deexcite to the fine-structure level of the ground state, i.e., level 2. This emits a photon with a redder wavelength, likely to leave the resonance, and creating the fluorescent emission (solid line on the right). *Arrows at 1264.74* Å *and rightmost arrow*: since there are ions in the ISM populating the fine-structure level, it is possible to have absorptions from level 2. The most likely channel is the one going from level 2 to level 4, always leading to a decay back to level 2. There is a probability of around 10% that a photon with wavelength between 1264.74 Å and 1265 Å excites the ion from level 2 to level 3 instead of level 4. In this case, the ion will deexcite like after a "normal" absorption from level 1 to level 3. This absorption and reemission around the fine-structure level sometimes creates an absorption component at the wavelength of the fluorescence, and tends to shift the fluorescent line to the red side (Mauerhofer et al., 2021).

In Figure 3.10 I show in the left panel the fraction of C^+ ions that are excited in their fine-structure level as a function of electron density and temperature. We see that for a non-negligible fraction of ions to be in this level, the temperature must be at least 25 [K] and the density 0.03 [cm^{-3}]. The fraction can be high even at temperatures above 10^5 [K], but in this regime the density of C^+ itself becomes very small due to collisional ionisation, as we saw in Figure 3.7. The fraction reaches its maximum of 66% in the region of the plot above a density of $\sim$ 20 [cm^{-3}] and temperatures between 100 [K] and 10^5 [K]. Those are quite high density regions, which are in general neutral in our simulations. The right panel shows the comparison between silicon and carbon, i.e. the ratio between the fraction of Si^+ populating the fine-structure level and the same fraction for C^+. The first thing to notice is that the fraction of Si^+ in its fine-structure level is always smaller than the fraction of C^+. The maximum of this ratio is around 95%, for very high densities and temperatures

between 1000 and 10^5 [K]. I also tested iron and only a negligible fraction of Fe^+ is not in the true ground state.

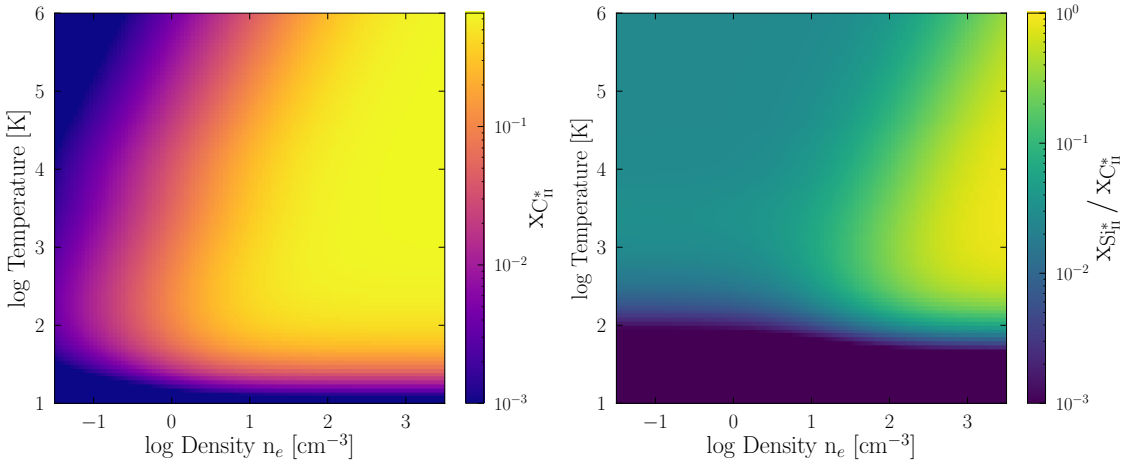

Figure 3.10: Left panel: Fraction of C^+ ions that are populating their fine-structure level, as a function of electron density and temperature. Right panel: Ratio of fraction of Si^+ ions populating their fine-structure level over the same for C^+.

To see the effect that the absorptions from the fine-structure level has on the observed spectra, I refer the reader to Section 3 of Mauerhofer et al. (2021) (Chapter 4). In summary, they do not alter the absorption lines, but modify the equivalent width of the fluorescence in a complex way, and shift the fluorescent line to redder wavelengths.

3.5 The distribution and opacity of dust, a central ingredient for observables

In this work we adopt the dust model of Laursen et al. (2009), where the dust optical depth is the product of a pseudo-density of dust and a cross-section. The pseudo-density is computed as follows:

$$n_{\text{dust}} = \frac{Z}{Z_0} \times (\text{n}_{\text{H I}} + f_{\text{ion}} \text{n}_{\text{H II}}), \tag{3.14}$$

where Z_0 and f_{ion} are free parameters, the latter determining how much dust there is in ionised regions. We use fiducial values of $Z_0 = 0.005$ and $f_{\text{ion}} = 0.01$. The optical depth of dust in each cell of a simulation is then given by

$$\tau_{\text{dust}}(\lambda) = n_{\text{dust}} \sigma_{\text{dust}}(\lambda) d, \tag{3.15}$$

where σ_{dust} is the cross-section taken from Gnedin et al. (2008), shown in Figure 3.11, and d is the distance crossed by a photon in the cell.

This cross-section of the interaction of a photon with dust is the sum of an absorption component and a scattering component. The albedo is the parameter that represents the probability that the photon is scattered rather than destroyed. In practice, if there is an interaction with dust during the Rascas run, a random number r_1 is selected between 0 and 1. If $r_1 <$ albedo, the photon is scattered, and it is destroyed otherwise. The albedo as a function of wavelength is represented with the purple data points in Figure 3.11.

Finally, the asymmetry parameter g is used to determine the outgoing direction of propagation of a photon after it scatters on a dust grain. When a scattering event happens, with an initial direction of the photon $\vec{k}_{\text{in}}$, the outgoing direction $\vec{k}_{\text{out}}$ is described by two angles θ and ϕ. The angle ϕ

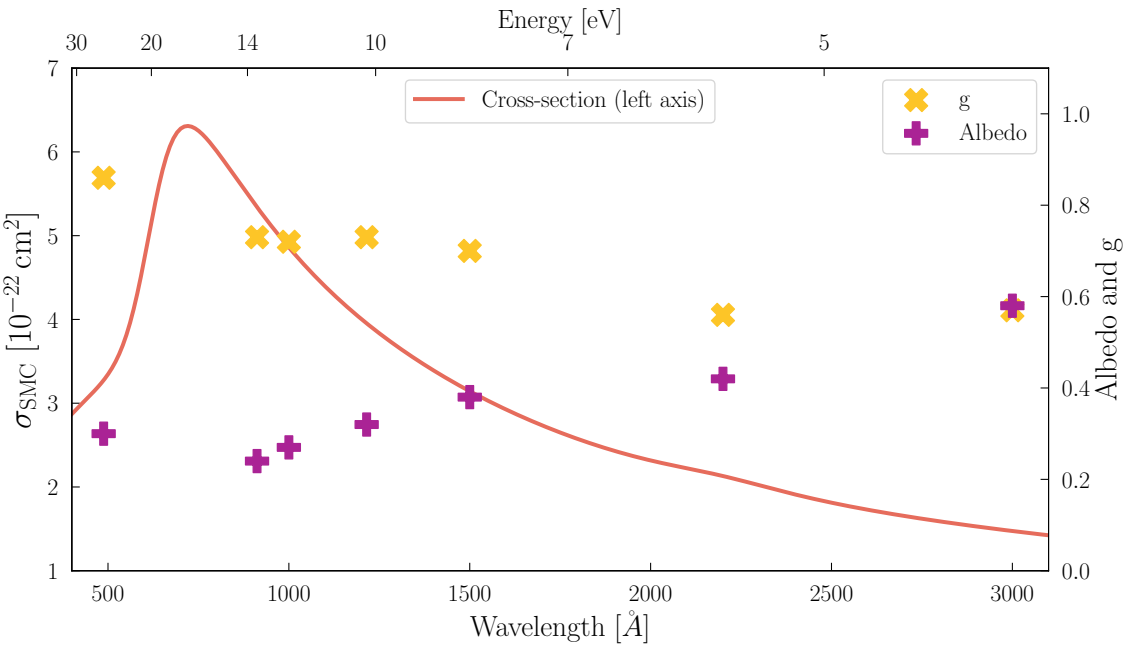

Figure 3.11: Dust cross-section, albedo and asymmetry parameter g as a function of wavelength (or energy). The cross-section, modelled from studies of the Small Magellanic Cloud (SMC), is the orange line with values on the left y-axis and comes from Gnedin et al. (2008). The albedo and g parameters, purple and yellow symbols, with values on the right y-axis, are from Li & Draine (2001).

describes the orientation of the plane containing $\vec{k}_{\rm in}$ and $\vec{k}_{\rm out}$, and has a uniform distribution between 0 and 2π. The angle θ, for which we use the variable $\mu \equiv \cos\theta$, describes the angle between $\vec{k}_{\rm in}$ and $\vec{k}_{\rm out}$, and depends on what is called the Henyey-Greenstein function $P_{\rm HG}(\mu)$ (Henyey & Greenstein, 1941):

$$P_{\rm HG}(\mu) = \frac{1-g^2}{2\,(1+g^2-2g\mu)^{\frac{3}{2}}}. \tag{3.16}$$

When $g = 0$, the angle θ also follows a uniform distribution, and when $g = 1(-1)$, $\vec{k}_{\rm out}$ is parallel (anti-parallel) to $\vec{k}_{\rm in}$. The values of g as a function of wavelength are plotted in Figure 3.11, taken from Li & Draine (2001).

Again, to derive the value of μ for a given scattering event, we first have to compute the cumulative distribution function $P(<\mu) = \int_{-1}^{\mu} P_{\rm HG}(\mu_1)\,{\rm d}\mu_1$. This can be done analytically, and gives

$$P(<\mu) = \frac{1-g^2}{2g\sqrt{1+g^2-2g\mu}} - \frac{1-g}{2g}.$$

Then one has to draw a random number r_1 between 0 and 1, and inverse the function, i.e. find μ such that $r_1 = P(<\mu)$. Finally, this results in

$$\mu = \frac{1}{2g}\left(1+g^2-\left(\frac{1-g^2}{1-g+2gr_1}\right)^2\right), \tag{3.17}$$

which, along with ϕ, determines $\vec{k}_{\rm out}$.

I discuss in detail the implications of this model in our simulations in Chapter 5.

3.6 Emitting light from stellar particles

Until now I have presented how to treat the gas in the simulation so that we have all the necessary data to model its interactions with photons. It is now time to switch to the other side of the picture and describe how we model the light itself.

Stars in Ramses-RT simulations are treated as particles that are formed when the right conditions are met in the ISM (see Section 2.1.6). Those particles have three properties: age, metallicity and mass. Using those properties, one can get a full spectrum of light for each particle using a stellar SED library. Typical examples, used in e.g. Rosdahl et al. (2018), are Galexev (Bruzual & Charlot, 2003) or BPASS (Eldridge et al., 2008; Stanway et al., 2016). Starburst99 (Leitherer et al., 2014) is another example. They are complex codes which compute the evolution of stars of different masses and metallicities over their lifetime. BPASS also has the additional feature to include the evolution of binary stars, which can lead to different SEDs compared to using only single stars.

Our main goal being the study of absorption lines imprinted by the ISM and CGM on the stellar continuum, we adopt a method to smooth the intrinsic stellar spectra to prevent contamination from stellar absorption features in the SED models. For this we use a Python code developed by Jérémy Blaizot which fits a region the stellar continuum (for each metallicity and age) using a power-law approximation of the form

$$F_\lambda = F_0 \left(\frac{\lambda}{\lambda_0}\right)^\beta , \tag{3.18}$$

where F_λ is the luminosity in erg s^{-1} Å^{-1} M$_\odot^{-1}$, F_0 is the fitted luminosity parameter, β the fitted slope and λ_0 [Å] a reference wavelength, the choice of which has no impact on the numerical results. In practice, for absorption lines, we need only narrow wavelength ranges, so this power-law approximation is equivalent to having a flat continuum. There are spectral regions where stellar absorption lines are strong, like around Lyα or C IV $\lambda\lambda$1548, 1550, which could potentially affect the power-law fit. Hence there is the option to remove those regions with strong absorption lines from the fit. This is what I use in Mauerhofer et al. (2021), even though there is no strong stellar features around the wavelength of LIS lines. The power-law fit allows us to obtain spectra consisting of only two parameters (normalisation and slope) for each metallicity and age in the stellar library. The result is written in what we call a "power-law table".

Once this table is done, we are ready to compute the photon initial conditions that will then be used by Rascas to run the radiative transfer (Section 3.7). The main idea is to choose a number $N_{\rm phot}$ of so-called "photon packets", which will sample the luminosity of the stars in the galaxy, and which all represent a number of photons per second. I do this with a code from the Rascas distribution: "PhotonFromStars", which begins by reading all the stellar particles in a given domain of the simulation. There are two main options for this code, "PowLaw", to use the power-law approximation I described above, or "Table" to use the stellar SED directly, without smoothing, which is useful when one wants to include the stellar absorption features. I focus here on the first method, which I used more extensively for my work. PowLaw assigns to each stellar particle a beta slope β and a normalisation flux f_0, by interpolating on the closest ages and metallicities available in the stellar library. The next step is to compute the total number of photons emitted per second for each stellar particle, $\dot{N}_{\rm star}$, in the chosen wavelength range $\lambda_{\rm min}$ to $\lambda_{\rm max}$. This gives

$$\dot{N}_{\rm star}\,[{\rm s}^{-1}] = \frac{M_{\rm star} F_{0,\rm star} \lambda_0^2}{10^8\, hc} \begin{cases} \ln\left(\frac{\lambda_{\rm max}}{\lambda_{\rm min}}\right) & \text{if } \beta_{\rm star} = -2 \\ \frac{1}{2+\beta_{\rm star}} \left(\left(\frac{\lambda_{\rm max}}{\lambda_0}\right)^{2+\beta_{\rm star}} - \left(\frac{\lambda_{\rm min}}{\lambda_0}\right)^{2+\beta_{\rm star}} \right) & \text{otherwise.} \end{cases} \tag{3.19}$$

where $M_{\rm star}$ is the mass of the stellar particle in solar masses and $F_{0,\rm star}$ and $\beta_{\rm star}$ are the power-law parameters of the stellar particle.

The code attributes the $N_{\rm phot}$ photon packets to the stellar particles based on their luminosity $\dot{N}_{\rm star}$, using Monte-Carlo sampling. To do so, we first build a list lum$_{\rm i}$ of the luminosities of every

star, normalised by the total luminosity:

$$\mathrm{lum}_{\mathrm{i=1}}^{\mathrm{n_{star}+1}} = 0,\ \frac{\dot{N}_{\mathrm{star},1}}{L_{\mathrm{tot}}},\ \frac{\dot{N}_{\mathrm{star},1} + \dot{N}_{\mathrm{star},2}}{L_{\mathrm{tot}}}, ...,\ \frac{\sum_{\mathrm{i=1}}^{n_{\mathrm{star}}-1} \dot{N}_{\mathrm{star,i}}}{L_{\mathrm{tot}}},\ 1.0 \tag{3.20}$$

where n_{star} is the number of stellar particles, and L_{tot} is the sum of all luminosities. This can be seen as a list, normalised between 0 and 1, of the cumulative contribution of all stars to the total luminosity. A photon packet will have a higher probability to be launched from a star with a higher contribution. To assign a photon packet to a star, we first pick a random number r_1 between 0 and 1[6], and then find the largest index i in our "lum" list such that $\mathrm{lum_i} < r_1$. This index i is the index of the star from which the photon will be launched. In practice, the index is computed with a bisection method, which is on average the fastest way to find this index.

To recapitulate, we have the initial position of the photon packets, which is the position of the star particle it was assigned to, as well as the initial direction of propagation, chosen randomly from an isotropic distribution. Now, the last step is to determine the wavelength of the photon, based on the power-law parameters of the assigned star particle. To do this, we start by computing the cumulative distribution function for the power-law spectrum of a given star. That is to say, given a wavelength $\lambda \in [\lambda_{\mathrm{min}}, \lambda_{\mathrm{max}}]$, what is the probability $P(<\lambda)$ that a random photon emitted by the star has a wavelength between λ_{min} and λ. This is computed similarly as for Equation 3.19, by doing an integral of the number of photons per unit wavelength, $\dot{N}_\lambda$, which results in:

$$P(<\lambda) = \begin{cases} \frac{\ln\left(\frac{\lambda}{\lambda_{\mathrm{min}}}\right)}{\ln\left(\frac{\lambda_{\mathrm{max}}}{\lambda_{\mathrm{min}}}\right)} & \text{if } \beta_{\mathrm{star}} = -2 \\ \frac{\lambda^{2+\beta_{\mathrm{star}}} - \lambda_{\mathrm{min}}^{2+\beta_{\mathrm{star}}}}{\lambda_{\mathrm{max}}^{2+\beta_{\mathrm{star}}} - \lambda_{\mathrm{min}}^{2+\beta_{\mathrm{star}}}} & \text{otherwise.} \end{cases} \tag{3.21}$$

Then, we draw a random number r_1 between 0 and 1, and find the wavelength for which $P(<\lambda) = r_1$, which is equivalent to inverse the function $P(<\lambda)$. We find:

$$\lambda = \begin{cases} \lambda_{\mathrm{min}} \left(\frac{\lambda_{\mathrm{max}}}{\lambda_{\mathrm{min}}}\right)^{r_1} & \text{if } \beta_{\mathrm{star}} = -2 \\ \sqrt[2+\beta_{\mathrm{star}}]{\lambda_{\mathrm{min}}^{2+\beta_{\mathrm{star}}} + r_1 \left(\lambda_{\mathrm{max}}^{2+\beta_{\mathrm{star}}} - \lambda_{\mathrm{min}}^{2+\beta_{\mathrm{star}}}\right)} & \text{otherwise.} \end{cases} \tag{3.22}$$

This determines all the initial properties of continuum photons launched from stellar particles with the PowLaw method.

This process can be used in a similar fashion for initial conditions of ionising photons, which is useful for making mock spectra of Lyman continuum with Rascas, in order to compute the escape fraction of ionising photons in post-processing, as I do in Mauerhofer et al. (2021). The only difference is that I take a larger wavelength range, typically from 10 to 920 Å, and hence the spectrum is not approximated by a power-law. Rather, Rascas has a "Table" routine that directly uses the SED from the library without fitting it to a power-law. As a consequence, there are no more analytical expressions for both the total luminosity of the stellar particles, contrary to Equation 3.19, and the wavelength to draw from a random number as it is done in Equation 3.22. The code has to employ more sophisticated numerical schemes to replace those two equations.

3.7 The transfer of photons with Rascas

At this point we are ready to transfer the photons in the simulation using Rascas (Michel-Dansac et al., 2020). We want to propagate photons, or photon packets, in a domain of the simulation. This is called the "computational domain" and can take the shape of a sphere, a shell, a cube or an

[6] more details on the random number generation in Michel-Dansac et al. (2020)

infinite slab. The photons' initial conditions, position, wavelength and direction of propagation, were previously determined by one of the modules of Rascas, as explained in Section 3.6 for the case of stellar continuum. The next step is to propagate all the photons, and the run is finished when each photon was either destroyed by dust or escaped the domain. The modelling of dust and even the light-dust interactions were described in 3.5, so I now explain the propagation of photons through the AMR grid and the interaction with the ions, taking Si^+ as an example. First, to understand the goal of Rascas, here is a list of what the code is outputting for each photon packet:

1. Status: did the photon escape the domain or was it destroyed by dust?
2. Position of last scattering
3. Final wavelength
4. Final direction of propagation
5. Number of scattering events
6. Physical time of propagation (not computation time).

These quantities allow the user to build a spectrum of the galaxy as a whole, averaged over the full 4π solid angle, to study the statistics of the number of scattering events or build a healpix map of the outgoing photons (map of a sphere projected on a 2d oval, like the Planck CMB map of Figure 1.1). However, to do an actual "mock observation", i.e. simulate an observer somewhere in the Universe taking the spectrum of the galaxy from a fixed angle of observation, this output alone is not practical. It is in principle possible to select only the photon packets whose final direction of propagation is close enough to the direction of the observer, but there is a more efficient and accurate way: the peeling-off algorithm. Therefore, after explaining the functioning of Rascas, I present this algorithm and how it is implemented in our code.

3.7.1 Rascas main program

Rascas propagates photons by doing a series of computations of optical depths, probabilities of interactions, and outcomes of interactions. At the start of the propagation of one photon, and after each scattering event, the photon has a position in a cell, a wavelength, a direction of propagation, and an assigned optical depth at which it will interact next. The optical depth measures how much a medium absorbs photons. More explicitly, for a photon with wavelength λ going through a medium with optical depth τ_λ, the probability to not be absorbed is $e^{-\tau_\lambda}$. Rascas uses Monte Carlo sampling, so in order to know when a photon packet will interact, the usual three steps method is applied: drawing a random number r_1 between 0 and 1, integrate the probability function $e^{-\tau_\lambda}$ and inverse it. The cumulative distribution function is: $P(<\tau) = \int_0^\tau e^{-\tau_1}\,\mathrm{d}\tau_1 = 1 - e^{\tau}$. Then we look for τ such that $r_1 = P(<\tau)$. This gives an optical depth of $\tau_{\rm abs} = -\ln(1 - r_1)$. This $\tau_{\rm abs}$ is the optical depth after which our photon packet will interact. Following are the steps to compute the propagation of a photon packet.

1. From its position, determine in which cell of the simulation the photon packet is, and retrieve the gas density, velocity and temperature (or Doppler width).
2. Determine whether the cell is fully in the computation domain and compute the distance to the cell border along the direction of propagation, or to the domain border if it is shorter.
3. Compute the total optical depth $\tau_{\rm cell}$ due to dust and spectral lines along this distance.
4. Determine whether there is an interaction in the cell. If $\tau_{\rm cell} < \tau_{\rm abs}$, there is no interaction, otherwise there is.
5. If there is no interaction and the border of the domain is closer than the border of the cell, the photon leaves the computational domain and its propagation is finished. If the border of the cell is closer than the border of the domain, the photon goes on to the next cell, and the process starts over from 1 with a new position and an updated $\tau_{\rm abs} = \tau_{\rm abs} - \tau_{\rm cell}$.

6. If at 5, there is an interaction, i.e. $\tau_{\text{cell}} > \tau_{\text{abs}}$, determine what the photon is interacting with. There is a channel of interaction for dust, and one for each spectral line included in the transfer. Among the n channels c_n, one is chosen by drawing a random number r_1 and looking at the largest index i such that $\sum_{j=1}^{i} \tau_{c_j} < r_1$.
7. Determine the outcome of the interaction, i.e. whether the photon survives, and at which wavelength and direction of propagation it is reemitted. In case of absorption by dust, as we saw in Section 3.5, the photon can either be destroyed, in which case its propagation is finished, or scattered. In case of an interaction with a spectral line, the photon can be reemitted resonantly, or through a fluorescent channel (more details below).
8. If the photon survives the interaction, restart the process at 1. with an updated position, wavelength, direction of propagation and $\tau_{\text{abs}}(r_1)$, with a new random number r_1.

Some of the steps require more explanation, in particular how the optical depth of interaction with a spectral line is computed, and how the outcome of such an interaction is determined.

To compute the optical depth, several quantities have to be known. First, the atomic parameters of the lines: the wavelength λ_{line} [Å], the Einstein coefficient A_{line} [s^{-1}] and the oscillator strength f_{line}. The Einstein coefficient is a measure of the rate at which an excited state of an ion, which has absorbed a photon, decays back to a lower level through some channel (there can be several possible channels, as seen in Figure 3.9), and will also be important for the determination of the outcome of an interaction between a photon and an ion. The oscillator strength f_{line} is a number that dictates the propensity of an ion to absorb photons through the channel corresponding to the line.

Next, we need to know the wavelength of the photon packet in the reference frame of the cell. The Doppler effect will impact the optical depth depending on the velocity of the gas in the cell. There is also a Doppler effect during the setup of the photon's initial conditions, due to the velocity of the stars. Since the equations of the Doppler effect apply more directly to frequencies than to wavelengths, I switch to the former quantities. Let us call ν_0 the frequency of the photon obtained from the method described in Section 3.6, in the frame of the star that emitted the photon, $\vec{V}_{\text{star}}$ the velocity of the star and $\vec{k}_{\text{in}}$ the direction of propagation of the photon. For a reference frame outside the simulation box, this photon has a frequency

$$\nu_{\text{ext}} = \left(1 + \frac{\vec{V}_{\text{star}} \cdot \vec{k}_{\text{in}}}{c}\right) \nu_0. \tag{3.23}$$

In the reference frame of a cell with a gas velocity $\vec{V}_{\text{gas}}$, the wavelength becomes

$$\nu_{\text{cell}} = \left(1 - \frac{\vec{V}_{\text{gas}} \cdot \vec{k}_{\text{in}}}{c}\right) \nu_{\text{ext}}. \tag{3.24}$$

In case the photon has already scattered at least once on an ion, it has been reemitted by the gas, hence the ν_{ext} in Equation 3.24 will be computed as in Equation 3.23 but with $\vec{V}_{\text{gas}}$ of the cell where the interaction happened rather than $\vec{V}_{\text{star}}$.

Now, we have to define the Doppler parameter b, which is also important for the optical depth. It depends on the gas temperature T and the turbulent velocity v_{turb}, which in my work is a free parameter that accounts for the subgrid gas movements that are not resolved by the simulation and for smoothing the discontinuities of the velocity field when jumping from one cell to another. More details about the turbulent velocity are in Mauerhofer et al. (2021) (Chapter 4). The Doppler parameter b is then

$$b\,[\text{cm/s}] = \sqrt{\frac{2k_b T}{m_{\text{ion}}} + V_{\text{turb}}^2}, \tag{3.25}$$

where m_{ion} is the mass of the ion interacting with the photon. b is a quantity that sets the width of the cross-section as a function of wavelength. The bigger b is, the more photons with wavelengths away from line center λ_{line} will interact.

Finally, we define new dimensionless variables, that simplify the notations:

$$a = \frac{A_{\text{line}} \lambda_{\text{line}}}{4\pi b},$$

$$x = \frac{c}{b} \frac{\nu_{\text{cell}} - \nu_{\text{line}}}{\nu_{\text{line}}},$$

and the Voigt function:

$$H(a, x) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2}}{(y - x)^2 + a^2} \, \mathrm{d}y. \tag{3.26}$$

The variable x is a normalised wavelength shift from the line center, and a tells us whether it is the intrinsic quantum width of the line or the Doppler width, caused by the movements of the atoms, that dominate the width of the line. If $a \ll 1$, the intrinsic quantum width of the line is negligible.

With those quantities in mind, we can now describe the optical depth of interaction of a photon packet with a spectral line. It is the product of a cross-section and a column density, $\tau_{\text{line}} = \sigma_{\text{line}} N_{\text{ion}}$. The column density is the product of the ion density and the distance the photon travels in a cell, $N_{\text{ion}} = n_{\text{ion}} d$. The cross-section is given by the formula

$$\sigma_{\text{line}}(a, x) = \frac{\sqrt{\pi} e^2 f_{\text{line}}}{m_e c} \frac{\lambda_{\text{line}}}{b} H(a, x). \tag{3.27}$$

The main computational difficulty is to compute the Voigt function, especially since it has to be done each time a photon goes through a cell, so potentially millions or even billions of times. It is an integral of an intricate function with two variables. There have been several implementations of numerically efficient ways to compute this function (e.g. Humlícek, 1982; Tasitsiomi, 2006; Smith et al., 2015). Michel-Dansac et al. (2020) tested those three methods and concluded that they all have advantages and disadvantages depending on the precise values of a and x but that the method from Smith et al. (2015) seems to be consistently faster and accurate, which is why I use this method in my Rascas runs.

Now that we have seen how to compute the optical depth, the last things to understand are how Rascas transfers photon packets in the simulation and what is the outcome of an interaction between a photon and an ion.

We assume that the direction of propagation of a photon after it interacts with an ion is isotropic. What remains to be understood is the output wavelength of a photon after interaction with an ion. First, there can be several channels of de-excitation of an ion after absorbing a photon. As I explained in Section 3.4, many metallic ions have a ground state split into several spin states, typically two, as for C^+ and Si^+. For each such spin state, including the true ground state, there is a decay channel of the ion after interaction with a photon. To determine which one is taken, one has to compare their Einstein coefficients A_{channel}. The probability of decaying through a certain channel is $A_{\text{channel}}/A_{\text{tot}}$, where A_{tot} is the sum of Einstein coefficients of all possible channels. As usual, we use a random number between 0 and 1 to decide which channel is taken.

The decay channel is not enough to determine the exact wavelength of the outgoing photon, one also has to know the velocity of the interacting atom with respect to the surrounding gas. I already mentioned the importance of the Doppler effect due to the velocity of the sources of photons or the velocity of the absorbing gas. Now we have to go to smaller scales, and look at the Doppler effect due to the motion of the atom that absorbs the photon. Its movement has a random nature, but is dictated by the temperature and quantum physics laws. Using our usual Monte-Carlo techniques, we have to determine the velocity of the absorbing atom, both parallel and perpendicular components, $V_{\parallel}$ and $V_{\perp}$ respectively. Let us call $u_{\parallel} = V_{\parallel}/b$ and $u_{\perp} = V_{\perp}/b$.

Then, following Michel-Dansac et al. (2020), the probability function for $u_\parallel$ is:

$$f(u_\parallel) = \frac{a}{\pi H(a, x_{\text{in}})} \frac{e^{-u_\parallel^2}}{a^2 + (x_{\text{in}} - u_\parallel)^2}. \tag{3.28}$$

This function is too complex to integrate or inverse, so it requires numerical tricks and approximations to draw a value of $u_\parallel$ from random numbers. This is an intricate process, and I refer the reader to Smith et al. (2015), Michel-Dansac et al. (2020) or the thesis of Dr. Maxime Trebitsch[7]. For $u_\perp$, the probability function is

$$f(u_\perp) = \frac{e^{-u_\perp^2}}{\sqrt{\pi}}. \tag{3.29}$$

Following the Box–Muller transform, we can sample this function analytically with two random numbers r_1 and r_2 between 0 and 1 via the relation:

$$u_\perp = \sqrt{-\ln r_1} \cos(2\pi r_2). \tag{3.30}$$

Once $V_\parallel$ and $V_\perp$ are known, one can easily compute the outgoing wavelength by taking the reference frame of the interacting atom. If the scattering event is resonant, i.e. no fluorescent decay, quantum physics dictates that it has to be coherent scattering. In other words, the wavelength of the photon is the same before and after scattering, *in the reference frame of the atom*. As a reminder, ν_{ext} is the frequency of the photon as seen from outside the simulation box, ν_{cell} the frequency in the reference frame of the cell where the interaction happens and ν_{atom} in the reference frame of the atom. As seen from the atom, the frequency of the incoming photon is

$$\nu_{\text{atom}} = \nu_{\text{ext}} \left(1 - \frac{(\vec{V}_{\text{gas}} + \vec{V}_{\text{atom}}) \cdot \vec{k}_{\text{in}}}{c}\right) = \nu_{\text{cell}} - \nu_{\text{ext}} \frac{\vec{V}_{\text{atom}} \cdot \vec{k}_{\text{in}}}{c} = \nu_{\text{cell}} - \nu_{\text{ext}} \frac{V_\parallel}{c}.$$

In the resonant case, the outgoing frequency is the same as the incoming one, in the reference frame of the atom, hence the outgoing frequency for an external observer is

$$\nu_{\text{ext}}^{\text{out}} = \nu_{\text{atom}} \left(1 + \frac{(\vec{V}_{\text{gas}} + \vec{V}_{\text{atom}}) \cdot \vec{k}_{\text{out}}}{c}\right). \tag{3.31}$$

Note that $\vec{V}_{\text{atom}} \cdot \vec{k}_{\text{out}} = \mu V_\parallel + \sqrt{1 - \mu^2} V_\perp$, where $\mu = \vec{k}_{\text{in}} \cdot \vec{k}_{\text{out}}$. This gives us $\nu_{\text{ext}}^{\text{out}}$, the quantity that we were looking for, as a function of known quantities ($\vec{V}_{\text{gas}}$, $\vec{V}_{\text{atom}}$, μ, $V_\parallel$ and $V_\perp$).

In the case of fluorescent decay, we have the same equation for the outgoing frequency, except that ν_{atom} is simply the exact frequency of the fluorescent channel. In this case the scattering is not coherent, and the atom loses the information of the incoming photon's frequency.

Finally, we need to address the transfer of ionising photons, which I developed for accurate determination of escape fractions at different wavelengths, as it is presented in Mauerhofer et al. (2021). When transferring Lyman continuum, everything I describe in this section applies as well, except for the computation of the optical depth. I implemented both the approximation of Verner et al. (1996) and the analytical formula from Osterbrock & Ferland (2006):

$$\sigma(E) = \frac{6.3 \times 10^{-18}\,[\text{cm}^2]}{Z^2} \left(\frac{Z^2 E_0}{E}\right)^4 \frac{\exp\left[4 - \frac{4 \arctan(\epsilon)}{\epsilon}\right]}{1 - \exp\left(-\frac{2\pi}{\epsilon}\right)}, \text{ for E} > \text{E}_0, \tag{3.32}$$

where E is the energy of the photon, Z is the atomic number, 1 for hydrogen and 2 for helium, $E_0 = 13.6$ eV is the ionising energy of neutral hydrogen and $\epsilon = \sqrt{E/E_0 - 1}$.

This completes all the computations that Rascas has to make for propagating photons in the simulation.

[7] https://www.theses.fr/2016LYSE1119

3.7.2 Peeling-off algorithm

My main goal during this PhD is to bring observations and simulations closer, by making accurate and realistic mock observations of simulated galaxies. To this end, the radiative transfer code Rascas as I described until here is not fully satisfactory, since it computes the global output properties of a collection of photons, in all directions, and not in one particular direction such as in real observations. Luckily, there is an algorithm to simulate the presence of observers in given directions of observations and faithfully produce mock observations. The cost for this is a larger computation time. The memory use can also be increased significantly, especially when building mock data cubes (images with a spectrum in each pixel) and using many directions of observations. Detailed reports on this method are in e.g. Dijkstra (2017).

The parameters for the peeling-off algorithm are the directions of observation, which are unit vectors, the coordinates of the center of the mock observation, and the kind of observation to build, along with its associated parameters. The kind of observation can be a spectrum, as if acquired by an instrument with a circular aperture. It is parametrised by a wavelength range, a spectral resolution, and a radius for the aperture. The resolution is in principle as high as the user wants, but better resolutions need more photon packets to avoid the numerical, Monte-Carlo, noise. One can also make an image, for which the parameters are the number of pixels and the size of the image. Finally, the user can choose a data cube, which combines a spectrum and an image. In principle more complex setups can be imagined, like a data cube with a better spatial resolution enhanced in the central region, or a spectral resolution enhanced around the wavelength of spectral lines, but it is not yet implemented in Rascas.

The basic idea of the peeling-off method is to compute the contribution of each emission and scattering event to the "telescope" of an observer located toward a direction $\vec{k}_{\rm obs}$. Since a photon packet represents a collection of real photons, each emission (first emission or after a scattering event) does contribute to the light arriving in the telescope of the observer, rather than contributing to one single direction, the one Rascas chooses to give the photon.

At each emission of a photon, be it the initial emission from the source or after any scattering event, the peeling-off algorithm saves the position, initial wavelength and initial direction of propagation (before the scattering actually happens) in what we call a list of peels. When this list reaches a threshold size, or at the end of the run, all those peels are treated and their contributions to the mock observation are computed. First, the code verifies if the position of the peel, as seen from the chosen direction of observation, falls in the aperture of the mock observation. If not, its contribution is not computed. If it does fall on the aperture, the code computes the output frequency of the interaction in an external reference frame, $\nu_{\rm ext}$, just as described in the last section, but forcing the output direction to $\vec{k}_{\rm obs}$ rather than determining it by a Monte-Carlo draw. This frequency then serves to compute the optical depth $\tau_{\nu_{\rm ext}}$ from the position of the peel to the border of the computation domain along the direction $\vec{k}_{\rm obs}$. The contribution of this peel to the mock observation is then

$$w_\mu L_{\rm photon} e^{-\tau_{\nu_{\rm ext}}},$$

where w_μ is a weight that applies in case the directions of propagation after a scattering event are not symmetric, to increase or decrease the contribution if $\vec{k}_{\rm obs}$ is a probable or improbable direction, and $L_{\rm photon}$ is the luminosity of a photon packet (which is the same for each packet). In our case, we assume isotropic scatterings, so that $w_\mu = 1$. Finally, this contribution is added to one of the pixels, either spectral, spatial or 3D, for spectra, images or cubes respectively, depending on the position of the peel in the aperture of observation and on $\nu_{\rm ext}$.

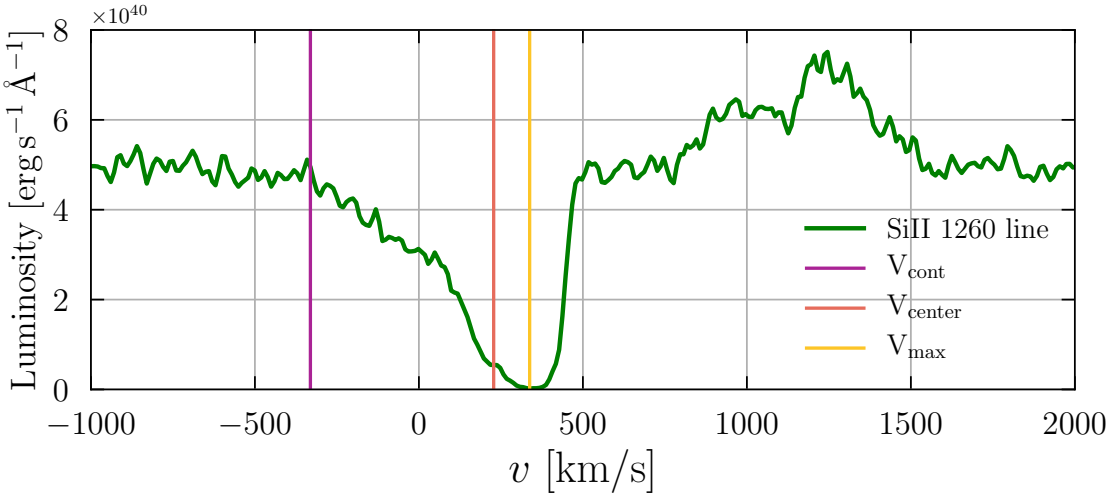

Figure 3.12: Si^+ spectrum of the second most massive halo of the Sphinx20 simulation, as seen from one direction of observation (with the peeling-off algorithm). The x-axis is in velocity space, with $v = c\,(\lambda - \lambda_{\rm line})/\lambda_{\rm line}$.

3.8 Line properties from high resolution spectra

I now define some properties of absorption lines and present how we compute them on our mock observations. Figure 3.12 shows an example of a metallic absorption line, namely a Si^+ line from the second most massive halo of the Sphinx20 simulation. Here are definitions of interesting properties of absorption lines, and values for the one in Figure 3.12:

- The continuum luminosity $L_{\rm cont}$ is obtained by averaging the luminosities of the flat part on both sides of the line. Its value is 4.91×10^{40} [erg s^{-1} Å^{-1}].
- The equivalent width of the absorption lines, $\mathrm{EW_{abs}}$, is the area of the absorption region, below the continuum values, normalised by the continuum luminosity. In other terms, $\mathrm{EW_{abs}}$ [Å] $= \int_{\lambda_{\rm min}}^{\lambda_{\rm max}} \frac{L_{\rm cont} - L_\lambda}{L_{\rm cont}}\,{\rm d}\lambda$, where $\lambda_{\rm min}$ and $\lambda_{\rm max}$ are wavelength chosen to encompass the full absorption feature and L_λ is the luminosity as a function of wavelength. $\mathrm{EW_{abs}}$ can be seen as the width of the rectangle of height equal to $L_{\rm cont}$ and area equal to the absorption line. In Figure 3.12, $\mathrm{EW_{abs}} = 1.85$ Å.
- The quantity $V_{\rm max}$ is the velocity at which the line reaches its minimum luminosity. In this case, it is equal to 337.7 km s^{-1}. When $V_{\rm max}$ is positive we say that the line is redshifted, and it is blueshifted otherwise.
- $V_{\rm center}$ is the velocity at which the area of the absorption line is cut in half. If the line is symmetric, it is equal to $V_{\rm max}$. Our example of an absorption profile has $V_{\rm center} = 228$ km s^{-1}. It is smaller than $V_{\rm max}$ because our line is asymmetric, extended toward the blue side. To compute it, I build the cumulative integral $C_i = \int_{\lambda_{\rm min}}^{\lambda_i} \frac{L_{\rm cont} - L_\lambda}{L_{\rm cont}}\,{\rm d}\lambda$ for each wavelength λ_i in our (discretised) spectrum. Then, I look for the index i for which $C_i = \mathrm{EW_{abs}}/2$, and this is the index of $V_{\rm center}$.
- The so-called "continuum velocity", $V_{\rm cont}$ is the velocity at which the left side of the spectrum reaches the value of the continuum, here equal to -330.8 km s^{-1}. To compute it I loop from the minimum between $V_{\rm max}$ and $V_{\rm center}$, going downward, until the spectrum reaches the value of the continuum. For noisy spectra, I first estimate the average noise in the same spectral range where I compute the continuum luminosity, and loop downwards until reaching $L_{\rm cont}$ − noise.
- All quantities, from equivalent width to the different velocity parameters, can in principle

be computed for the fluorescent emission. In practice I compute only the equivalent width EW_{fluo}, which is 0.89 Å for this spectrum.

- Finally, one of the most important parameters is the residual flux R. It is defined as the ratio between the luminosity at the deepest of the absorption and the continuum luminosity, $R = L_{min}/L_{cont}$. It is then also equal to the luminosity at V_{max}, in the normalised spectrum. A small residual flux means that the line is deep or even saturated, while a number close to one indicates that the line is weak. In Figure 3.12, the residual flux is 0.004, the line is essentially saturated.

As an illustration of other kind of mock observations that can be done with Rascas, I show in Figure 3.13 mock images of the zoom simulation presented in Chapter 2, at redshift 3.2. The mocks were done by launching a continuum between wavelengths 920Å and 2750Å from the stellar particles with the *Table* method. The peeling-off algorithm was set to build a mock data cube from a total of 108 directions of observation. The resulting data cube is then treated to get images as if this galaxy would be at redshift 0.3 and observed with the Hubble COS imager. The details of the treatment of the data cube are in the caption of the figure. We see that the galaxy is dominated by one bright region that can be seen from several angles. However, there are directions of observation from which the bright region is hidden by dust, resulting in a more homogeneous image, with low signal-to-noise ratio.

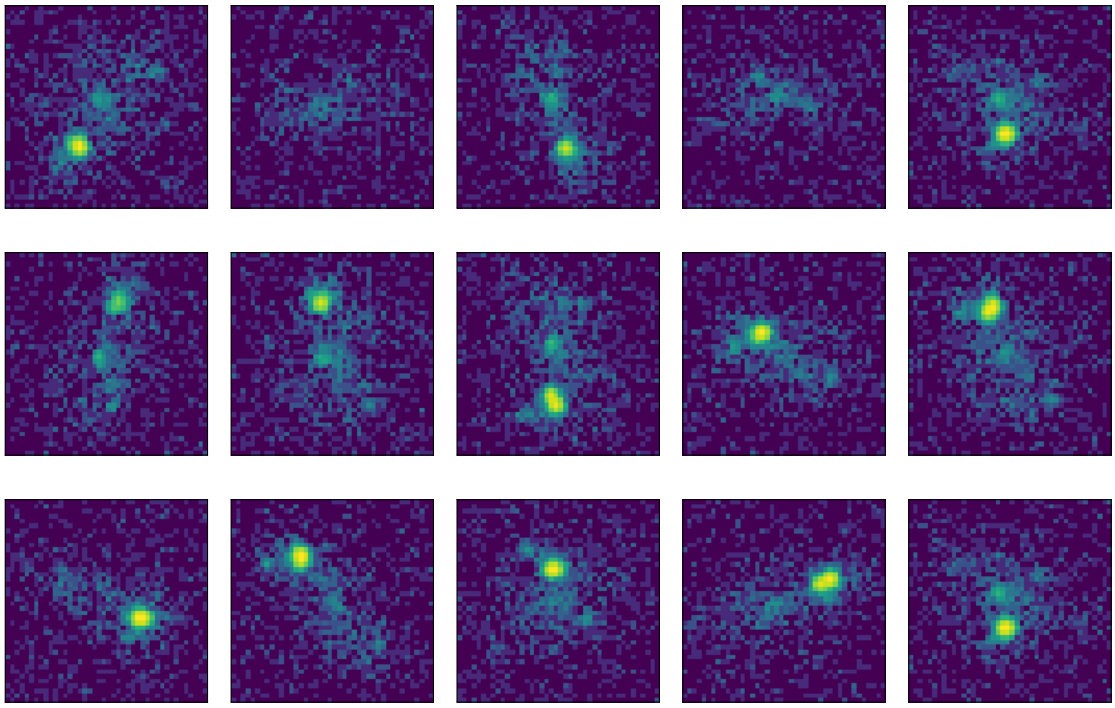

Figure 3.13: Images of the galaxy from the zoom simulation at redshift 3.2, seen from different angles. The mock images are done as if the galaxy was at redshift 0.3, observed with the Hubble COS imager with an exposure time of 300 seconds. The units are counts/s, in a logarithmic color scale. The NUV spectrum in each pixel is convolved with the transmission function of the instrument[8]. The area of the mirror is 45'250 cm^2. The pixel size is 23.5 mas (105 pc). The PSF, assumed to be Gaussian, has a width of 1.97 pixels. The sky background is 3.5×10^{-4} counts/s and the dark current 7×10^{-4} counts/s. The actual number counts in each pixel is a realisation of the Poisson distribution with a mean equal to the expected number of photons in this pixel.

As a comparison, I show in Figure 3.14 the same images as in Figure 3.13 but with a higher

[8] https://hst-docs.stsci.edu/cosihb/chapter-6-imaging-with-cos/6-2-sensitivity

resolution and without the observational degradations such as the Point Spread Function (PSF) or the limited throughput. As a result, the bright regions are less obvious because the absence of PSF makes them much smaller. One can also see that they are made of several small bright points. Additionally, dust lanes become visible as dark patches with complex shapes.

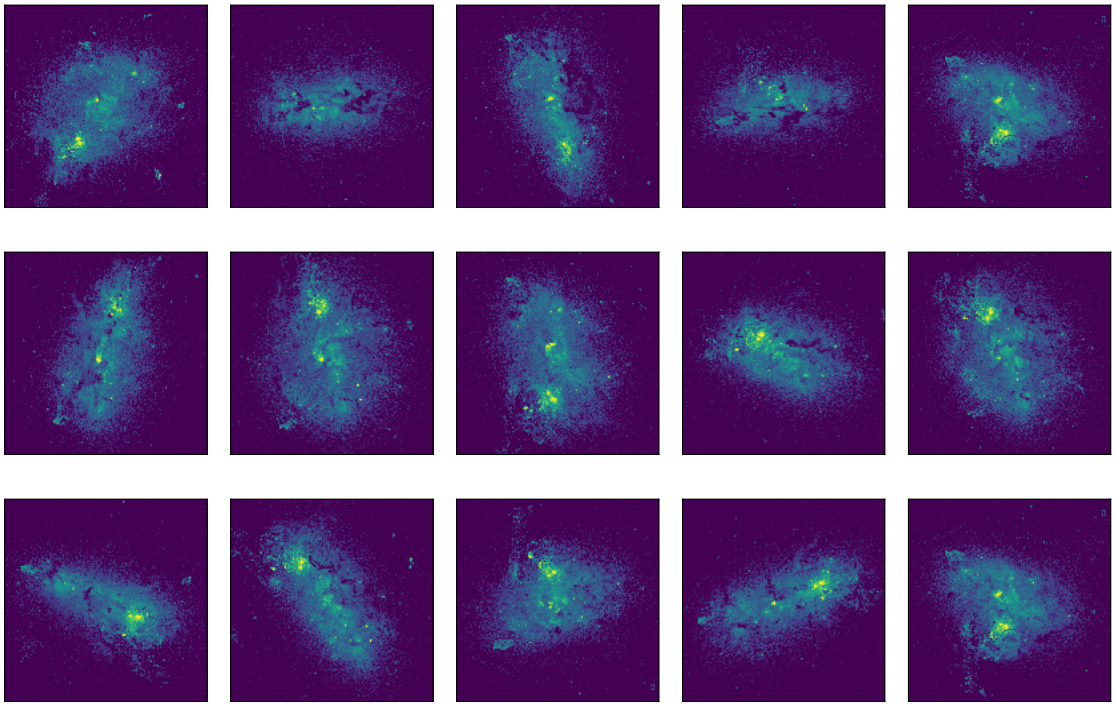

Figure 3.14: Same galaxy and orientations as in Figure 3.13, but with no PSF, a 100% throughput, no background noise and 25 times more pixels.

4 UV absorption lines as a tracer of escape fractions of ionising photons

As discussed in the introduction, the main goal of the post-processing pipeline presented in Chapter 3 is to produce accurate mock down-the-barrel absorption lines to test whether we can use them to predict the escape fraction of ionising photons. In this section, I include my first published paper discussing this topic. In this paper, I use the galaxy from the zoom simulation at three different times, redshift 3.2, 3.1 and 3.0, separated by 80 Myr. I make mock observations of absorption lines along with computations of the escape fraction of ionising photons, both in 1728 directions of observation for each snapshot.

After the abstract and introduction of the paper, I give details about the methods, similarly as in Chapter 3. Scientific results start with Section 3, which is a study of the properties of absorption lines. I discuss how realistic my mock spectra are, test the effect of changing free parameters in the post-processing pipeline and show how different physical processes shape the lines. In Section 4, I plot the escape fraction of ionising photons from different angles, and compare the fraction at 900 Å and the global fraction. The comparison between line properties and escape fractions is done in Section 5. There, I focus on the residual flux of the lines, defined as the ratio between the flux at the bottom of the absorption and the flux of the continuum. I show that a dust correction of the continuum leads to a better correspondence between the residual flux of metallic lines and the escape fractions. In the same section, I explain my results in detail, including tests of whether the picket-fence model applies well in a realistic environment such as our simulations. Finally, I summarise the results in Section 6.

4.1 The paper

A&A 646, A80 (2021)
https://doi.org/10.1051/0004-6361/202039449

Astronomy & Astrophysics

UV absorption lines and their potential for tracing the Lyman continuum escape fraction

V. Mauerhofer[1,2], A. Verhamme[1,2], J. Blaizot[1], T. Garel[1,2], T. Kimm[3], L. Michel-Dansac[1], and J. Rosdahl[1]

[1] Univ. Lyon, Univ. Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, 69230 Saint-Genis-Laval, France
[2] Observatoire de Genève, Université de Genève, Chemin Pegasi 51, 1290 Versoix, Switzerland
e-mail: valentin.mauerhofer@unige.ch
[3] Department of Astronomy, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

Received 16 September 2020 / Accepted 3 December 2020

ABSTRACT

Context. The neutral intergalactic medium above redshift ∼6 is opaque to ionizing radiation, and therefore indirect measurements of the escape fraction of ionizing photons are required from galaxies of this epoch. Low-ionization-state absorption lines are a common feature in the rest-frame ultraviolet (UV) spectrum of galaxies, showing a broad diversity of strengths and shapes. As these spectral features indicate the presence of neutral gas in front of UV-luminous stars, they have been proposed to carry information on the escape of ionizing radiation from galaxies.
Aims. We aim to decipher the processes that are responsible for the shape of the absorption lines in order to better understand their origin. We also aim to explore whether the absorption lines can be used to predict the escape fraction of ionizing photons.
Methods. Using a radiation-hydrodynamical cosmological zoom-in simulation and the radiative transfer postprocessing code RASCAS we generated mock C II λ1334 and Lyβ lines of a virtual galaxy at $z = 3$ with $M_{1500} = -18.5$ as seen from many directions of observation. We also computed the escape fraction of ionizing photons in those directions and looked for correlations between the escape fraction and properties of the absorption lines, in particular their residual flux.
Results. We find that the resulting mock absorption lines are comparable to observations and that the lines and the escape fractions vary strongly depending on the direction of observation. The effect of infilling due to the scattering of the photons and the use of different apertures of observation both result in either strong or very mild changes of the absorption profile. Gas velocity and dust always affect the absorption profile significantly. We find no strong correlations between observable Lyβ or C II λ1334 properties and the escape fraction. After correcting the continuum for attenuation by dust to recover the intrinsic continuum, the residual flux of the C II λ1334 line correlates well with the escape fraction for directions with a dust-corrected residual flux larger than 30%. For other directions, the relations have a strong dispersion, and the residual flux overestimates the escape fraction for most cases. Concerning Lyβ, the residual flux after dust correction does not correlate with the escape fraction but can be used as a lower limit.

Key words. radiative transfer – line: formation – dark ages, reionization, first stars – ultraviolet: galaxies – galaxies: ISM – scattering

1. Introduction

Decades of observations show that the intergalactic medium (IGM) was completely ionized by redshift $z \sim 6$, about one billion years after the Big Bang (e.g., Fan et al. 2006; Schroeder et al. 2013; Sobacchi & Mesinger 2015; Bañados et al. 2018; Inoue et al. 2018; Ouchi et al. 2018; Bosman et al. 2018; Kulkarni et al. 2019). This implies that, before this redshift, sources of ionizing radiation must have emitted enough photons that managed to escape their galactic environment and reach the IGM to photoionize it. The main candidates for the source of reionization are active galactic nuclei (AGNs) and massive stars. Although recent work suggests AGNs could be an important driver of reionization (Madau & Haardt 2015; Giallongo et al. 2019), a consensus seems to be emerging on the idea that their contribution is negligible compared to that of stars (Grissom et al. 2014; Parsa et al. 2018; Trebitsch et al. 2020).

Models based on observations show that stars alone could drive reionization under reasonable assumptions on the Lyman continuum (LyC) production efficiency and on the escape fraction of ionizing photons ($f_{\rm esc}$), for example in Atek et al. (2015), Gnedin (2016) and Livermore et al. (2017). The escape fraction is the least constrained quantity, and is often fixed to a value of between 10 and 20% (e.g., Robertson et al. 2015). Finkelstein et al. (2019) argue that reionization can also be explained when assuming lower escape fractions. Additionally, simulations have successfully reionized the Universe with stellar radiation, either at very large scales (e.g., Iliev et al. 2014; Ocvirk et al. 2016), or at smaller scales with high enough resolution to resolve the escape of ionizing radiation from galaxies (e.g., Kimm et al. 2017; Rosdahl et al. 2018).

In order to better understand the process of reionization and to be able to compare observations and simulations, it is essential that we obtain accurate constraints on the LyC escape fractions of observed galaxies. Recently, several low-redshift galaxies were found to be leaking ionizing photons (e.g., Bergvall et al. 2006; Leitet et al. 2013; Leitherer et al. 2016; Puschnig et al. 2017). When this ionizing radiation flux is detected, it is possible to use a direct method to infer the escape fraction (e.g., Borthakur et al. 2014; Izotov et al. 2016a,b). The total LyC production is computed as the addition of the observed LyC photons and the ones absorbed in the galaxy, deduced from the flux of a (dust corrected) Balmer line, assuming that this line is produced

by recombinations of photoionized hydrogen ions. The ratio of the observed ionizing flux to the assessed intrinsic LyC production gives the escape fraction.

However, the ionizing radiation escaping galaxies during the epoch of reionization is not directly observable because of IGM absorption, and therefore this method cannot be used. To circumvent this limitation, different ways to infer the escape fractions indirectly have been proposed in the literature. The connections between Lyman α (Lyα) and Lyman continuum emission have been studied (Verhamme et al. 2015; Dijkstra et al. 2016; Kimm et al. 2019) with promising results at $z \sim 3$ (Steidel et al. 2018) or in the local Universe (Verhamme et al. 2017; Izotov et al. 2018). Above $z \sim 6$, the intergalactic neutral hydrogen may alter Lyα too much, except for galaxies that have such big ionized bubbles that Lyα is redshifted enough before reaching the neutral IGM (Dijkstra 2014; Mason et al. 2018; Gronke et al. 2020). A possible alternative is to use Mg II $\lambda\lambda$2796, 2803 (Henry et al. 2018), a resonant emission line doublet which is a candidate proxy of Lyα and has the advantage of not being absorbed by the neutral IGM of the epoch of reionization. A second method is to use O32, which is defined as the line ratio [O III λ5007]/[O II $\lambda\lambda$3726,3729], and is thought to correlate with the escape fraction (Jaskot & Oey 2013; Nakajima & Ouchi 2014; Izotov et al. 2016b). However, recent studies shed doubt on this method (Izotov et al. 2018; Bassett et al. 2019; Katz et al. 2020). A high O32 ratio seems to be a good indicator of a nonzero escape fraction, but the correlation between the two quantities is weak.

A third way to infer the escape fractions indirectly is to use down-the-barrel absorption lines from low-ionization states (LISs) of metals, such as C II, O I or Si II lines. The shapes of these lines are used as an indicator of the covering fraction of the absorber. Indeed, column densities of metals in galaxies typically lead to saturated absorption lines (i.e., absorption profiles where the minimum flux reaches zero). When nonsaturated lines are observed, the residual flux is interpreted as being due to partial coverage of the sources by the absorbing material, meaning that the ratio of this residual flux to the continuum is used to deduce an escape fraction. As the LISs of metals are thought to be good tracers of neutral hydrogen, this escape fraction is then linked to the escape fraction of ionizing photons (e.g., Erb 2015; Steidel et al. 2018; Gazagnes et al. 2018; Chisholm et al. 2018)

There have not yet been observations of absorption lines at the epoch of reionization because it is challenging to detect the stellar continuum with a sufficiently high signal-to-noise ratio. However, there has been impactful progress in the study of absorption lines for two decades thanks to the observation of galaxies at ever higher redshifts and of local analogs of high-redshift galaxies. Detections of down-the-barrel absorption lines of individual galaxies with the highest redshifts, with $z \sim 3$–4, are carried out using gravitational lenses (e.g., Smail et al. 2007; Dessauges-Zavadsky et al. 2010; Jones et al. 2013; Patrício et al. 2016). Alternatively, without gravitational lensing, it is possible to study absorption lines at these redshifts by stacking spectra of many galaxies (e.g., Shapley et al. 2003; Steidel et al. 2018; Feltre et al. 2020). The most common lines observed at these high redshifts are Si II (λ1190, λ1193, λ1260, λ1304), O I λ1302 or C II λ1334, which are strong lines whose observed wavelengths fall in the optical range. There are also observations of slightly redder absorption lines, such as Fe II (λ2344, λ2374, λ2382, λ2586, λ2600) or Mg II $\lambda\lambda$2796, 2803, at intermediate redshifts between 0.7 and 2.3 (e.g., Finley et al. 2017; Feltre et al. 2018). Finally, there are many studies of absorption lines at $z < 0.3$ (e.g., Rivera-Thorsen et al. 2015; Chisholm et al. 2017, 2018; Jaskot et al. 2019).

Because most atoms and ions in the interstellar medium (ISM) are in their ground state, absorption lines are typically transitions from the ground state to a higher level. The energy of these transitions translates into lines which are mostly in the rest-frame ultraviolet (UV). Some lines are resonant (e.g., Lyα, Mg II $\lambda\lambda$2796, 2803, Al II λ1670), but most are not completely resonant because the ground state is split into different spin levels, and de-excitation may produce photons with different wavelengths. If the wavelength of the so-called fluorescent channel(s) is different enough from the resonant wavelength the emitted photon will leave the resonance.

Observed absorption lines show a large diversity of spectral profiles, for example in terms of depth, width, and complexity. A common feature is that absorption is often blueward of the systemic velocity, with a minimum of the absorption occurring between around -400 and $100\,\mathrm{km\,s^{-1}}$. There are also detections of fluorescent emission redward of the absorption (e.g., Shapley et al. 2003; Rivera-Thorsen et al. 2015; Finley et al. 2017; Steidel et al. 2018; Jaskot et al. 2019).

There are several physical processes that challenge the use of absorption lines to measure the covering fraction of neutral gas and the escape fraction of ionizing photons. One of them is the distribution of the velocity of the gas at galaxy scales. The gas in front of the sources moves at different velocities, leading to an absorption that is spread in wavelength. This can lead to an increase of the flux at the line center compared to a case where all the gas has the same velocity (Rivera-Thorsen et al. 2015). Another process is scattering: when photons are absorbed by an ion they are not destroyed but scattered. Thus, the flux in absorption lines can be increased by photons that were initially going away from the observer but that end up in the line of sight because of resonant scattering (Prochaska et al. 2011; Scarlata & Panagia 2015).

This paper investigates the relation between the escape fraction of ionizing photons and LIS absorption lines using mock spectra constructed from a radiation-hydrodynamic (RHD) simulation. There have been studies of various emission lines by post-processing simulations (e.g., Barrow et al. 2017, 2018; Katz et al. 2019, 2020; Pallottini et al. 2019; Corlies et al. 2020), and there are also studies on absorption lines in simulations, but these latter focus on quasar sightline absorption instead of down-the-barrel absorption (e.g., Hummels et al. 2017; Peeples et al. 2019). Concerning down-the-barrel lines, studies have been done using Monte-Carlo methods in idealized geometries (Prochaska et al. 2011) or using semi-analytic computations (Scarlata & Panagia 2015; Carr et al. 2018). Finally, there have been studies of mock down-the-barrel absorption lines in simulations (Kimm et al. 2011), but this present work is to our knowledge the first time that such lines are produced in a high-resolution RHD simulation and including resonant scattering. We focus in particular on the C II λ1334 line, which is among the strongest LIS absorption redward of Lyα, is not contaminated by absorption lines of other abundant elements (unlike O I lines), and is frequently used in the literature (e.g., Heckman et al. 2011; Rivera-Thorsen et al. 2015; Steidel et al. 2018). We also show results using the Si II λ1260 line. In addition, we study Lyβ to see if it is in general a better tracer of the escape fraction, even though it is not observable at the epoch of reionization because its wavelength is 1026 Å.

The paper is structured as follows: in Sect. 2 we provides details of our method, from the simulation that we use to the post-processing that is necessary to make mock observations

Table 1. Properties of the simulated galaxy at three different times.

Output	z	$R_{\rm vir}$ [kpc]	$M_{\rm h}$ [$M_\odot$]	$M_{\rm gas}$ [$M_\odot$]	M_* [$M_\odot$]	$Z_{\rm gas}$ [$Z_\odot$]	Z_* [$Z_\odot$]	SFR [$M_\odot\,{\rm yr}^{-1}$]	LyC $f_{\rm esc}$ [%]	M_{1500} [mag]	A_{1500} [mag]
A	3.2	27	5.09×10^{10}	5.15×10^{9}	1.88×10^{9}	0.40	0.36	5.0	$11.0^{+23.9}_{-10.5}$	$-19.0^{+1.0}_{-0.5}$	$1.3^{+1.1}_{-0.5}$
B	3.1	28	5.30×10^{10}	5.13×10^{9}	2.08×10^{9}	0.43	0.40	2.2	$3.8^{+8.2}_{-3.7}$	$-18.2^{+0.6}_{-0.4}$	$1.6^{+0.6}_{-0.4}$
C	3.0	29	5.56×10^{10}	4.83×10^{9}	2.27×10^{9}	0.43	0.42	4.2	$1.0^{+2.2}_{-1.0}$	$-18.5^{+0.5}_{-0.4}$	$1.5^{+0.5}_{-0.4}$

Notes. All the masses are computed inside the virial radius and the metallicities in a sphere of a tenth of the virial radius. The SFR is averaged over the last 10 Myr. LyC $f_{\rm esc}$ is the escape fraction of hydrogen ionizing photons computed at the virial radius (Sect. 2.6). M_{1500} is the absolute magnitude of the galaxy at 1500 Å, including dust attenuation. A_{1500} represents the attenuation by dust, following the equation: $F_{1500}^{\rm observed} = 10^{-0.4\,A_{1500}} \times F_{1500}^{\rm intrinsic}$. The escape fractions, magnitudes, and attenuation by dust, which are the only three quantities depending on the direction of observation, were computed for 1728 isotropically distributed directions. We show the mean of the results and the 10th and 90th percentiles.

and to compute escape fractions of ionizing photons. In Sect. 3 we present the resulting spectra and we assess their robustness against changes in modeling parameters. We also study the effect that various physical processes have on the spectra. In Sect. 4 we show the values of the escape fractions, and discuss the effects of helium and dust. In Sect. 5 we compare the line properties with the escape fractions to try to find correlations. We also explain the different sources of complexity of the relations we find. Finally, we summarize our results in Sect. 6.

2. Methods

In this section we provide details of the simulation that we use and the steps that are necessary to build the mock observations. In this paper we focus on the C II λ1334 absorption line but the method can be applied to any absorption line. In Sect. 5.5 we show the main results of the paper for Si II λ1260. Here we use one galaxy from a zoom-in simulation, and in future work we will study a statistical sample of galaxies from various simulations like SPHINX[1] (Rosdahl et al. 2018) or Obelisk (Trebitsch et al. 2020).

2.1. Simulation

We use a zoom-in simulation run with the adaptive mesh refinement (AMR) code RAMSES-RT (Teyssier 2002; Rosdahl et al. 2013; Rosdahl & Teyssier 2015), including a multi-group radiative transfer algorithm using a first-order moment method which employs the M1 closure for the Eddington tensor (Levermore 1984). The collisionless dark matter and stellar particles are evolved using a particle-mesh solver with cloud-in-cell interpolation (Guillet & Teyssier 2011). We compute the gas evolution by solving the Euler equations with a second-order Godunov scheme using the HLLC Riemann solver (Toro et al. 1994). The initial conditions are generated by MUSIC (Hahn & Abel 2013) and chosen such that the resulting galaxy at $z = 3$ has a stellar mass of around $10^9\,M_\odot$. The physics of cooling, supernova feedback, and star-formation are the same as in the SPHINX simulations (Rosdahl et al. 2018). Radiative transfer allows for a self-consistent and on-the-fly propagation of ionizing photons in the simulation, which provides an accurate nonequilibrium ionization state of hydrogen and helium, as well as radiative feedback. We use three energy bins for the radiation: The first bin contains photons with energies between 13.6 eV and 24.59 eV, which ionize H^0, the second bin between 24.59 eV and 54.42 eV, which ionize H^0 and He^0, and the last above 54.42 eV, which ionize H^0, He^0 and He^+. The ionizing radiation is emitted from the stellar particles following version 2.0 of the BPASS[2] stellar library spectral energy distributions (SEDs; Eldridge et al. 2008; Stanway et al. 2016). In this version of BPASS the maximum mass of a star is $100\,M_\odot$, and all stars are considered to be in binaries. To account for the ionizing radiation produced by external galaxies that are not present in the zoom-in simulation, we add an ionizing UV background (UVB) to all cells of the simulation with $n_{\rm H} < 10^{-2}\,{\rm cm}^{-3}$, following Faucher-Giguère et al. (2009). The outputs of the simulation have a time resolution of 10 Myr, going down to redshift $z = 3$. The maximum cell resolution around $z = 3$ is 14 pc. The halo mass, stellar mass, stellar metallicity, and gas metallicity of the galaxy at $z = 3$ are $M_{\rm h} = 5.6\times10^{10}\,M_\odot$, $M_* = 2.3\times10^9\,M_\odot$, $Z_\star = 0.42\,Z_\odot$, and $Z_{\rm gas} = 0.43\,Z_\odot$, respectively.

In this paper we focus on three outputs of the simulation that span a range of ionizing photon escape fractions; we call these outputs A, B, and C. There are roughly 80 Myr of separation between each output. Output C is the last of the simulation, at $z = 3$. Properties of the outputs, including the star formation rate (SFR), are shown in Table 1. In Fig. 1 we plot the neutral hydrogen column density and the stellar surface brightness at 1500 Å at the virial radius scale and the ISM scale (a tenth of the virial radius), for output C. The column density map in the upper left panel shows gas that is being accreted from the IGM, and regions where the supernova feedback disrupts the gas flows. The upper right panel shows that there are many small satellites orbiting the galaxy. The lower panels highlight the high resolution of the simulation, where small star-forming clusters and dust lanes are well resolved.

2.2. Computing densities of metallic ions

To be able to make mock observations of metallic absorption lines in the simulation, we have to compute the density of the ions of interest in post-processing. The cosmological zoom-in simulation does not trace densities of elements other than hydrogen and helium, but it follows the value of the metal mass fraction Z in every cell. Assuming solar abundance ratios, we obtain the carbon density via the equation $n_{\rm C} = n_{\rm H} A_{\rm C}^\odot\, Z/Z_\odot$, where $n_{\rm H}$ is the hydrogen density, $A_{\rm C}^\odot = 2.69\times10^{-4}$ is the solar ratio of carbon atoms over hydrogen atoms, and $Z_\odot = 0.0134$ is the solar metallicity (Grevesse et al. 2010). We do not consider the depletion of carbon into dust grains, which could alter the densities in the gas phase. However, Jenkins (2009) finds that, in the Milky

[1] https://sphinx.univ-lyon1.fr

[2] https://bpass.auckland.ac.nz/index.html

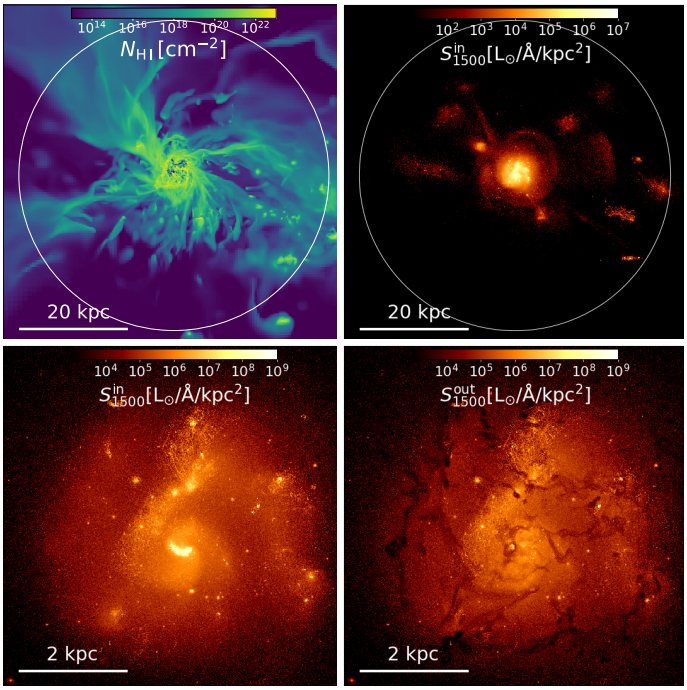

Fig. 1. Maps of neutral hydrogen column density and surface brightness at 1500 Å for a snapshot of the simulated galaxy at $z = 3$. *Upper left*: column density of neutral hydrogen atoms, with the virial radius highlighted by the white circle. *Upper right*: intrinsic surface brightness at 1500 Å. *Lower left*: zoom-in to a tenth of the virial radius. *Lower right*: surface brightness at 1500 Å after extinction by dust.

Way (MW), not more than 20−50% of the carbon is depleted into dust. Moreover, De Cia (2018) shows that the depletion of metals into dust is less important in high-redshift galaxies than in the MW, which is why we can assume that the depletion does not change the gas-phase carbon density significantly. In order to assess the impact of the uncertainties of abundance ratios and dust depletion on our conclusions, we test the effect of dividing and multiplying the density of C^+ by two in Sect. 3.2. The effect is visible but mild.

What remains is to compute the ionization fractions of carbon to get the densities of C^+. To do so we use KROME[3] (Grassi et al. 2014), which is a code that computes the evolution of any chemical network implemented by the user. As this computation is done in post-processing, we need to assume equilibrium values for the ionization fractions of carbon; we therefore use a routine of KROME that evolves the chemical network until it reaches equilibrium. For hydrogen and helium we use the (nonequilibrium) ionization fractions from the simulation.

The ionization fractions are determined by three main processes: recombination, collisional ionization, and photoionization. There can also be charge-transfer reactions between carbon ions and hydrogen or helium ions, but we omit them in this work, as we explain below. One of the major advantages of KROME is that the chemical network is completely customizable. The user chooses all the reactions and their rates as a function of temperature. We use the following rates:

– Recombination rates: Badnell (2006).
– Collisional ionization rates: Voronov (1997).
– Photoionization cross-sections: Verner et al. (1996).

We choose recombination and collisional ionization rates to reproduce the carbon ionization fractions as a function of temperature for a collisional ionization equilibrium setup in Cloudy[4] (Ferland et al. 2017), which is the state-of-the-art tool to compute ionization fractions (among other things), but is too slow to use on a full simulation. There are works that use Cloudy on Ramses simulations, but these adopt strategies to avoid running it on every cell. Katz et al. (2019) use Cloudy on a subset of cells and then extrapolate on the other cells using machine learning algorithms. Pallottini et al. (2019) create Cloudy grids that are then interpolated for each cell of the simulation. We opt to use KROME because it is fast enough to directly compute the ionization fractions in every cell of the simulation, although it does not compute line emissivities of the gas. The collisional ionization rates that we use are slightly different from the ones in Cloudy (Dere 2007). We show in Appendix A that we recover the same ionization fractions of carbon as a function of temperature as Cloudy, despite the fact that we lack charge-transfer reactions and have different collisional ionization rates.

Concerning photoionization, the cross-sections that we use are the same as in Cloudy. To obtain a photoionization rate, those cross-sections have to be multiplied by a flux of photons. Taking advantage of the radiation-hydrodynamics in the simulation, we directly use the inhomogeneous radiation field self-consistently computed by the simulation at energies higher than 13.6 eV. Below 13.6 eV, the simulation does not track the radiation field or its interaction with gas. We nevertheless need to account for this lower energy range as it may ionize metals (e.g., the ionizing potential of C is 11.26 eV). For this, we add a simple model for radiation in the Habing band, from 6 eV to 13.6 eV (Habing 1968), in post-processing, using the UVB as inferred by Haardt & Madau (2012). In Sect. 3.2.2 we look at the effect of dividing the UVB by two or multiplying it by 100, and we see that this does not affect the results of the mock observations. We also use the Haardt & Madau (2012) UVB at all wavelengths instead of using the radiation field of the simulation, for comparison. In this case there can be drastic changes, which demonstrates that it is important to use a simulation that treats radiative transfer on-the-fly. In Appendix B we show in more detail how we compute the photoionization rates using the Verner et al. (1996) photoionization cross-sections and the number density of photons in the simulation.

Another advantage of KROME is that it is possible to fix the state of hydrogen and helium. If we compute the ionization fractions with Cloudy, the convergence to equilibrium of hydrogen and helium ionization fractions modifies the electron density compared to the value of the simulation. This different electron density would lead to different carbon ionization fractions than with the simulation electron density. To summarize, we compute the ionization fractions of carbon at equilibrium, but with the hydrogen and helium ionization fractions fixed at the simulation values. Therefore we set all rates of reactions containing hydrogen and helium to zero in KROME. This is why we do not include charge transfer reactions in KROME, as they would also change the simulation values of hydrogen and helium ionization fractions.

2.3. Fine structure of the ground state

Almost all ions have a ground state that is split into several spin states. The C^+ ion has one such spin level, which we call level 2, just above the ground state (level 1), as illustrated in Fig. 2. A fraction of C^+ ions are populating level 2 due to collisions with electrons, and they are absorbing photons at 1335.66 Å or

[3] http://kromepackage.org

[4] https://nublado.org

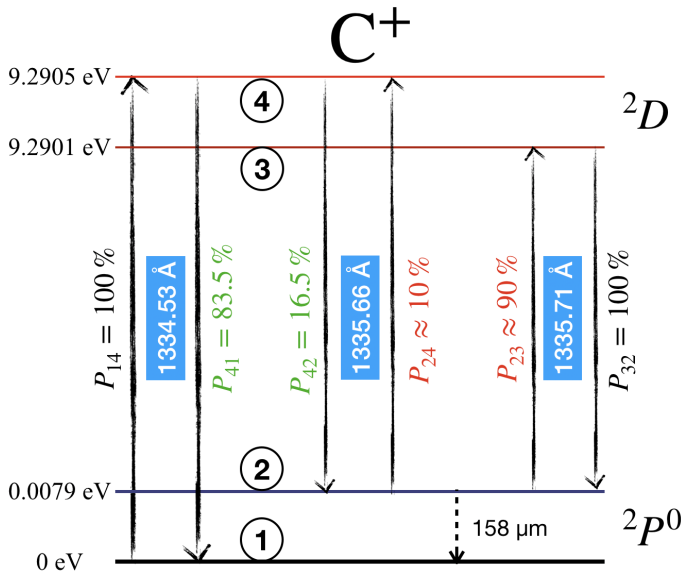

Fig. 2. Energy levels of the C^+ ion. $P_{14} = 100\%$ indicates that a C^+ ion in level 1 can be photo-excited only to level 4. The green numbers are the probabilities that a C^+ ion in level 4 radiatively de-excites to level 1 or 2. The red numbers are the probabilities that an incoming photon with a wavelength between 1335.66 Å and 1335.71 Å hitting a C^+ ion in level 2 excites it to level 3 or 4. $P_{32} = 100\%$ indicates that a C^+ ion in level 3 can only radiatively de-excite to level 2.

1335.71 Å instead of 1334.53 Å from level 1. We take this effect into account by computing the percentage of ions that are in level 2 using PyNeb[5] (Luridiana et al. 2015). For every cell of the simulation we give PyNeb the temperature (assumed to be both the gas temperature and the electron temperature) and the electron density taken from the results of KROME. The fraction of C^+ ions populating level 3 and 4 is extremely small, due to the short lifetime of these levels, and so we consider that the ions only populate levels 1 and 2.

The three lines at 1334.53, 1335.66, and 1335.71 Å are considered in our radiative transfer post-processing, as explained in Sect. 2.4. The effects of these different channels on the C II λ1334 line are studied in Sect. 3.3.

2.4. Radiative transfer of the lines

The radiative transfer post-processing is computed with the code RAdiation SCattering in Astrophysical Simulations (RASCAS[6]; Michel-Dansac et al. 2020). The stellar luminosity of the simulated galaxy is sampled with one million so-called photon packets, each having a given wavelength and carrying one millionth of the total luminosity in the wavelength range considered (a few angströms on each side of the line). We launch the photon packets from the stellar particles of the simulation with the following initial conditions:

– Direction: The initial direction of every photon packet is randomly drawn from an isotropic distribution.

– Position: Each photon packet is randomly assigned to a stellar particle, with a weight proportional to the stellar luminosity in the wavelength range considered. The initial position of the packet is the position of the stellar particle.

– Frequency: The SED of the stellar particle, given by BPASS using the age and metallicity of the particle, is fitted by a power-law approximation, removing any stellar absorption lines. The frequency is then drawn randomly from this power law.

Photon packets are then propagated through the AMR grid of the simulation. The total optical depth in each cell is the sum of four terms, corresponding to four channels of interaction:

$$\tau_{\rm cell} = \tau_{\rm C\,II\,\lambda1334} + \tau_{\rm C\,II^*\,\lambda1335.66} + \tau_{\rm C\,II^*\,\lambda1335.71} + \tau_{\rm dust}. \quad (1)$$

The optical depth of a line is the product of the cross-section and the column density, $\tau_{\rm line} = \sigma_{\rm line} N_{\rm C\,II}$. The column density is the product of the ion density and the distance to the border of the cell[7], $N_{\rm C\,II} = n_{\rm C\,II} d$. The cross-section for a given line of the ion is given by the formula

$$\sigma_{\rm line} = \frac{\sqrt{\pi} e^2 f_{\rm line}}{m_e c} \frac{\lambda_{\rm line}}{b} {\rm Voigt}(x, a), \quad (2)$$

where $f_{\rm line}$ is the oscillator strength of the line, $\lambda_{\rm line}$ is the line wavelength, and b is the Doppler parameter, which is defined below (Eq. (3)). The variable a is defined by $a = A_{\rm line}\lambda_{\rm line}/(4\pi b)$, where $A_{\rm line}$ is the Einstein coefficient of the line. The variable x is defined by $x = (c/b)(\lambda_{\rm line} - \lambda_{\rm cell})/\lambda_{\rm cell}$, where $\lambda_{\rm cell}$ is the wavelength of the photon packet in the frame of reference of the cell. Finally, the Voigt function is computed with the approximation of Smith et al. (2015).

When an interaction occurs, one of the four channels is randomly chosen with a probability $\tau_{\rm channel}/\tau_{\rm cell}$. If the photon packet interacts with dust, it can either be scattered or absorbed (see Sect. 2.5). If the photon packet is absorbed by C^+ in any channel, it will be re-emitted in a direction drawn from an isotropic distribution. If the absorption channel is 1334.53 Å, the photon packet will be re-emitted either via the same channel, which is called "resonant scattering", or via the 1335.66 Å channel, which is the fluorescent channel. The probabilities of the two channels are determined by their Einstein coefficients, and are shown on Fig. 2. Similarly, if the photon packet is absorbed by the channel 1335.66 Å, it can be re-emitted at 1334.53 Å or at 1335.66 Å. Finally, if the photon packet is absorbed by the channel 1335.71 Å, it will always be re-emitted at 1335.71 Å. The transfer ends when all the photon packets either escape the virial radius or are destroyed by dust. More details, for example on the position of interaction or on the frequency redistribution after scattering, are given in Michel-Dansac et al. (2020).

Mock observations are made from different directions of observation thanks to the peeling-off method (e.g., Dijkstra 2017). This method allows us to make mock images, spectra, or data cubes. We can choose a spatial and spectral resolution and an aperture radius to make a mock observation of the whole galaxy or only of a part of it.

We also make mock Lyβ absorption lines. For this we use the density of neutral hydrogen directly from the simulation. No fine-structure level is taken into account because the resulting differences on the Lyβ wavelength would be of only 0.5 mÅ, which is too small for the photon to leave resonance. Lyβ is also a "semi-resonant" line: photons at the Lyβ wavelength can scatter on H^0, leading to potential infilling effects, and, like in the case of C II λ1334, there is a way to leave the resonance, namely via an emission of an Hα photon. The probability for Lyβ to be resonantly scattered is 88.2%.

[5] https://pypi.org/project/PyNeb/
[6] http://RASCAS.univ-lyon1.fr

[7] We use the density of C^+ in level 1 for the line C II λ1334, and the density of C^+ in level 2 for C II λ1335.66 and C II λ1335.71 (Sect. 2.3).

In this paper, the Doppler parameter b, which enters in the computation of the optical depth, depends on the thermal velocity and the turbulent velocity $v_{\rm turb}$:

$$b = \sqrt{\frac{2k_bT}{m_{\rm ion}} + v_{\rm turb}^2}. \quad (3)$$

The turbulent velocity increases the probability of interaction of photons with wavelengths away from line center due to gas moving along or opposite to the direction of propagation, and therefore broadening the effective cross-section. This is important due to two limitations of the simulations that challenge a completely self-consistent treatment of absorption studies. The first limitation is the fact that the simulated galaxy is discretized on a grid, which means there are discrete jumps in the gas velocity from one cell of the grid to the next. Therefore, for some conditions, a photon packet with a wavelength corresponding to the velocity $v + \frac{\Delta v}{2}$ could go through two cells with projected velocities v and $v + \Delta v$ without being absorbed, while it would be absorbed at intermediate velocities if the grid was more refined. The turbulent velocity is a way to smooth out those discontinuities. It is possible to estimate this quantity in the simulation by comparing the velocity in a cell with the velocities of neighboring cells, as is done for the star formation criteria of Kimm et al. (2017) or Rosdahl et al. (2018). Doing so in the three outputs of Table 1 yields an average $n_{\rm H\,I}$-weighted turbulent velocity of approximately $12\,{\rm km\,s^{-1}}$.

The second limitation of simulations is the fact that there are always gas motions at scales unresolved by the simulation, which would also increase the probability that photons get absorbed. This is called "subgrid turbulent velocities", which is not taken into account in our simulation or in the previous computation of the turbulent velocity based on neighboring cells. Some other RAMSES simulations quantify those subgrid motions on-the-fly (e.g., Agertz et al. 2015; Kretschmer & Teyssier 2020). In this work we assume a uniform turbulence in every cell of the simulation, accounting for those two kinds of turbulence. We choose a fiducial value of $v_{\rm turb} = 20\,{\rm km\,s^{-1}}$. In Sect. 3.2 we investigate the effects of changing this parameter.

2.5. Dust model

As described in Michel-Dansac et al. (2020), we follow Laursen et al. (2009) to model the effect of dust on radiation. This is a post-processing method, because the dust is not followed directly in the simulation. In this model, the optical depth of dust is the product of a pseudo-density and a cross-section. The pseudo-density is

$$n_{\rm dust} = \frac{Z}{Z_0} \times (n_{\rm H\,I} + f_{\rm ion} n_{\rm H\,II}), \quad (4)$$

where Z_0 is a reference metallicity and $f_{\rm ion}$ sets the quantity of dust that is put in ionized regions. Throughout this paper we use $Z_0 = 0.005$ and $f_{\rm ion} = 0.01$, corresponding to the calibration on the Small Magellanic Cloud (SMC) of Laursen et al. (2009).

For the cross-section, we use the fits of Gnedin et al. (2008) in the SMC case. This cross-section accounts for both absorption and scattering events. RASCAS has an albedo parameter, representing the probability that a photon-packet that interacts with dust is scattered instead of destroyed. For this albedo we use the value of Li & Draine (2001), which is 0.276 for Lyβ and 0.345 for C II λ1334. Finally, the asymmetry parameter g of the Henyey-Greenstein function (Henyey & Greenstein 1941), which influences the probability distribution of the direction in which a scattered photon goes, is also taken from Li & Draine (2001), and is $g = 0.721$ for Lyβ and 0.717 for C II λ1334.

2.6. Computation of escape fractions

To compute the escape fractions of ionizing photons, we first create mock spectra of the galaxy between 10 Å and 912 Å using the method described in Sect. 2.4. The only changes are that, first, we use BPASS but without a power-law approximation of the SEDs, and second, the optical depth is computed differently. The optical depth for ionizing photons is a sum of contributions from H, He, He^+, and dust. The first three are given by Verner et al. (1996) and dust is treated exactly as in Sect. 2.5, including the scattering, with an albedo of 0.24. The mock spectra are then used to compute the total number of ionizing photons going out of the virial radius, which is divided by the intrinsic number to get the escape fraction. We also compute the escape fraction at 900 Å by averaging the spectra between 890 Å and 910 Å. Results are shown in Table 1 and in Sect. 4.

3. Mock spectra

In this section we show a selection of C II λ1334 and Lyβ spectra. From the 5184 mock observations that we made (i.e., 1728 directions of observation for the three outputs A, B, and C), we select seven directions that have absorption profiles that are as diverse as possible, shown in Fig. 3. We highlight that those spectra do not represent the average of the galaxy over time and direction of observation, but they rather depict the large range of possible shapes and properties for one galaxy. Most lines are similar to the one in the top row of Fig. 3, that is to say saturated and wide.

There is a very wide range of shapes of absorption lines for a single galaxy seen from different directions of observation and at time differences of 80 Myr. This shows that the distribution of gas in the galaxy is highly anisotropic. There is however an axis along which the galaxy is more flat, resulting in an irregular disk-like structure. There are slight trends between the properties of the lines and the inclination of this disk: edge-on angles have somewhat stronger absorption than face-on ones. We will explore those relations in a following paper.

The main properties of the lines on which we focus in this paper are the residual flux and the equivalent width (EW). The residual flux is

$$R = \frac{F_{\rm line}^0}{F_{\rm cont}^{\rm observed}}, \quad (5)$$

where $F_{\rm line}^0$ is the flux at the deepest point of the absorption and $F_{\rm cont}^{\rm observed}$ is the flux of the continuum next to the absorption line. The EW corresponds to the area of the absorption line after the continuum has been normalized to one. The EW of the fluorescence corresponds to the area of the emission part above the continuum. Both EWs are defined as positive. We highlight that in several spectra the EW of fluorescence is larger than that of the absorption. This is because fluorescence is an emission process through which the gas emits in all directions. There are directions of observation with almost no absorption but with a strong fluorescence, which is emitted by gas that is excited by photons traveling in other directions. Lastly, we note that to compensate the movement of the galaxy in the simulation box we add $\boldsymbol{v}_{\rm CM} \cdot \boldsymbol{k}_{\rm obs}$ to the velocity axis of Fig. 3, where $\boldsymbol{v}_{\rm CM}$ is the velocity of the center of mass of the stars and $\boldsymbol{k}_{\rm obs}$ is the unit vector pointing toward the observer.

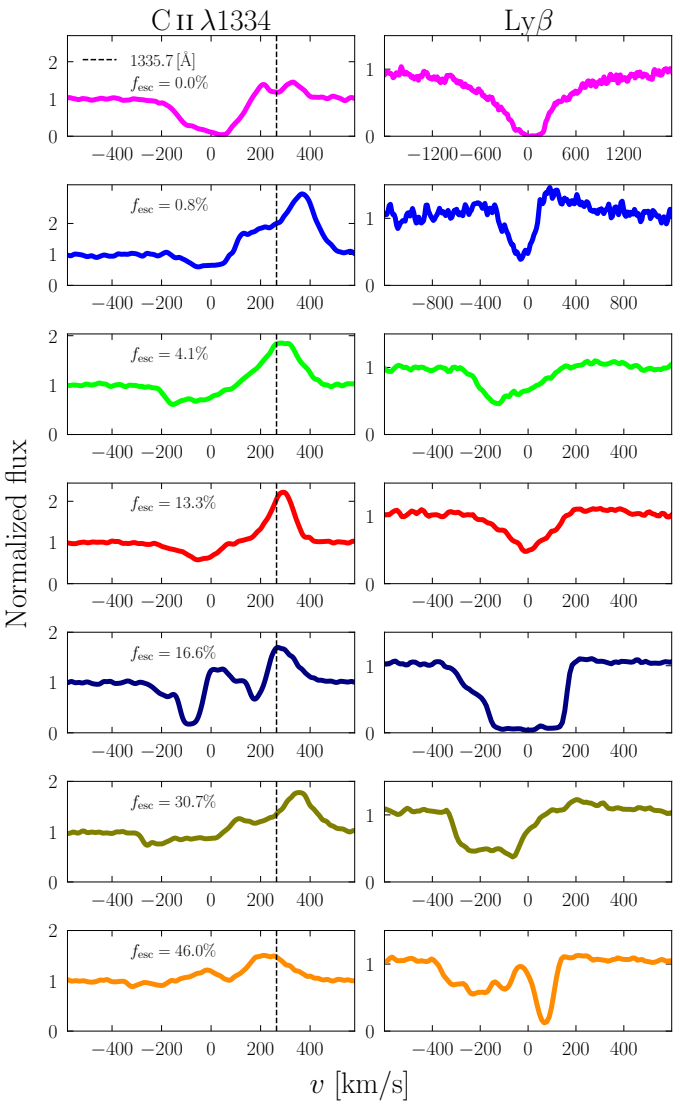

Fig. 3. C II λ1334 and Lyβ absorption lines for seven directions of observation. The vertical dashed lines highlight the wavelength of the fluorescent channels. The spectra are arranged in increasing order of escape fraction of ionizing photons in the direction of observation, f_{esc} (as computed in Sect. 2.6). The noise is Monte-Carlo noise, due to a sampling of the luminosity with a finite number of photon packets. The spectra are smoothed with a Gaussian kernel with $FWHM = 20\,\mathrm{km\,s^{-1}}$. The aperture diameter for the mock observation is a fifth of the virial radius, which corresponds to about 5.6 kpc, or 0.75″ at $z = 3$. No observational effects such as noise, spectral binning, Earth atmosphere, or MW alterations are considered. We note that the x-axis in the first two rows of the second column differ from the others.

3.1. Comparison with observations

In comparison with observed $z \sim 3$ galaxies, our Lyβ and C II λ1334 lines are slightly less wide. In our mock spectra we do not find absorption at very negative velocities. The median velocity at which the absorption reaches 90% of the continuum is $-200\,\mathrm{km\,s^{-1}}$, and the minimum is $-450\,\mathrm{km\,s^{-1}}$. This is to be contrasted with studies such as those of Steidel et al. (2010, 2018) and Heckman et al. (2015) where this velocity is often $-800\,\mathrm{km\,s^{-1}}$. We note however that these observations are averages over many galaxies and typically obtained with lower spectral resolution, meaning that a direct comparison is not really indicative. Furthermore, these absorptions at very negative velocities are produced in outflowing gas, and so an alternative explanation is that our simulation probes galaxies with weaker outflows. This is expected because our galaxy has a much lower mass and SFR than the Lyman Break Galaxies of those studies. However, it is possible that the simulation does not produce enough outflows to be realistic, as is argued in Mitchell et al. (2021), where they use a similar simulation. We will study the gas flows of our galaxy in a forthcoming paper.

In contrast, we find C II λ1334 profiles that are very similar to high-resolution observations of galaxies in the local Universe, as in Rivera-Thorsen et al. (2015) and Jaskot et al. (2019). These latter authors find similar absorption velocities as in our mock spectra, and a large variety of shapes, residual fluxes, and strengths of fluorescence. They even detect fluorescent lines with a P-Cygni profile (i.e., a blue part in absorption and a red part in emission), as in the fifth row of Fig. 3, due to the presence of C^+ in the excited fine-structure level.

Our Lyβ lines are rather different from current observations of either high or low redshift galaxies. Steidel et al. (2018) and Gazagnes et al. (2020) show Lyβ absorption lines that have a large EW and a significant residual flux. Our lines are either wide with a small residual flux, or have a small EW and a large residual flux. This could be due to low signal-to-noise ratios, spectral resolution, or stacking effects. Alternatively, these effects could come from stellar absorption features or Lyβ emission from the gas, neither of which is included in this work.

We note that some Lyβ spectra display some emission redward of the absorption, especially in the second row of Fig. 3, and very little in the last five rows. This is due to scattering effects, and has not been observed yet because it is rare (the first row, which is the most common, has no emission) and requires an excellent signal-to-noise ratio.

3.2. Robustness of the method

Our procedure to make mock observations involves several free parameters. Here we test the extent to which the spectra are affected if we change their values. We focus in particular on the effect on residual flux, which is the quantity that we use the most in this paper. Figure 4 shows the same spectra as in Fig. 3 but with different alterations, which we explain individually below. Figure 5 shows the statistics of the change of residual flux for the 5184 directions of observation, for every setup that we try.

3.2.1. Carbon abundance

As discussed in Sect. 2.2, the simulation does not trace carbon directly and we have to assume an abundance ratio for carbon to infer it in post-processing. This setup incurs uncertainties due to the poorly constrained carbon-to-hydrogen ratio and dust depletion factor in $z = 3$ galaxies, and therefore we vary the density of carbon to see if the result is sensitive to that variation. The first column of Fig. 4 shows the effects of dividing and multiplying the density of carbon by a factor of two. Overall the effect is small. The fluorescence of the second and sixth spectra is affected the most, while the absorption part is never drastically changed.

For a more global view, the left panel of Fig. 5 shows the effects on the 5184 directions. We see that directions with a very small or a very large residual flux are not significantly affected by the change in carbon density. This is because when the residual flux is either very small or very large the line photons have gone through mostly optically thick or thin media, respectively, which are not strongly affected by a change of a factor of two. Alternatively, when the residual flux is between ~0.2 and 0.8, the

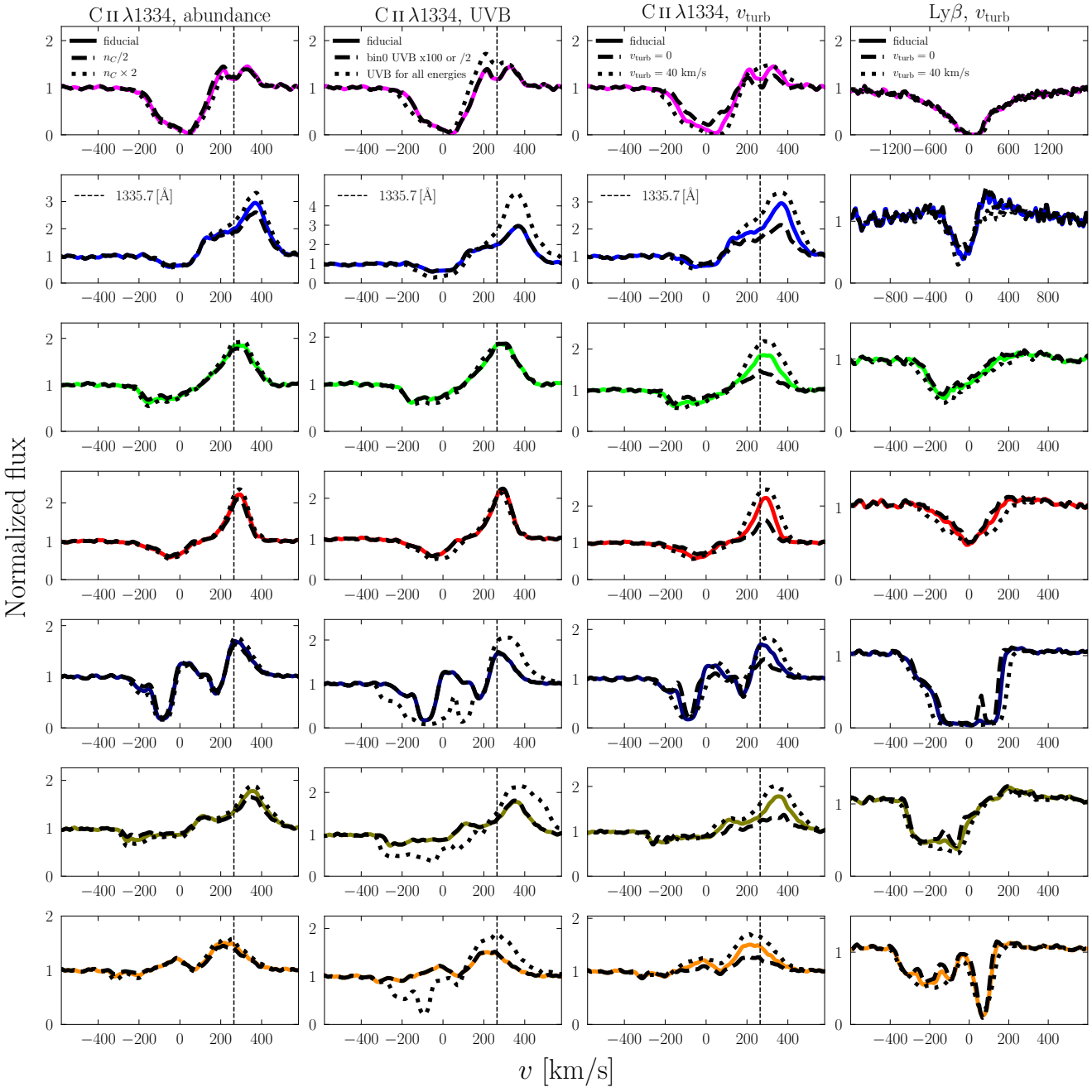

Fig. 4. Test of the robustness of the spectra against changing free parameters. The seven rows show the same seven directions as in Fig. 3. The vertical dashed lines highlight the wavelength of the fluorescent channels. *First column*: test of the effect of multiplying and dividing the fiducial carbon density by two. *Second column*: test of multiplying (dividing) the Habing band UVB by 100 (2) and comparison with the result with only UVB and no radiation field from the simulation. *Third* and *fourth columns*: test of changing the turbulent velocity for C II λ1334 and Lyβ respectively. The fiducial turbulent velocity is 20 km s^{-1}.

gas through which the photons went is neither completely thick nor thin, and therefore a change in density of a factor of two has more of an effect, with a difference of residual flux of up to 0.25. As the carbon-to-hydrogen ratio is poorly constrained, we use solar abundance ratios and test the relations between the absorption lines and the escape fraction of ionizing photons under this assumption.

3.2.2. Implementation of the radiation below 13.6 eV

As the simulation was done with only hydrogen- and helium-ionizing photons, we have to estimate in post-processing the flux of photons between 6 eV and 13.6 eV, which ionize neutral carbon, among other metals. Our fiducial model is to use the Haardt & Madau (2012) UVB in every cell of the simulation, except in cells where $n_{HI} > 10^2$ cm^{-3}, because the optical depth of dust in those cells is enough to extinguish the UVB. We note that removing the UVB in those cells does not impact the absorption lines at all because those cells are so dusty that they hide all the stellar continuum. Additionally, as explained in Sect. 2.2, we take the ionization fractions of hydrogen and helium from the simulation, and therefore we remove all reactions including hydrogen or helium from our KROME chemical network.

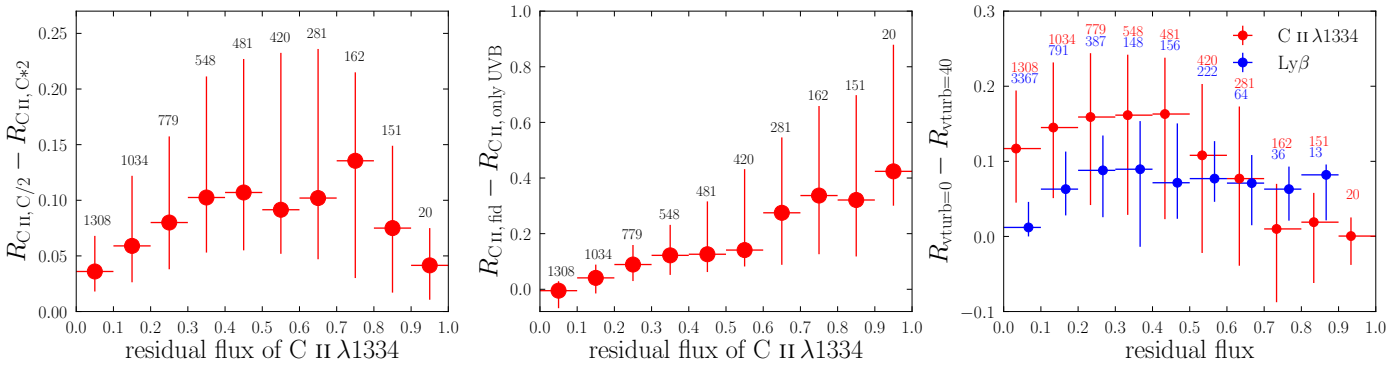

Fig. 5. Effect of different parameters of the modeling on the residual fluxes. The *y*-axes display differences of residual fluxes between two setups. The dots are the median of the differences in a given bin of residual flux, and the bars show the 10th and 90th percentiles. The numbers indicate how many spectra are in the corresponding bin of residual flux. *Left panel*: effects of carbon density, which is divided or multiplied by two. *Middle panel*: effect of UVB radiation for ionizing photons instead of the radiation field from the simulation. *Right panel*: effects of turbulent velocity. Turbulent velocities are varied between 0 and 40 km s^{-1}.

In the second column of Fig. 4 we compare the fiducial spectra with three other setups:

1. The same as fiducial, but dividing the flux between 6 eV and 13.6 eV by two in every cell. This is done to evaluate the effect of reduction of the UVB due to absorption in the galaxy.

2. The same as fiducial, but multiplying the flux between 6 eV and 13.6 eV by 100 in every cell. This is done to simulate the contribution of a young metal-poor stellar population of 1000 $M_\odot$ at a distance of 10 pc (which is roughly the resolution in the ISM) at this energy range.

3. Using the Haardt & Madau (2012) UVB at all energies, ignoring the radiation field from the simulation. We let KROME compute the ionization fractions of every species, including hydrogen and helium, unlike in the fiducial setup. We compute the photoionization rate that the UVB imposes on H^0, He^0, He^+, C^0 and C^+ at the redshift of the output, and we apply this photoionization rate in every cell with a hydrogen density smaller than a certain threshold. This threshold is 10^{-2} cm^{-3} for H^0 photoionization, 4.3×10^{-2} cm^{-3} for He^0, 3.3×10^{-1} cm^{-3} for He^+, 10^2 cm^{-3} for C^0 and 4.3×10^{-2} cm^{-3} for C^+. For each ion this threshold is chosen so that a cell with a size of 10 pc having this density would have an optical depth of around one for photons having the smallest ionizing energy of the ion. We test this setup to assess the importance of using an RHD simulation instead of a hydrodynamics simulation where only a UVB can be used for the photoionization.

The two first setups are represented by the dashed lines in the second column of Fig. 4. They are identical to the fiducial spectra. This is because the Haardt & Madau (2012) UVB below 13.6 eV is high enough to completely photoionize C^0 to C^+ (in all the regions with $n_{\mathrm{HI}} < 10^2$ cm^{-3}), and therefore multiplying it by 100 does not change the result, and dividing it by two is not enough to stop photoionizing C^0. To confirm this we multiplied the UVB by a factor 10 000, because locally the radiation field below 13.6 eV can be even stronger than 100 times the UVB, and the spectra are still not affected. On average, over the 5184 directions the residual flux is changed by only 0.002, both when we divide the UVB by two or multiply it by 100.

In the third setup, the dotted lines in the second column of Fig. 4, the effects are more noticeable. The first, third, and fourth rows are not affected significantly, while the second and the last three are drastically changed. The middle panel of Fig. 5 shows the differences for the 5184 directions. We see that the directions with small residual fluxes are not affected significantly, such as the purple spectrum in the first row of Fig. 4, but that the directions with high fiducial residual fluxes tend to have a much deeper absorption line when we use the UVB instead of the radiation field of the simulation, which is the case for the last three rows of the second column of Fig. 4. This shows that using a UVB at all energies instead of a self-consistent ionizing radiation field simulated on-the-fly does not photoionize C^+ to C^{++} efficiently enough, and therefore globally overestimates the density of C^+, which leads to deeper absorption lines. Another factor that explains the change of C^+ density is the modification of the C^+ recombination and collisional ionization rates due to the change of electron density when we let KROME compute the equilibrium ionization fractions of hydrogen and helium instead of using the values from the simulation. Those results show that it is important to use an RHD simulation to accurately compute the ionization fractions of metals in post-processing.

3.2.3. Turbulent velocity

Turbulent velocity (with a fiducial value of 20 km s^{-1}) is a parameter we add to smooth out the discontinuities of the velocity field in the simulation and to model the subgrid motion of the gas, as we introduced in Sect. 2.4. The third and fourth columns of Fig. 4 show the effect on C II λ1334 and Lyβ of changing the turbulent velocity to 0 and 40 km s^{-1}. We see that Lyβ is almost unchanged, and for C II λ1334, the absorption parts of our seven spectra are not very sensitive to the turbulent velocity either, except the one in the first row. However, fluorescence grows quickly with turbulent velocity, indicating that there is more absorption overall in the galaxy, leading to a stronger fluorescent emission in every direction of observation.

The right panel of Fig. 5 shows the effects of changing the turbulent velocity on our 5184 mock spectra. The C II λ1334 line shows large differences of residual fluxes, of up to 0.25. An average difference of 0.15 in residual flux is seen for directions with a low residual flux, and a smaller difference is seen for directions with a high residual flux. There are even directions where the residual flux is larger for a turbulent velocity of 40 km s^{-1}, which may seem counter-intuitive. This is because a larger turbulent velocity results in more absorption globally in the galaxy, and so there are also more scattering events and more infilling effects in some directions of observation. Some directions are sensitive to infilling, as we show in Sect. 3.3.2, and this infilling can increase the residual fluxes. Figure 5 confirms that Lyβ is less affected by

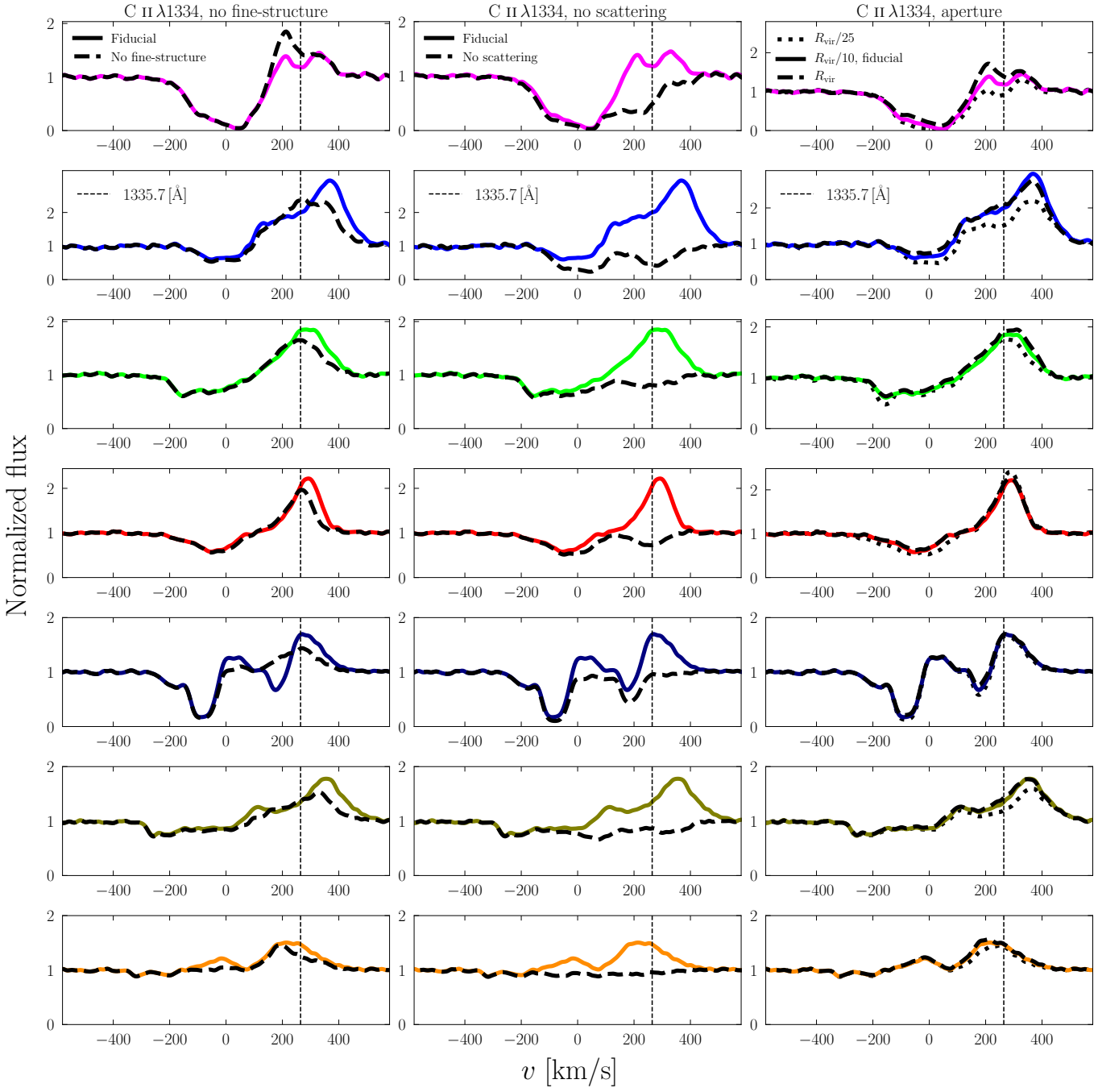

Fig. 6. Illustration of the effects of different physical processes on the same spectra as Fig. 3. The vertical dashed lines highlight the wavelength of the fluorescent channels. *First column*: comparison with and without a fraction of C^+ ions populating their ground-state fine-structure level. *Second column*: effect of removing the scattering of photons. *Third column*: effect of changing the aperture of the observation.

the turbulent velocity parameter than C II λ1334. The difference of residual fluxes with a turbulent velocity of 0 and 40 km s^{-1} is on average smaller than 0.1. The turbulence affects Lyβ less than C II λ1334 because the thermal velocity of hydrogen is about 3.5 times larger than that of carbon.

3.3. Effects of various physical processes

In this section we analyze the effects of various physical processes on the spectra to show that the formation of the line is complex. In Figs. 6, 9, and 10 we show the same spectra as in Fig. 3 but altered by removing a number of processes that are taken into account in the fiducial method. Figures 7 and 8 show the statistics of the change of residual flux for the 5184 directions of observation for some of the processes that we analyze. Below we explain these processes one by one.

3.3.1. Fine-structure level

Around 5% of C^+ ions in the ISM of our galaxy are excited in the fine-structure level of the ground state, leading to absorption at a different wavelength from the true ground state, as explained in Sects. 2.3 and 2.4. One of the resulting differences is that C^+ ions in the excited state can resonantly scatter fluorescent photons at

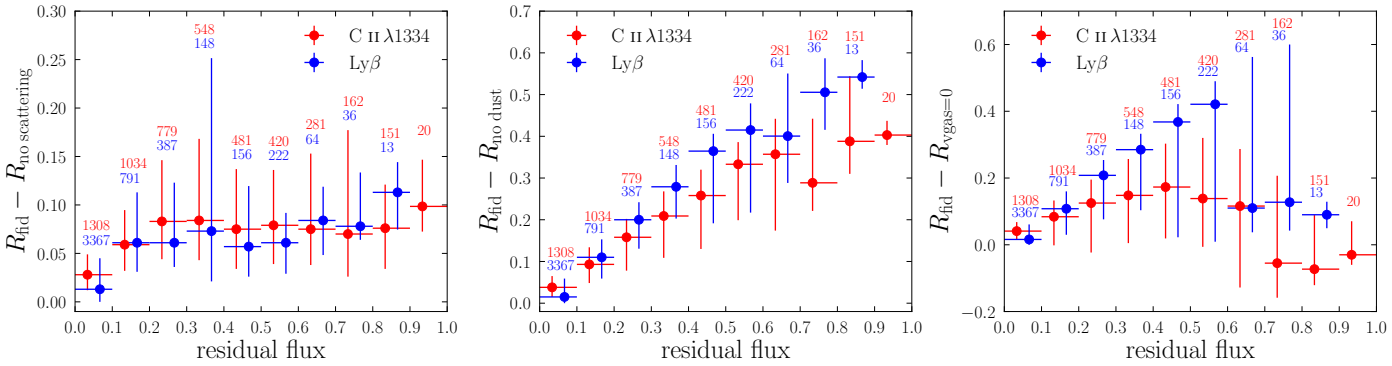

Fig. 7. Same as Fig. 5, to show the effects of different physical processes. *Left*, *middle*, and *right panels*: effects of scattering, dust, and gas velocity, respectively.

1335.66 Å and 1335.71 Å. The first column of Fig. 6 shows how spectra change if we assume instead that all C^+ ions are in the true ground state.

The absorption parts of the spectra are almost unaffected, with an average difference of residual flux of less than 0.01. However, there is significant modification of the fluorescent emission. The distribution of EW of fluorescence is redistributed among the different directions of observations when ions in the fine-structure level are considered because of scattering of resonant photons on those ions. This leads to some directions having a larger EW of fluorescence and other directions a smaller one.

Furthermore, the peak of the fluorescence is redshifted when the fine structure is present, once more due to the scattering of photons in the 1335.66 Å and 1335.71 Å channels. Those photons are trapped by C^+ ions in the fine-structure level; they can only escape if their wavelength is redshifted. The ones that are blueshifted are absorbed by C^+ ions in their true ground state. This behavior is reflected in the statistics of the number of scattering events of the photons. Without fine-structure level splitting there is up to a maximum of 65 scattering events per photon, while this number becomes ∼27 000 when the fine-structure level is added. Finally, the presence of the fine-structure level can lead to fluorescence with a P-Cygni profile, as illustrated in the first column and fifth row of Fig. 6.

3.3.2. Infilling

The scattering of C II λ1334 or Lyβ photons can lead to an effect called infilling. Without scattering, all the photons that are absorbed are destroyed. However, when scattering is taken into account, the photons are re-emitted, either as an Hα photon for Lyβ or a fluorescent photon for C II λ1334, or at the wavelength of the line. In this case, photons that are absorbed can still contribute to the spectra if they are re-emitted resonantly and manage to escape the galaxy. Furthermore, scattered photons can contribute to the spectrum of a different direction of observation than the initial emission. This is the infilling effect. In order to see how much infilling occurs in our mock spectra, we generate spectra in which any absorption leads to the destruction of the photon. This means that the resulting spectra are the sum of Voigt absorption profiles for every cell in front of every stellar particle in the simulation. In this case, no fluorescent photons are emitted, and so the fluorescent C II* λ1335 line disappears.

The second column of Fig. 6 for C II λ1334 and the first column of Fig. 9 for Lyβ show the fiducial spectra and the spectra without scattering. For C II, we see a second absorption around

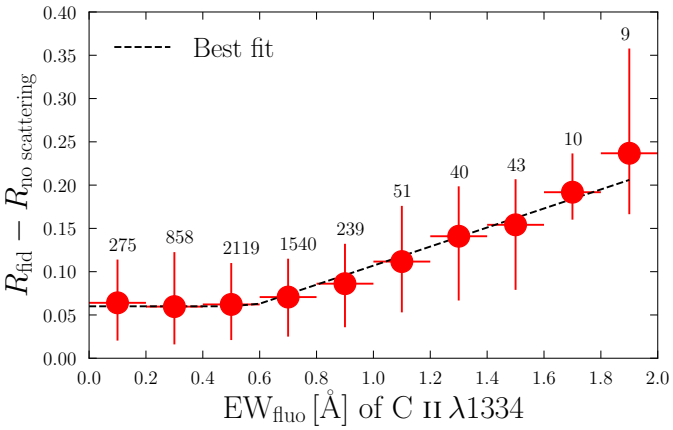

Fig. 8. Same as left panel of Fig. 7, in bins of EW of the fluorescent line instead of residual flux. The best fit is 0.06 for $EW_{fluo} < 0.6$ Å and $0.11\,EW_{fluo} - 0.003$ otherwise.

250 km s^{-1} coming from the presence of C^+ ions in their fine-structure level. The infilling in the absorption is strong for the second row, with a residual flux that is doubled when scattering is included. For the other spectra, the residual flux is only increased by a few percent.

The left panel of Fig. 7 shows the differences of residual flux with and without scattering for the 5184 directions. Both C II λ1334 and Lyβ show the same trends. For residual fluxes below 0.1 the median of the difference is small (less than 0.03). For very high residual fluxes, the difference is larger, with a median of 0.12 for Lyβ and 0.1 for C II λ1334. In between, for residual fluxes from 0.1 to 0.8, the median of the difference varies between 0.05 and 0.1. Overall, the effect of infilling is not very significant on average, but there are extreme directions where the infilling increases the residual flux by more than 0.3.

Figure 8 shows the differences of residual flux with and without scattering, similar as in Fig. 7, but in bins of EW of fluorescence. We choose these bins to show that the effect of infilling is more important for directions with a strong fluorescence. This highlights the fact that there are directions where many scattered photons escape the galaxy. Indeed, both fluorescent photons and photons creating the infilling effect must have scattered at least once. Figure 8 shows that fluorescent photons and resonantly scattered photons both escape the galaxy in preferential directions. This can be used to get a rough estimate of the infilling effect based on the fluorescent line (see Fig. 8).

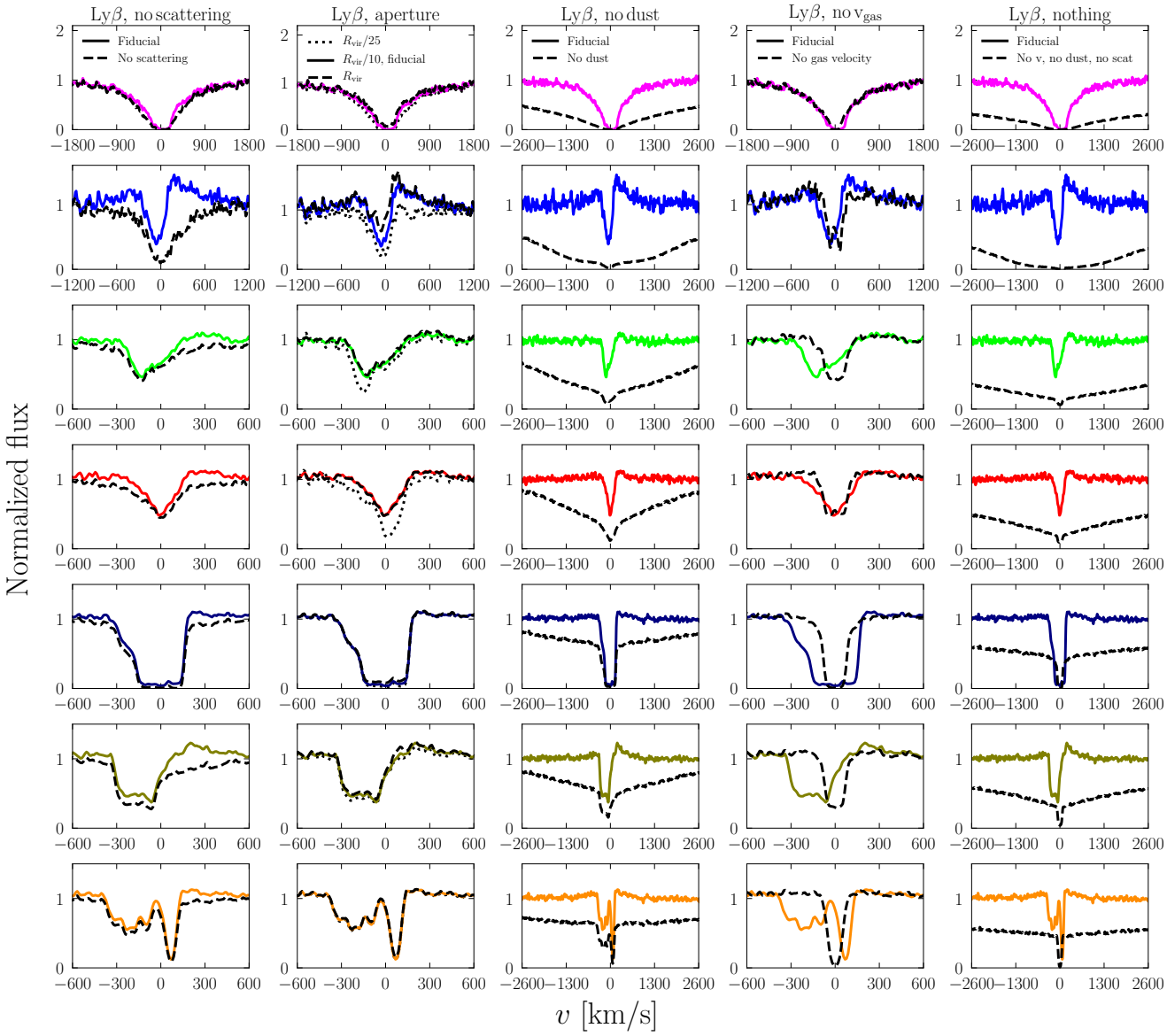

Fig. 9. Same as Figs. 6 and 10 but for Lyβ. No fine-structure level is considered here.

3.3.3. Aperture size

The effect of infilling depends on the aperture used for the observation, because photons can potentially travel far before being scattered and redirected toward the direction of observation (Scarlata & Panagia 2015). The third column of Fig. 6 for C II λ1334 and the second column of Fig. 9 for Lyβ compare the fiducial spectra with different aperture radii: $R_{vir}/25$, $R_{vir}/10$ and R_{vir}, which are approximately 1.1, 2.8, and 28 kpc, or angular diameters of about 0.″3, 0.″7, and 7″ at $z = 3$. Our fiducial spectra are made with the $R_{vir}/10$ aperture, because it includes as much as ≈95% of the stellar continuum emission. The R_{vir} aperture is used to assess the effect of adding the light scattered in the circum-galactic medium (CGM) and the $R_{vir}/25$ aperture, which contains ≈40% of the stellar continuum, is useful to show what happens if we do not enclose the whole stellar-continuum -emitting region in the mock observation. We highlight that the aperture is a property of the mock observation, and it does not affect the radius of propagation of the photon packets, which is always one virial radius (Sect. 2.4).

The result is that the residual flux is often barely affected by the aperture of observation, except for the small aperture in the second, third, and fourth rows of Fig. 9. It is noticeable in general that smaller apertures lead to smaller residual fluxes. This is in part because a smaller aperture misses some of the scattered photons, but is mainly because a smaller aperture misses flux from stars that are facing optically thin gas. The light from those stars is not absorbed, and therefore it increases the residual flux if the stars are in the aperture of the observation. Similarly, the fluorescence increases for larger apertures, as does the emission part of Lyβ, because there are photons scattering far away that are missed with smaller apertures. Nevertheless, the differences when taking a larger aperture are small, and the aperture size has little importance as long as it encompass more than ≈90% of the stellar continuum. For example, an aperture corresponding to 1″ gives a result very similar to our fiducial aperture. As discussed in Mitchell et al. (2021), where a simulation similar to ours is used, the small effect of increasing the aperture may be due to a lack of neutral gas in the CGM, possibly due to imperfect supernova feedback modeling.

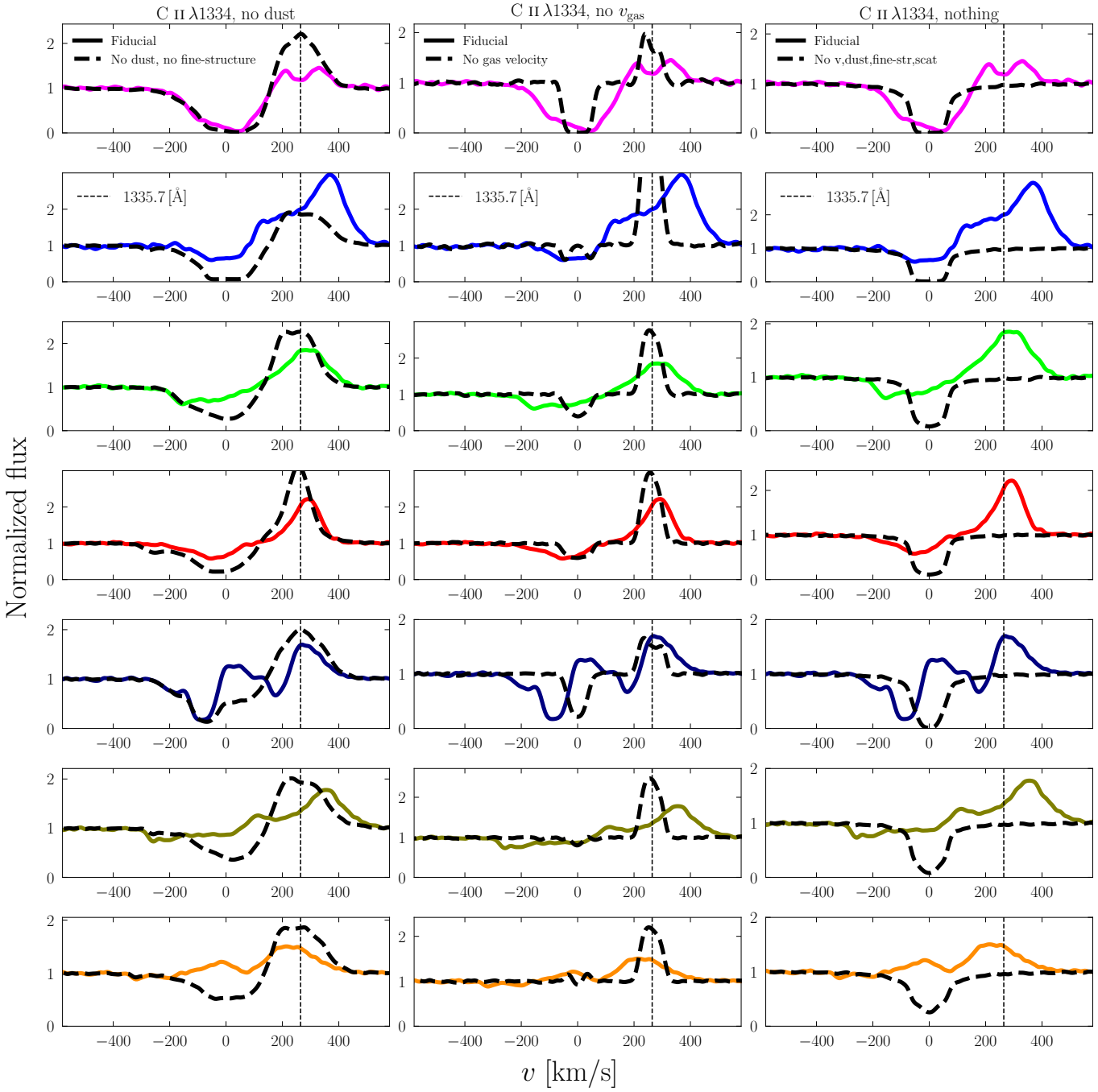

Fig. 10. Same as Fig. 6, with other processes. *First column*: effect of removing dust. We also remove the fine-structure level to simplify the interpretation. *Second column*: spectra with and without gas motion in the galaxy. Infilling effects give complex shapes to the *second* and *seventh rows*. *Third column*: comparison of fiducial spectra with spectra with neither gas motion, dust, scattering, nor fine-structure level.

3.3.4. Dust

Now we study the effects of dust on the spectra. The third column of Fig. 9 for Lyβ and the first column of Fig. 10 for C II λ1334 compare the fiducial spectra with the spectra made without dust. To avoid obtaining a complex shape for the fluorescence, which can be misleading in this case, we omit here the fine-structure of C^+. Every C^+ ion is assumed to be in the ground state. All the spectra that we plot are normalized, and so we do not see the difference in the continuum when dust is removed, which would be significantly attenuated. Averaging over the 5184 spectra, the continuum is reduced by 74% at 1330 Å and by 83% at 1020 Å.

One striking difference in the spectra is that the Lyβ lines have especially large EWs when dust is removed. This is because there are stars embedded in optically thick regions, with $N_{\rm HI} > 10^{22}\,{\rm cm}^{-2}$, causing absorption up to high velocities, yet those stars do not contribute to the spectra when dust is present because they are completely attenuated. Another result is that the absorption lines have smaller residual fluxes when dust is absent. This is because the flux at the deepest point of the line does not change significantly with or without dust, whereas the flux of the continuum does. Thus the residual flux, which is the ratio of those two fluxes, is larger with dust than without. The flux of the line center does not change significantly because dust

and H^0 or C^+ all reside in the neutral regions, and therefore the line center photons have generally a higher probability of being absorbed by the gas than by dust.

The middle panel of Fig. 7 shows the difference in residual flux between the spectra with and without dust for the 5184 directions. For very small fiducial residual fluxes, the difference is almost zero. However, for directions with higher residual fluxes, the spectra without dust have a much smaller residual flux than the ones with dust.

3.3.5. Gas velocity

The gas in front of the stars moves at different velocities, leading to an absorption that is spread in wavelength. This can lead to an increase of the residual flux compared to a case where all the gas has the same velocity. For example, even if each star is facing a column density high enough to saturate the absorption line, the residual flux can be larger than zero if the gas is not going at the same velocity for each star. For the line to be saturated, with zero residual flux, there must be at least one velocity regime for which all the stars are facing a sufficiently high column density of gas going at this (projected) velocity.

To show that the spectra in our simulation are sensitive to the velocity of the gas, the fourth column of Fig. 9 for Lyβ and the second column of Fig. 10 for C II λ1334 show the spectra with and without gas velocity. The EW is generally smaller without gas velocity, because the absorption is not spread in wavelength. Moreover, there are scattering effects that give complex shapes to the absorption profiles, especially in the second row for both lines and in the seventh row for C II λ1334.

The right panel of Fig. 7 shows the difference of residual flux between the spectra with and without gas velocity. As for the case without dust, the difference is negligible for very small fiducial residual fluxes and increases for higher residual fluxes, especially for Lyβ. For very high residual fluxes, above 0.6, the difference becomes smaller again. For C II λ1334, there are even directions where the residual flux is larger when the gas velocity is zero, because of scattering effects.

3.3.6. All processes removed

Finally, we show in the last column of Fig. 9 for Lyβ and of Fig. 10 for C II λ1334 the spectra with all four processes removed: fine-structure level, scattering, dust absorption, and gas velocity. Those spectra are more saturated because all four processes that we remove tend to increase the residual flux. The spectra still do not resemble Voigt profiles because of the fact that there are many sources, each facing a different column density of gas.

To summarize this section, we find that our predictions for the shape of the C II λ1334 and Lyβ lines are rather robust against the free parameters of our modeling, such as metal abundances, nonionizing UVB estimations, and subgrid turbulent velocity. However, there are different processes that are crucial for the formation of the absorption lines. The absorption by dust and the velocity field are the most important, while the absorption from the fine-structure level, the infilling, and the aperture effects alter the lines to a lesser extent.

Additionally, dust has a counter-intuitive impact on the shape of absorption lines: young and luminous stars are often not contributing to the spectrum because of selective suppression of the stellar continuum emerging from the dense and dusty star-forming regions (see also Sect. 5.1), which is why dust increases the residual flux.

4. Escape fractions of LyC

We now focus on the escape fractions of ionizing photons in the simulated galaxy. In order to compare the absorption lines and the escape fractions of ionizing photons, we compute the escape fraction from the galaxy as seen from different directions in post-processing, as explained in Sect. 2.6. The correlations between the mock absorption lines and the escape fractions are studied in Sect. 5.

4.1. Distribution of escape fractions

The resulting escape fractions averaged over all directions are 11% for output A ($z = 3.2$), 3.8% for output B ($z = 3.1$) and 1% for output C ($z = 3.0$). Output A has the largest escape fraction of all outputs with $z < 7$ in this simulation. Around 10% of the last 35 outputs, at $z < 3.5$, are similar to output B, with an average $f_{\rm esc} \sim 3$–4%. The vast majority, around 90% of outputs, have similar escape fractions to that of output C, with an average $f_{\rm esc} \leq 1\%$. Such large variations of the escape fraction with time are expected because the escape of ionizing photons is regulated by small-scale processes, essentially the disruption of star-forming clouds by radiation and supernova explosions, which have typical timescales of the order of a few million years (e.g., Kimm et al. 2017; Trebitsch et al. 2017). Also, at output A, the galaxy has been forming stars relatively intensely over the past ∼10 Myr, at around $5\,M_\odot\,{\rm yr}^{-1}$, relative to the average of around $3\,M_\odot\,{\rm yr}^{-1}$ from $z = 3.5$ to $z = 3$, and is thus undergoing strong feedback. As can be seen from Sect. 5.3.1 below, this results in lowering the covering fraction of neutral gas in front of stars, and therefore increasing the escape of ionizing radiation. We choose these outputs, in particular output A, to study the largest variety of escape fractions possible, despite the fact that output A is not representative of the typical state of the galaxy.

Figure 11 shows the histogram of escape fractions in the 1728 directions of observation of the galaxy, for the three epochs A, B, and C. The three outputs show very different distributions of escape fractions, even though they are the same galaxy with 80 Myr between the outputs. As for the spectra in Sect. 3, there is a large variety of escape fractions depending on the direction of observation of the galaxy. This large variety implies that a measurement of escape fraction of a galaxy does not necessarily give the correct contribution of the galaxy to the ionizing photon background (e.g., Cen & Kimm 2015). One needs a statistical study of a large sample of galaxies to get information on the potential of galaxies to reionize the Universe.

Additionally, we show in Fig. 12 the comparison between the global escape fraction of ionizing photons and the escape fraction at 900 Å, $f_{\rm esc}^{900}$. We find that $f_{\rm esc}^{900}$ is mostly smaller than $f_{\rm esc}$, from a few percent to around 12% lower. There are a few points with small escape fractions where $f_{\rm esc}^{900}$ is slightly larger than $f_{\rm esc}$ because of the effect of helium absorption.

4.2. The effect of helium and dust

As described in Sect. 2.6, the computation of escape fractions includes helium and dust in addition to hydrogen. We want to highlight that in our simulation the role of helium and dust is very subdominant, as also found by Kimm et al. (2019). If we remove helium, the escape fractions smaller than 5% increase on average by only about 0.15% (absolute difference), while the larger escape fractions increase by around 1–3%. Those differences are small because helium absorbs only ionizing photons

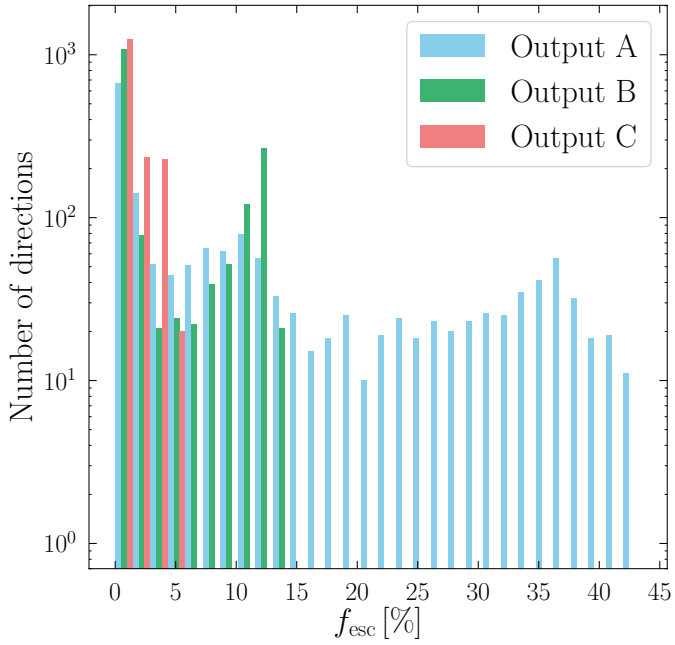

Fig. 11. Histogram of LyC escape fractions for the three outputs, computed as in Sect. 2.6.

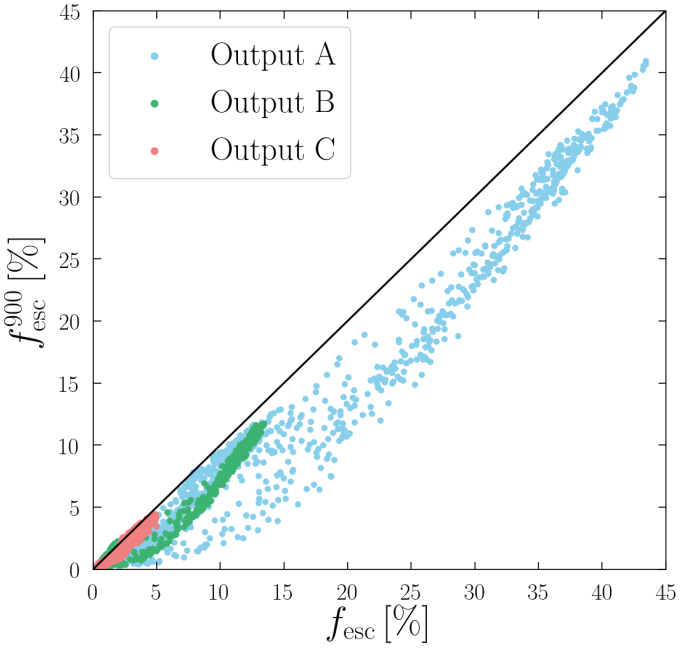

Fig. 12. Comparison of the escape fraction at 900 Å and the global escape fraction of ionizing photons.

with energies larger than 24.59 eV, which are less numerous than the ones between 13.6 eV and 24.59 eV.

When dust is removed, the change is even smaller, with an average increase of escape fractions of 0.05%, and a maximum increase of 1.1%. This small effect comes from our dust distribution model (Eq. (4)), where the density of dust is proportional to the density of neutral hydrogen. This implies that for ionizing photons the optical depth of dust is always much smaller than the optical depth of H^0. In other words, dust alone would have a strong impact on the ionizing photons, but neutral hydrogen absorbs them much more efficiently, such that dust becomes ineffective. This is in agreement with other studies that find little or no effect of dust on the escape fraction (Yoo et al. 2020; Ma et al. 2020). However, in Ma et al. (2020) the escape fractions in the most massive galaxies are substantially affected by dust, but these galaxies are more massive than ours. Additionally, their dust modeling puts dust in gas with a temperature of up to 10^6 K, whereas in our simulation dust is proportional to $n_{\rm H\,I} + 0.01 n_{\rm H\,II}$, and so the dust is almost completely absent in gas hotter than 3×10^4 K. We note that a model of dust creation and destruction in the simulation could yield a different dust distribution, for example with more dust in ionized regions, which could then affect ionizing photons more, but this is beyond the scope of this paper.

5. The link between escape fractions and absorption lines

Now we turn to the search of correlations between properties of absorption lines and the escape fraction of ionizing photons. Figure 13 shows the relations between the escape fractions and some properties of C II λ1334 on the left and Lyβ on the right, for the three outputs and the 1728 directions of observation of each output. We find that none of the observables give a clear indication of escape fraction in all ranges of $f_{\rm esc}$. We can however see some general trends.

For C II, the escape fraction is almost always smaller than the residual flux, which may therefore be used as an upper limit for the escape fraction. This means that a zero residual flux implies $f_{\rm esc} = 0$. The EW of absorption may also be used to get an upper limit because the following relation holds approximately (for our simulated low-mass galaxy): $f_{\rm esc} < 0.5 - 0.4 \times {\rm EW_{abs}}$ [Å]. In particular, directions of observation with ${\rm EW_{abs}} > 1.25$ Å have $f_{\rm esc} < 2\%$. The EW of the fluorescence is a poorer indicator of $f_{\rm esc}$ than the EW of the absorption, but a high fluorescent EW also implies small $f_{\rm esc}$. More precisely, ${\rm EW_{fluo}} > 1$ Å implies $f_{\rm esc} < 2\%$. This is compatible with Jaskot & Oey (2014), who find a strong fluorescent line in a galaxy whose Lyα profile hints at a nonzero escape of LyC. Finally, the three velocity indicators, $v_{\rm max}$, $v_{\rm cen}$, and v_{90} show the least correlation with $f_{\rm esc}$. One potential relation is that directions with $v_{\rm cen} < -250\,{\rm km\,s^{-1}}$ have $f_{\rm esc} \gtrsim 10\%$, which is a similar behavior to what is observed by Chisholm et al. (2017). However we do not have enough data points in this velocity regime to be confident that $f_{\rm esc}$ could not be smaller.

For Lyβ, the relations are even more scattered. The residual flux does not seem to trace $f_{\rm esc}$ at all. A high EW again implies a small escape fraction. Indeed, we find that $f_{\rm esc} < 2\%$ when ${\rm EW_{abs}} > 3$ Å. The directions with ${\rm EW_{abs}} < -0.6$ Å, that is to say with Lyβ dominantly in emission, have $f_{\rm esc} < 2\%$. However, there are not many points, and so this regime might be undersampled. As in the case of C II, a very small $v_{\rm cen}$ implies a high $f_{\rm esc}$, but again the number of points is too small to be confident. Finally, a very small v_{90} is a sign of low $f_{\rm esc}$: $v_{90} < -700\,{\rm km\,s^{-1}}$ implies $f_{\rm esc} < 2\%$, except for two outliers which are above this relation. However, those rare values of v_{90} are arising only when the EW is very large, and therefore they do not bring new information.

In summary, there is one regime where the absorption lines can provide reliable information about $f_{\rm esc}$. When the lines are saturated and wide, with one attribute generally accompanying the other, $f_{\rm esc}$ is smaller than 2%. For Lyβ, being saturated is not a sufficient condition to have a low escape fraction. In addition, the EW has to be larger than 3 Å. A large EW of fluorescence

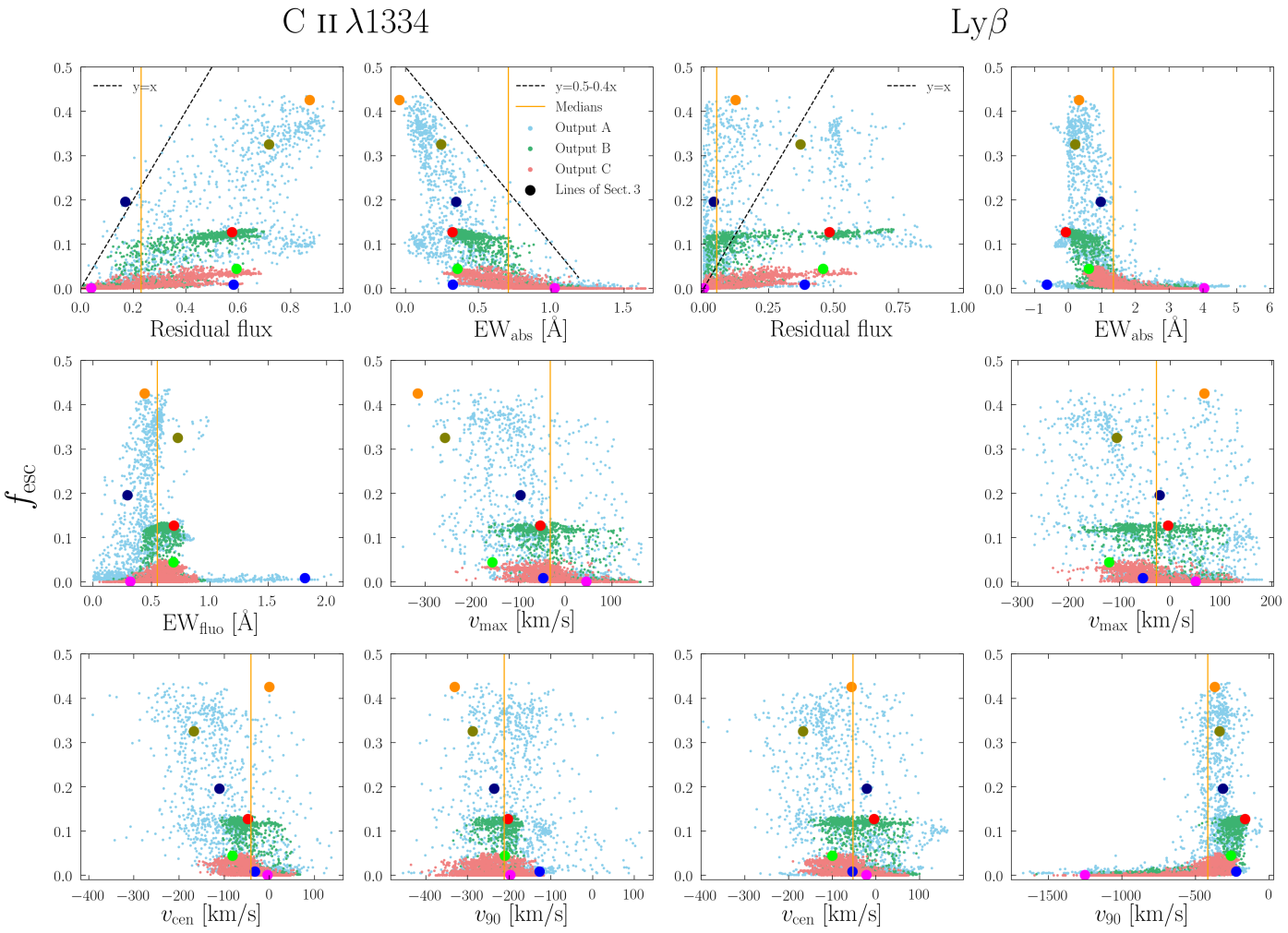

Fig. 13. Scatter plots comparing the escape fractions for the three outputs in the 1728 directions of observation to the properties of the C II absorption line on the *left* and the Lyβ line on the *right*. The seven colored dots show the position of the spectra displayed in Sect. 3. EW_{abs} is the EW of the absorption line (negative EW_{abs} means more emission than absorption), EW_{fluo} is the EW of the fluorescent emission, v_{max} is the velocity of the deepest point in the spectrum, v_{cen} is the velocity at which the absorption line has half its EW, and v_{90} is the velocity where the blue side of the absorption line reaches 90% of the continuum value. The orange vertical lines show the median value of the x-axis variable over the 5184 directions. There is no fluorescent channel for Lyβ, and so one panel is missing.

also implies a low f_{esc}, and is sometimes not linked with a large EW of absorption, as is the case for the large blue dot in Fig. 13. For the other regimes, there is always significant scatter and no predictive relations. To gain more confidence in the relations we find here, they must be verified in other simulated galaxies with different masses and metallicities, which we postpone to future work.

It might be surprising that there is little correlation between the escape fractions and the properties of the absorption lines, especially for the residual flux. In contrast, it is common in the literature to use the residual flux of LIS absorption lines to evaluate the covering fraction of neutral gas and deduce an estimate of the escape fraction of ionizing photons (e.g., Heckman et al. 2001; Shapley et al. 2003; Grimes et al. 2009; Jones et al. 2013; Borthakur et al. 2014; Alexandroff et al. 2015; Reddy et al. 2016; Vasei et al. 2016; Chisholm et al. 2018; Steidel et al. 2018). This is because there are idealized geometries where the escape fraction of ionizing photons and the residual flux of the LIS absorption lines are equal. If we imagine a plane of stars covered by a slab of gas composed of optically thick clouds and empty holes, the ionizing photons and the line photons can only escape through the holes, and not through the thick clouds. In this case, the escape fraction of ionizing photons and the residual flux of the lines are the same, and are equal to the fraction of the surface of the sources that is behind holes, which is referred to as the covering fraction. This is usually called the picket-fence model, implemented with some variations (e.g., Reddy et al. 2016; Steidel et al. 2018; Gazagnes et al. 2018). Several authors have warned that the residual flux can only give a lower limit on the covering fraction, for example due to the velocity distribution of the gas (e.g., Heckman et al. 2011; Jones et al. 2012; Rivera-Thorsen et al. 2015). We argue that there are other effects that complicate the relations between the lines and the escape fractions of ionizing photons, that we detail in this section.

There are several assumptions linked with the picket-fence model. The sources of ionizing photons and nonionizing UV continuum have to be the same so that they trace the same screens of gas. If the continuum of the lines and the ionizing photons are passing through different media, it is impossible for the absorption lines to be a perfect tracer of the escape fraction of ionizing photons. Concerning dust, there are different

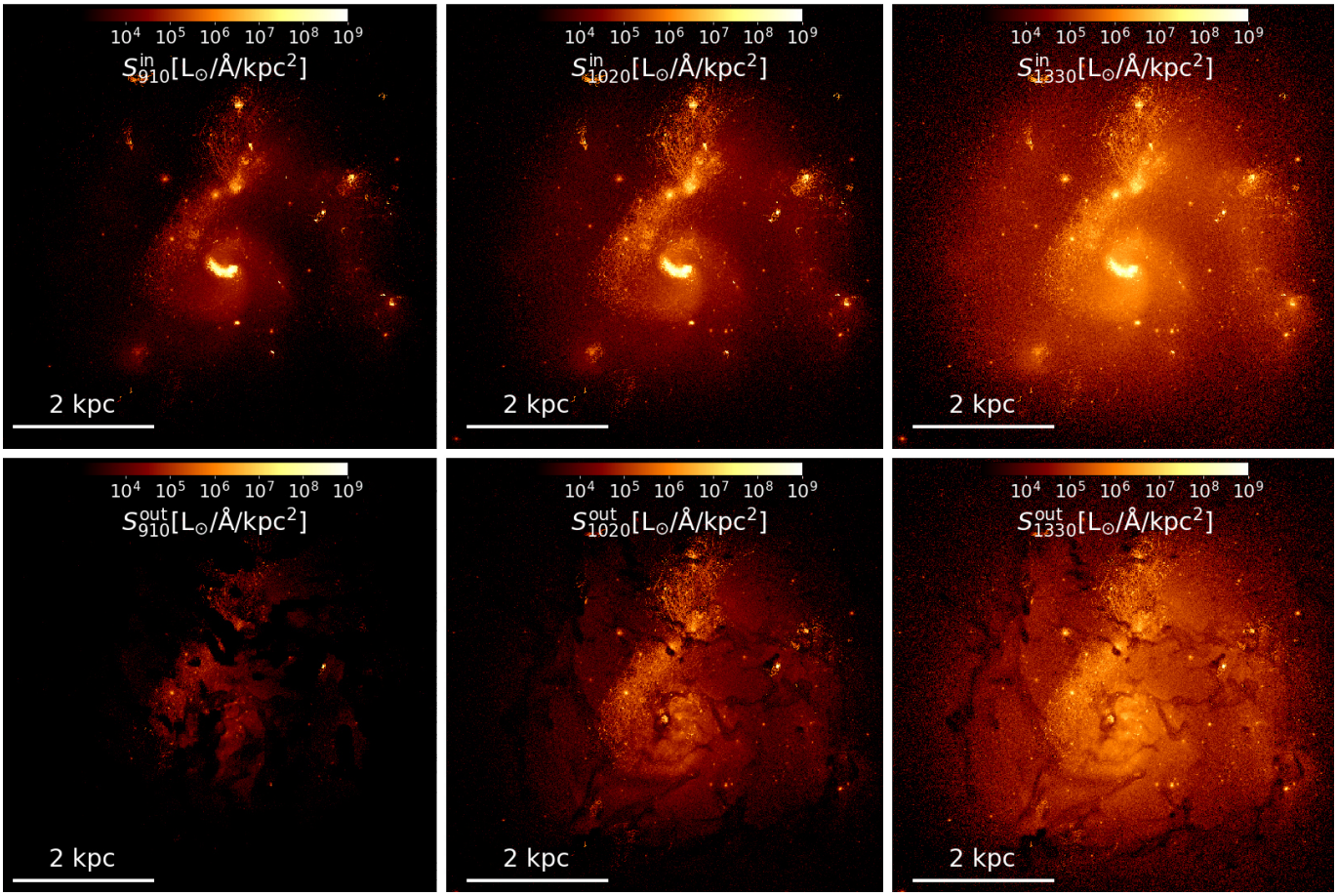

Fig. 14. Surface brightness maps of output C (z = 3.0). *Upper left*: intrinsic surface brightness at 910 Å. *Upper middle*: intrinsic surface brightness at 1020 Å. *Upper right*: intrinsic surface brightness at 1330 Å. *Lower left*: surface brightness at 910 Å after H^0 and dust absorption. *Lower middle*: surface brightness at 1020 Å after dust extinction. *Lower right*: surface brightness at 1330 Å after dust extinction.

treatments in the literature. Some studies do not apply dust corrections, similarly as in Fig. 13. The assumption for the residual flux of absorption lines to be equal to the escape fraction of ionizing photons without dust correction is based on dust being homogeneous, inside and outside of the holes, and the idea that the ionizing photons are affected by dust in the same way as the nonionizing UV continuum. In other studies, the residual flux is corrected under the assumption that the holes have no dust, or to account for the different effects of dust on ionizing and nonionizing photons (e.g., Steidel et al. 2018; Gazagnes et al. 2018; Chisholm et al. 2018). Additionally, one assumption is that the distribution of column densities has to have optically thick parts and holes empty enough to be optically thin for both the ionizing photons and the line photons. In this setup, it is straightforward to define a covering fraction, to measure it with the residual flux of absorption lines, and to deduce the escape fraction as one minus the covering fraction. If there are column densities such that the optical depth is around one, the concept of covering fraction is less well defined, and the escape fraction is no longer one minus the covering fraction. A final assumption is that the infilling effects and the gas velocities do not significantly affect the absorption lines, which would alter the measurement of the covering fraction.

In the following, we compare our simulation with the picket-fence model, addressing these assumptions one by one to decipher the similarities and differences and to explain the large dispersion in the relations between the escape fraction of ionizing photons and the residual flux of the absorption lines. In particular, we apply a dust correction to the residual fluxes using observable quantities, in order to improve the correlations.

5.1. Distribution of the sources

In this section we assess whether the ionizing photons and the Lyβ or C II λ1334 photons are emitted from the same place in order to decipher whether or not they probe the same regions of gas. From stellar models, we know that only stars younger than ~10 Myr are bright at <910 Å (i.e., ionizing energies), whereas older stars can still be bright at 1020 Å and 1330 Å, corresponding to the continuum next to the absorption lines. This is the main driver for the difference between sources of ionizing photons and nonionizing photons. The first row of Fig. 14 illustrates those differences. It shows surface brightness maps for one particular direction of observation of the output C, at different wavelengths. Compared to the 910 Å map, we see that at 1020 Å there are more stars shining in an extended disk around the central part of the galaxy, and even more so at 1330 Å.

To quantify those differences, we note that the brightest 11% of the stellar particles emit 99% of the 910 Å photons, while they emit around 98% of the 1020 Å photons and 85% of the 1330 Å photons. More detailed results for the three outputs are provided in the first two columns of Table 2. While the sources of 1020 Å photons are almost the same as the 910 Å photon sources, the

Table 2. Fraction of the light that is produced by the stellar particles responsible for 99% of the emission of 910 Å photons.

Output	1020 Å	1330 Å	Visible 1020 Å	Visible 1330 Å
A	97.8%	85.1%	$95.1^{+3.4}_{-9.6}$%	$76.8^{+10.8}_{-26.9}$%
B	97.9%	87.5%	$87.9^{+6.4}_{-9.7}$%	$68.4^{+7.3}_{-10.1}$%
C	98.0%	85.6%	$84.1^{+7.8}_{-13.2}$%	$58.9^{+8.8}_{-13.9}$%

Notes. "Visible" means after dust extinction. The results with error bars indicate the mean and the 10th and 90th percentiles.

distribution of sources of 1330 Å photons is noticeably different from that of 910 Å photons. Around 15% of the luminosity at 1330 Å is emitted by stars that do not contribute to the ionizing photon budget.

However, it is the dust-attenuated 1020 Å or 1330 Å photons that are directly observable, and not the intrinsic luminosities. Here, we refer to the UV nonionizing photons after dust extinction as the "visible" photons. Once dust is added, the most luminous stars, namely the young stars in dense, neutral regions, are strongly attenuated. As those stars are the main sources of ionizing photons, the distribution of sources of visible photons becomes very different from the distribution of ionizing radiation sources. The middle and right panels of the second row of Fig. 14 illustrate the effect of dust on 1020 Å and 1330 Å photons. Various bright regions are strongly attenuated, leading to a substantially different distribution of sources than for 910 Å photons. The bottom left plot of Fig. 14 shows the 910 Å surface brightness after H^0 and dust absorption for illustration.

We now compute the fraction of visible 1020 Å and 1330 Å photons emitted by the stars responsible for 99% of ionizing photon emission to assess whether or not the visible photons are tracing the gas screens through which ionizing photons travel. The results are listed in the last two columns of Table 2. When the attenuation by dust is included, the result depends on the direction of observation, and so we present the mean, and the 10th and 90th percentile of the results. In summary, we find that around 15% –but sometimes up to more than 30%– of visible 1020 Å photons are emitted by stars that do not contribute to the budget of ionizing photons, and therefore those photons are affected by screens of gas that do not influence the escape fraction of ionizing photons. For visible 1330 Å photons, the effect is even more pronounced, with around 30% and up to 60% of photons passing through gas that does not screen LyC sources. This is one reason why the residual flux of absorption lines before dust correction of the continuum cannot be a good proxy of the escape of ionizing photons. It is also one motivation to introduce dust correction of the UV nonionizing continuum, as is often done in the literature (e.g., Steidel et al. 2018; Gazagnes et al. 2018; Chisholm et al. 2018).

5.2. Dust correction

In addition to what we see in Sect. 5.1, another motivation to apply a dust correction is the different effects that dust has on both ionizing photons and on the nonionizing UV continuum in our galaxy. We highlight in Sect. 4.2 that dust absorbs very few ionizing photons because the optical depth of neutral hydrogen is always much larger than the optical depth of dust. However, the flux of the continuum next to the absorption line, which enters in the definition of the residual flux of the line, significantly depends on dust. This suggests that a better proxy of the escape fraction of ionizing photons is the ratio of the flux at the bottom of the line to the flux of the intrinsic continuum, rather than the residual flux we plot in Fig. 13, which uses the continuum after dust extinction. To obtain this new ratio, that we refer to here as the dust corrected residual flux, we multiply the residual flux (see Eq. (5)) by the dust extinction factor of the continuum:

$$R^{\text{dust corrected}} = \frac{F^0_{\text{line}}}{F^{\text{intrinsic}}_{\text{cont}}} = R \times \frac{F^{\text{observed}}_{\text{cont}}}{F^{\text{intrinsic}}_{\text{cont}}}, \quad (6)$$

where $F^{\text{intrinsic}}_{\text{cont}}$ is the value of the continuum next to the absorption line before dust attenuation. Fortunately this dust extinction factor $F^{\text{observed}}_{\text{cont}}/F^{\text{intrinsic}}_{\text{cont}}$ can be estimated by SED fitting, and therefore the dust correction is feasible in practice. The extinction factor is usually parametrized as $10^{-0.4\,E_{B-V}\,k_\lambda}$, where E_{B-V} models the strength of the dust attenuation and k_λ is a law of dust attenuation as a function of wavelength. However, the extinction factor can be obtained directly from the simulation after the transfer of the stellar continuum with RASCAS, which is why we do not use the observational method to apply the dust correction. This dust correction is very similar to the ones in Steidel et al. (2018), Gazagnes et al. (2018) and Chisholm et al. (2018). The only difference is that we use the extinction factor at the wavelength of the continuum next to the absorption lines, while these latter authors use the extinction factor at 912 Å. This is conceptually different but gives similar results.

Figure 15 shows the new relations between the escape fraction of ionizing photons and the dust-corrected residual flux of C II λ1334, on the left, and Lyβ, on the right. The results are more promising than before dust correction, in particular for C II λ1334. The corrected residual flux still shows significant scatter but now follows the one-to-one relation more closely, especially for directions with high escape fraction. Very low corrected residual fluxes still indicate very low escape fractions. For directions with dust-corrected residual fluxes of between 0.02 and 0.3, the prediction of f_{esc} using the residual flux is unfortunately often too high. Only around 18% of those directions have an escape fraction that is at least equal to 80% of the corrected residual flux. Above a corrected residual flux of 0.3, the predictions are more successful, with around 80% of the directions having an escape fraction at least equal to 80% of the corrected residual flux.

After dust correction, Lyβ still traces f_{esc} more poorly than C II λ1334 does, but we see that the dust-corrected residual flux of Lyβ now provides a good lower limit for f_{esc}, whereas before dust extinction, in the right panel of Fig. 13, the points were extremely dispersed.

5.3. Covering fractions and escape fractions

We now turn to the question of the distribution of the column densities and the concept of covering fractions. At the beginning of Sect. 5 we explain that in order to have equality between the residual flux of absorption lines and the escape fraction of ionizing photons, the screen of gas must be split in two parts: some optically very thick regions and some empty holes. The escape fraction is then equal to one minus the covering fraction f_C, which is the fraction of sources that are covered by optically thick gas. If there is gas with a column density such that $\tau_{910} \approx 1$, the covering fraction becomes harder to define and it is no longer true that $f_{\text{esc}} = 1 - f_C$.

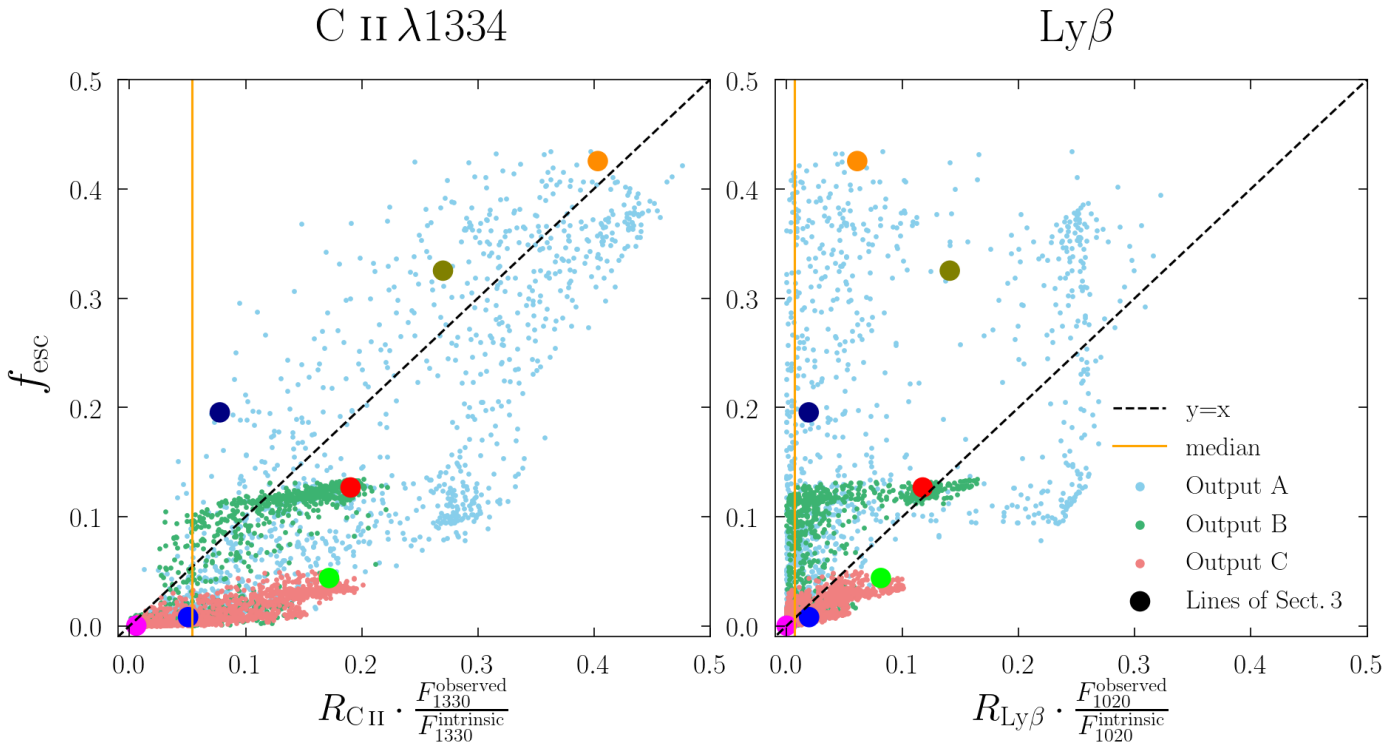

Fig. 15. Scatter plot of the escape fraction of ionizing photons and the dust-corrected residual flux for C II λ1334 on the *left* and Lyβ on the *right*. The residual fluxes are multiplied by the ratio between the flux of the continuum with dust extinction and the flux of the continuum without dust extinction.

In the first part of this section, we study the distribution of neutral hydrogen column densities that the ionizing photon sources are facing. We then define the covering fraction in the simulated galaxy and see how well it traces f_{esc} in the same three outputs and 1728 directions of observation as in the previous sections.

An additional difficulty arises when using absorption lines to probe this covering fraction: Lyβ does not have the same optical depth as ionizing photons for a given column density of H^0 and C^+ is not a perfect tracer of H^0, which is why C II λ1334 also has a distribution of optical depths that is different from that of ionizing photons. We discuss these issues in Sects. 5.3.2 and 5.3.3.

5.3.1. Covering of ionizing radiation sources

To study the distribution of column densities in a given direction of observation we first compute the column density in front of every stellar particle. As we want to study the coverage of photons rather than the coverage of stars, we give a weight to each stellar particle equal to its luminosity at 910 Å. We then compute the fraction of photons behind a given amount of column density. A distribution perfectly matching the picket-fence model would show two distinct populations (for example, 30% of the ionizing photons facing less than $N_{HI} = 10^{15}\,cm^{-2}$ and 70% facing more than $N_{HI} = 10^{19}\,cm^{-2}$).

We show in Fig. 16 distributions of column densities from three of the seven directions of observation that we studied in Sect. 3. The pink histogram is typical of the majority of our directions, where the escape fraction is zero. All the photons are emitted behind optically thick gas, with $N_{HI} > 10^{20}\,cm^{-2}$. As there are no photons with $\tau_{910} \approx 1$, the covering fraction is well defined: it is equal to one and the escape fraction is zero. The orange histogram shows a clear bimodality, with a (double)-peak of the curve around $N_{HI} = 10^{15}$–$10^{16}\,cm^{-2}$, which is the "holes" part, and another peak at high column densities, which is the "cloud" part. There is only a tiny fraction of 910 Å photons that are facing gas with $\tau_{910} \approx 1$, and so this is also a case where the picket-fence model is a good description. However, the dark-blue histogram is an example of a situation where there is a significant fraction of photons facing gas such that $\tau_{910} \approx 1$. The photons are partially absorbed in this regime, which makes the notion of the covering fraction unclear. The cyan dashed line corresponds to the average of all 5184 directions, and it shows that the most common configurations have almost exclusively high column densities, and thus very low escape fractions.

Let us define the covering fraction as the fraction of 910 Å photons facing more than $N_{HI} = 10^{17.2}\,cm^{-2}$, which is the column density for which $\tau_{910} = 1$:

$$f_C^{HI} = \frac{\sum_{N_{HI}^{star} \geq 10^{17.2}\,cm^{-2}} L_{star}^{910}}{\sum_{\text{all stars}} L_{star}^{910}}, \quad (7)$$

where N_{HI}^{star} is the column density of neutral hydrogen in front of a stellar particle and L_{star}^{910} is the luminosity of the stellar particle at 910 Å. We note that f_C^{HI} depends on the direction of observation of the galaxy. In an idealized picket-fence geometry with empty holes and optically thick clouds, we have the relation $f_{esc} = 1 - f_C^{HI}$. In the general case, photons that are covered in the sense of Eq. (7) can still have a transmission factor of up to $e^{-1} = 37\%$. Indeed, if all the gas has exactly $N_{HI} = 10^{17.2}$, the photons are considered as covered, but the optical depth in this case is $\tau_{910} = 1$, and so 37% of photons can still escape. Similarly, photons that are not covered can still be absorbed; we can only say that their transmission factor is larger

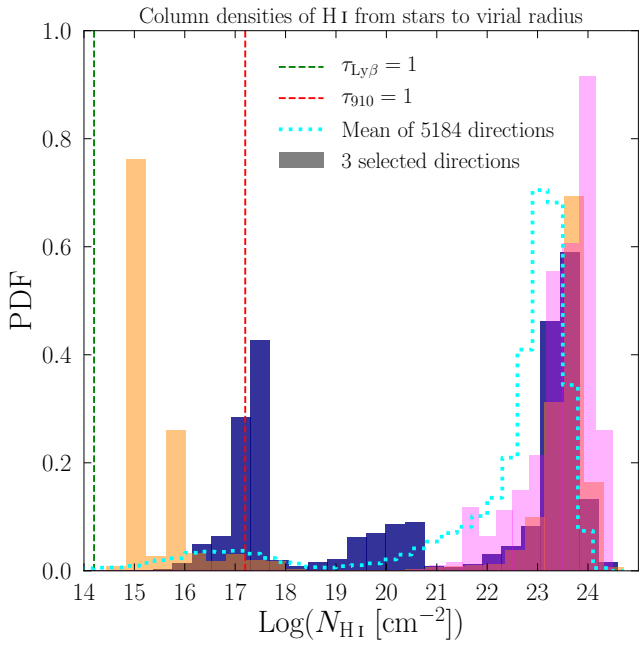

Fig. 16. Probability distribution of column densities faced by 910 Å photons. The three filled histograms correspond to the three directions that have the same color in Sect. 3. The cyan dashed line is the average distribution over the 1728 directions of the three outputs. It has a covering fraction of 93%. Molecular hydrogen chemistry is not considered in the simulation, and so the real atomic hydrogen column densities should saturate around 10^{23} cm^{-2}.

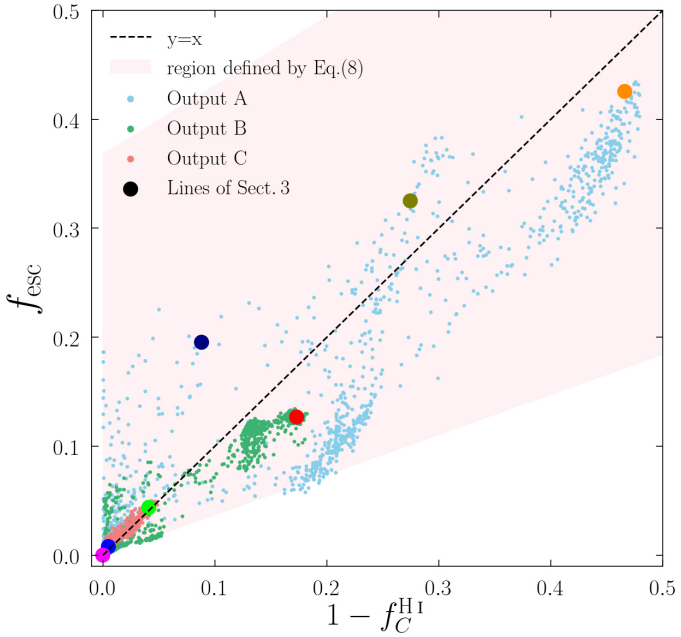

Fig. 17. Relation between the covering fraction of ionizing photons, defined by Eq. (7), and the escape fraction of ionizing photons for the 1728 directions of observation of the three outputs, A, B, and C. The seven colored dots are the directions that we study in Sect. 3. The pink area highlights the region of the figure that is allowed by our definition of covering fraction.

than 37%. The equation for the allowed region in the relation $f_{\rm esc} \leftrightarrow (1 - f_C^{\rm H\,I})$ is:

$$(1 - f_C^{\rm H\,I}) \cdot e^{-1} < f_{\rm esc} < f_C^{\rm H\,I} \cdot e^{-1} + (1 - f_C^{\rm H\,I}). \quad (8)$$

Figure 17 displays the relation between $f_{\rm esc}$ and $1 - f_C^{\rm H\,I}$ for our 5184 mock observations and shows that indeed not all the directions are perfectly described by the picket-fence model, in the sense that we do not always have $f_{\rm esc} = 1 - f_C^{\rm H\,I}$. We note that a few points are below the relation given by Eq. (8) because of absorption by helium and dust that was not considered in this formula. The directions of observations corresponding to the pink and the orange histograms in Fig. 16 are close to the one-to-one relation in Fig. 17, as in the picket-fence geometry. This was expected because those directions have almost no stars facing gas such that $\tau_{910} \approx 1$. However, the dark-blue dot sits further from the one-to-one relation, which is explained by the significant fraction of photons facing gas with $\tau_{910} \approx 1$ in the corresponding histogram of Fig. 16. We see that overall the directions are close to the one-to-one relation. Therefore, the picket-fence geometry is on average a reasonable model of the covering of ionizing photons in our galaxy.

Here we considered the description of the escape of ionizing photons in the picket-fence framework and compared to the simulation, but the goal is to use absorption lines to infer this escape fraction. This leads to additional limitations due to the different optical depths of Lyβ and C II λ1334 compared to LyC, and the difference in distribution of C^+ compared to H^0. We now explain those new limitations for the two lines.

5.3.2. The case of Lyβ

In order for the residual flux of Lyβ to be an indicator of the covering fraction, the "holes" in the picket-fence have to be optically thin for both LyC and Lyβ. However the column density at which $\tau_{Ly\beta} = 1$, with a turbulent velocity of 20 km s^{-1}, is $N_{\rm H\,I} = 10^{14.22}$ cm^{-2}, which is three orders of magnitude below the column density at which $\tau_{\rm LyC} = 1$. For example, the orange distribution of Fig. 16 represents a case where the picket-fence model works well to describe the escape of ionizing photons, in the sense that $f_{\rm esc} = 1 - f_C^{\rm H\,I}$, but Ly$\beta$ residual flux is not a good indicator of this $f_{\rm esc}$ because the part that is optically thin for LyC has $N_{\rm H\,I}$ of around 10^{15}–10^{16} cm^{-2}, which is already optically thick for Lyβ. This is confirmed in the right panel of Fig. 15, where the orange dot has a residual flux of only 7%, while the escape fraction is as high as 47%. The difficulty of the optical depth of Lyβ being larger than that of LyC cannot be circumvented. A saturated Lyβ line can very well correspond to a direction with a high escape fraction of ionizing photons. This is why Lyβ is a good lower limit of $f_{\rm esc}$, but not a direct tracer.

5.3.3. The case of C$_{\rm II}\lambda$1334

We now analyze the optical depth of C II λ1334 in comparison with the one of 910 Å photons. The column density of C^+ at which the optical depth of the center of C II λ1334 is one for a turbulent velocity of 20 km s^{-1} is $N_{\rm C\,II} = 10^{13.89}$ cm^{-2}. For 910 Å photons, it is $N_{\rm H\,I} = 10^{17.19}$ cm^{-2}, which means that if $N_{\rm C\,II}/N_{\rm H\,I} = 5 \times 10^{-4}$, the two optical depths are the same. This ratio is close to the ratio of carbon to hydrogen in the Sun, but metallicity and ionization effects can shift $N_{\rm C\,II}/N_{\rm H\,I}$ further from this value.

In Fig. 18, we show the ratio of carbon to hydrogen column densities in the upper panel and of C^+ over H^0 column densities in the lower panel for one direction of our simulated galaxy. We see that the C^+ over H^0 ratio varies by about eight orders of magnitude from region to region. However, the carbon-to-hydrogen

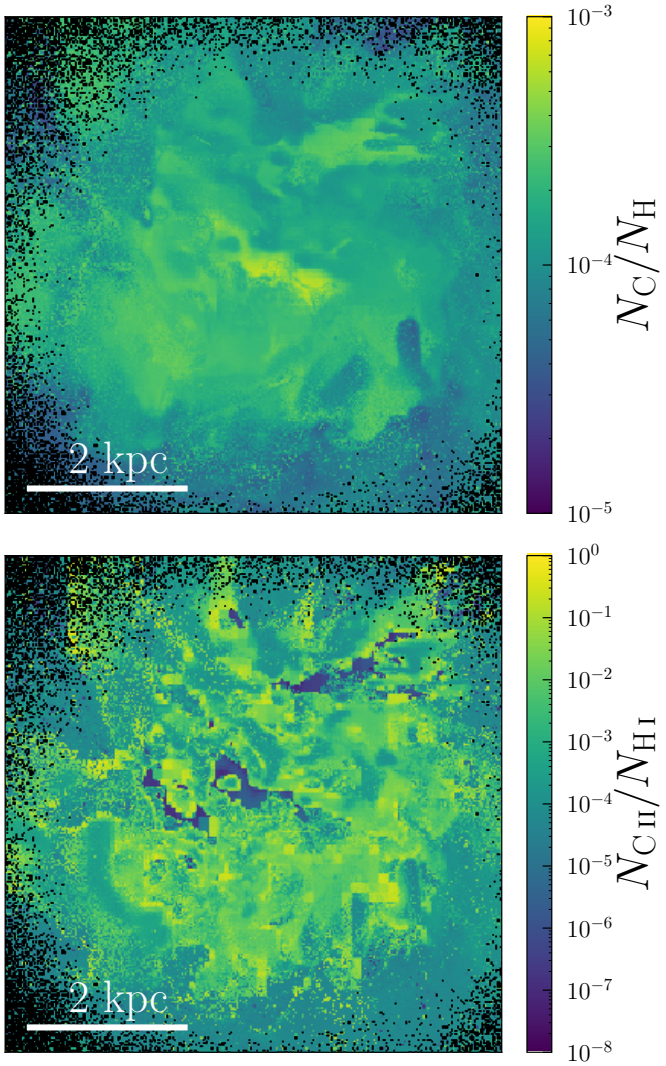

Fig. 18. Ratio of column densities in the same output and direction as in Figs. 1 and 14. *Upper panel*: ratio of carbon to hydrogen column densities. *Lower panel*: ratio of C^+ to H^0 column densities. The column densities are computed from the stellar particles to the virial radius. In every pixel, the ratio is averaged over all stellar particles in this pixel and weighted by the luminosity of the stellar particle at 1334 Å. The black pixels have no stellar particles in them.

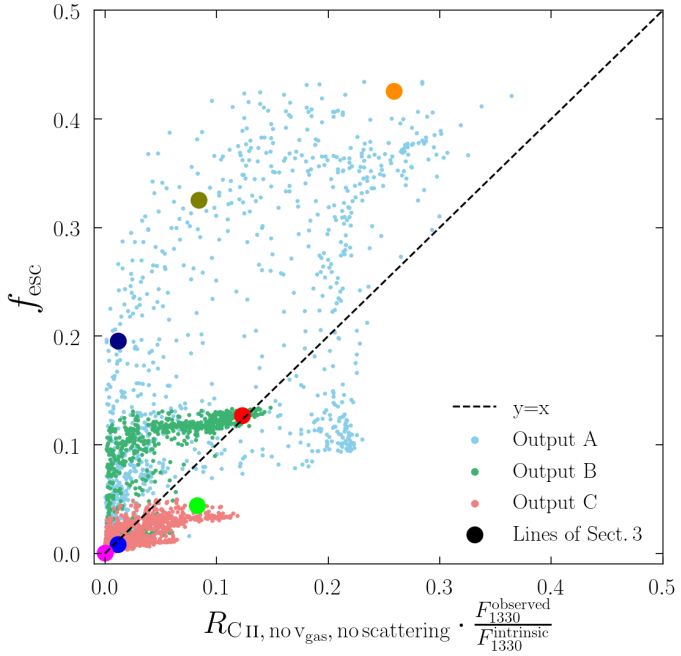

Fig. 19. As in the left panel of Fig. 15, but omitting scattering and setting gas velocity to zero.

ratio is more uniform, with an average of 1.5×10^{-4}, and varying by less than two orders of magnitude. Thus, the variation of the C^+-to-H^0 ratio comes primarily from ionization effects. There are regions abundant in H^0 but lacking C^+, because most carbon is in the neutral state. In other regions, C^+ is overabundant compared to the N_C/N_H ratio, which occurs when hydrogen is mostly ionized and carbon is mainly in the C^+ state, which is possible because the ionization energy of C^+ is higher than that of H^0.

For the continuum of stars behind screens with $N_{C\,II}/N_{H\,I} < 5 \times 10^{-4}$, the C II λ1334 absorption is weaker than the absorption of ionizing photons, which leads to the residual flux of C II λ1334 underestimating f_{esc}. Conversely, when stars are facing gas such that $N_{C\,II}/N_{H\,I} > 5 \times 10^{-4}$, the residual flux of C II λ1334 overestimates f_{esc}. Therefore, the variation of the value of $N_{C\,II}/N_{H\,I}$ in the galaxy causes inevitable dispersion in the relation between the residual flux of C II λ1334 and the escape fraction of ionizing photons. This dispersion changes with time and orientation, depending on the alignment between the brightest stars and the screens of gas with different $N_{C\,II}/N_{H\,I}$ ratios.

5.4. Processes affecting the residual flux

Lastly, we highlight the extent to which processes during the radiative transfer of the line affect its properties; in particular the residual flux. We show in Sect. 3.3 that the absorption lines are sensitive to the velocity distribution of the gas and sometimes to the effect of infilling, especially for directions with strong fluorescence. However, correcting for scattering and gas velocity dispersion does not improve the tightness of the relation between the residual flux of C II λ1334 and the escape fraction of ionizing photons. Figure 19 shows this relation with the scattering ignored and the gas velocity set to zero. We see that the relation has as much dispersion as in Fig. 15. Moreover, the correction of scattering and gas velocity brings the points further away from the one-to-one relation, leading to C II λ1334 being a lower limit of f_{esc}, similar to Lyβ. In other words, C II λ1334 globally saturates at column densities of gas that are too low for ionizing photons to be significantly absorbed, but the effect of infilling and the dispersion of gas velocity increase the residual flux of C II λ1334 so that it gets closer to the value of f_{esc}.

5.5. Another metallic line: Si II λ1260

We find that C II λ1334 traces the escape fractions of ionizing photons better than Lyβ does. We now show that all the results for C II λ1334 apply to Si II λ1260, another LIS absorption line. The modeling of the distribution of Si^+ is slightly less robust than that of C^+, first because silicon depletes more onto dust, which we do not model here, and second because of charge transfer reactions (Appendix A). However, the spectra are not very sensitive to a change in density, as shown in Sect. 2.2. First, Fig. 20 shows the seven directions of observations that we selected in this paper, with the spectra of Si II λ1260 on the left and that of C II λ1334 on the right. We see that they have similar properties:

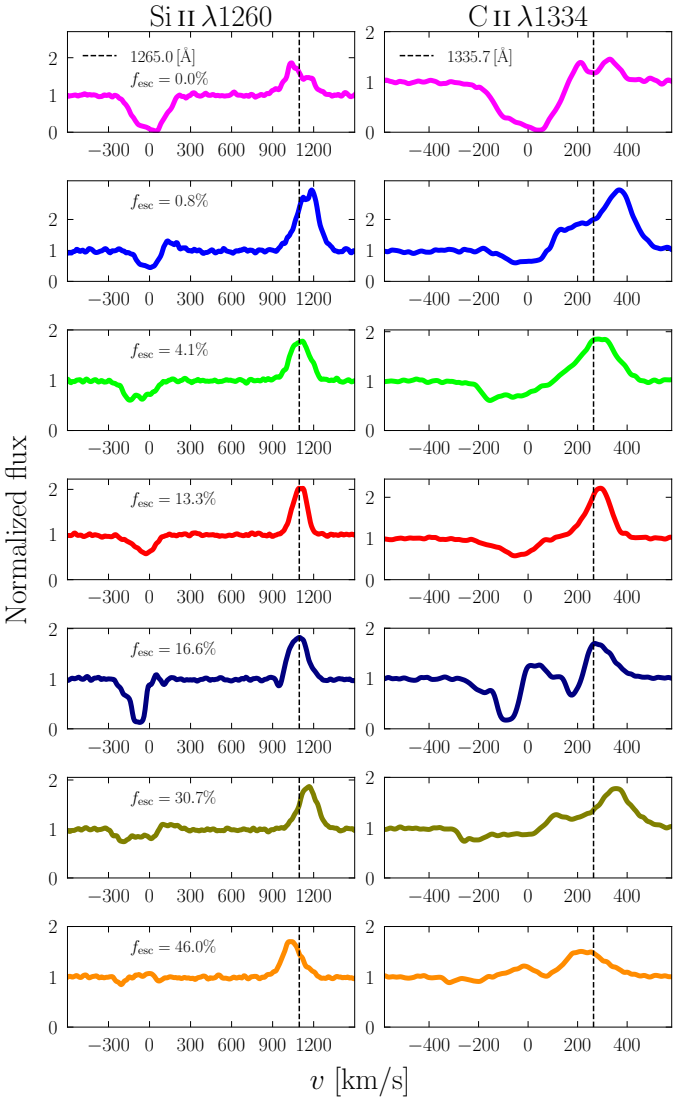

Fig. 20. Same as Fig. 3 but with Si II λ1260 instead of Lyβ. The vertical dashed lines highlight the wavelengths of the fluorescent channels.

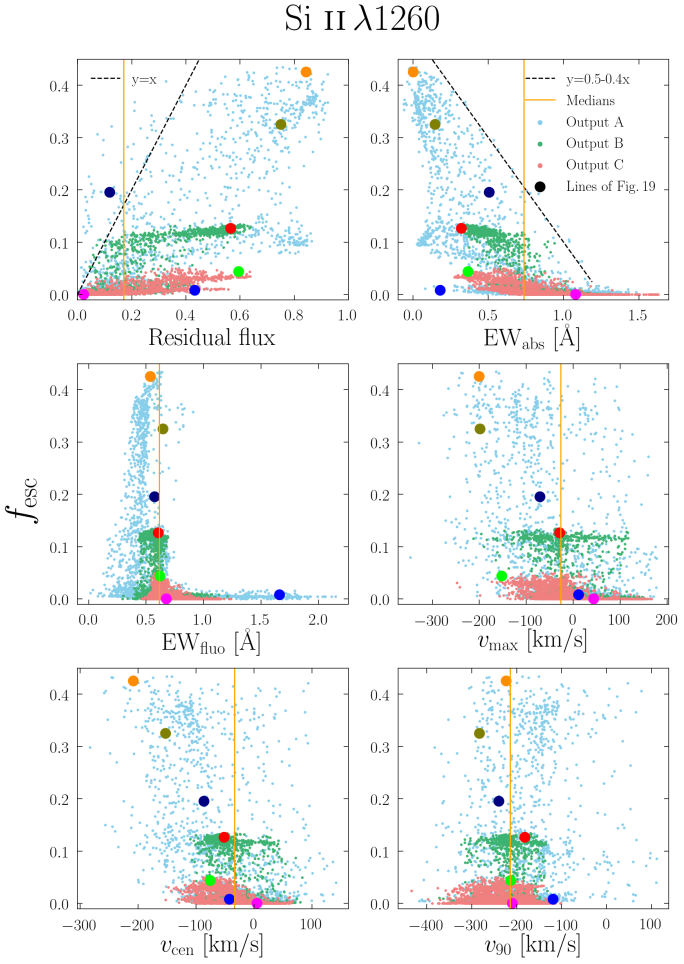

Fig. 21. Same as Fig. 13 but for Si II λ1260.

the residual flux and EWs are almost the same. This is confirmed in Fig. 21 where we plot the relation between the escape fractions and the different properties of the Si II λ1260 line. All the relations are similar to those seen in the C II plot in the left panel of Fig. 13. Finally, as Fig. 22 shows, the relations after dust correction are also closer to the one-to-one relation. Si II λ1260 is therefore as good a tracer of f_{esc} as C II λ1334, providing relatively good predictions when the (dust corrected) residual flux is large, but unfortunately predictions that often overestimate reality when the residual flux is smaller than 0.3.

6. Summary

In this paper we present a method to post-process radiative hydrodynamics simulations in order to build mock observations of LIS UV absorption lines. We applied this method to a cosmological zoom-in simulation of a galaxy at $z = 3$ with a stellar mass of $2.3 \times 10^9\,M_\odot$ in order to study the processes that affect the formation of absorption lines and to find correlations between the escape fraction of ionizing photons and the properties of the absorption lines. Our main results can be summarized as follows:

1. Our procedure reproduces metal absorption lines with a diversity of shape and strength (Fig. 3) that is similar to observations; such as for example those of Jaskot et al. (2019). The Lyβ absorption lines of our galaxy seem to differ from the ones in Steidel et al. (2018) or Gazagnes et al. (2020), probably because our simulated galaxy is less massive and has a lower SFR.
2. Many properties of the observed radiation depend highly on the direction of observation. There are variations of up to more than one magnitude at 1500 Å (Table 1). The LyC escape fractions vary from 0% to 47% (Fig. 11). The residual flux goes from 0% to 95% (Fig. 13).
3. The precise implementation of the radiation between 6 eV and 13.6 eV does not impact the mock absorption lines. There are dense regions where the carbon ionization fraction depends on this implementation, but they are very dusty and do not contribute to our spectra. However, using a UV background at all energies instead of the ionizing radiation from the simulation fails to reproduce the same observed spectra (Fig. 4). This shows that radiative transfer of ionizing photons in the simulation is crucial for this kind of study.
4. The fluorescent lines C II* λ1335 and Si II* λ1265 are clearly visible (Fig. 20), with an EW of the same order as for the absorption. The shape and size of the fluorescent lines is sensitive to the modeling of the fine-level structure of the

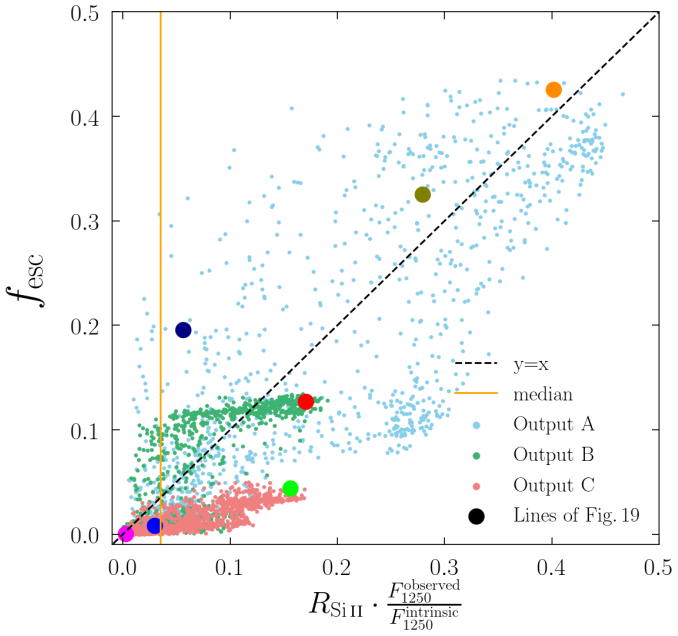

Fig. 22. Same as Fig. 15 but for Si II λ1260.

ground state of their respective ions (Fig. 6). The presence of absorption by ions in the fine-level structure can lead to the formation of a P-Cygni profile in the fluorescence.

5. The infilling effect increases the residual fluxes by around 0.06 on average. There are however directions where the effect is much stronger, with a change of up to more than 0.3 (Fig. 7). A fluorescent line with a large EW is an indicator that the infilling effect is stronger than on average (Fig. 8). Infilling depends very little on the aperture size of the mock observation, as long as it encompasses more than ≈90% of the stellar continuum (Figs. 6 and 9).
6. The large-scale distribution of the gas velocity spreads the absorption over a larger wavelength range than if all the gas had the same velocity (Figs. 9 and 10). This spread leads to an increase in the residual flux of the lines (Fig. 7), which is one factor that challenges the use of absorption lines as tracers of the escape fraction of ionizing photons.
7. The properties of the absorption lines without dust correction correlate poorly with the LyC escape fraction (Figs. 13 and 21). The residual fluxes of C II λ1334 or Si II λ1260 cannot predict the escape fraction, but are good upper limits. A large EW of absorption, either of a metallic line or of Lyβ, or a large EW of fluorescence, indicate an escape fraction smaller than 2%.
8. Multiplying the residual flux of C II λ1334 or Si II λ1260 by the extinction of the continuum by dust makes it a better indicator of escape fraction, following globally the one-to-one relation, although with strong dispersion (Figs. 15 and 22). This dust correction is similar to the ones made in the literature (e.g., Steidel et al. 2018; Gazagnes et al. 2018; Chisholm et al. 2018). For directions with dust-corrected residual flux between ∼2% and 30%, the escape fraction is often overestimated. Concerning Lyβ, multiplying its residual flux by the extinction by dust makes it a good lower limit of the escape fraction, but not a direct proxy (Fig. 15).
9. The covering fraction, defined as the fraction of photons that are facing $N_{HI} \geq 10^{17.2}$ cm^{-2}, gives a reasonable estimate for the escape fraction of ionizing radiation for our simulated galaxies, although the scatter in the relation is not negligible (Fig. 17).
10. Many geometrical and physical complexities make it difficult to use absorption lines as reliable tracers of the escape fraction. The different distributions of sources for ionizing photons and nonionizing UV continuum (Fig. 14), the complex distribution of column densities in front of stars (Fig. 16), the different optical depth between the ionizing photons and the lines, the tangled distribution of C^+ with respect to H^0 (Fig. 18), and the effects of infilling and gas velocity distribution are not correctable and explain the dispersion in the scatter plots, and the fact that dust-corrected residual fluxes below ∼30% overestimate the escape fraction.

In summary, metallic absorption lines can predict the escape fraction of ionizing photons accurately only in two extreme regimes: when $f_{esc} = 0$, with saturated absorption lines, or when the escape fraction is very high, with a dust corrected residual flux larger than 30%. As a result, even if it may be difficult to estimate the escape fraction of ionizing radiation from a galaxy, metallic absorption lines have the potential to rule out LyC emission from line-saturated galaxies, and to select promising candidates with high residual fluxes.

We are happy to share the mock observations of our simulated galaxy and encourage those interested to contact us.

Acknowledgements. The authors are grateful to the anonymous referee for her/his very careful and constructive report. We also thank J. Kerutt for her detailed revision of the manuscript. VM, AV, and TG are supported by the ERC Starting Grant 757258 "TRIPLE". TK was supported in part by the National Research Foundation of Korea (NRF-2019K2A9A1A0609137711 and NRF-2020R1C1C100707911) and in part by the Yonsei University Future-leading Research Initiative (RMS2-2019-22-0216). This work was supported by the Programme National Cosmology et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES. The simulation could be run thanks to our access to computing resources at the Common Computing Facility (CCF) of the LABEX Lyon Institute of Origins (ANR-10-LABX-0066).

References

Agertz, O., Romeo, A. B., & Grisdale, K. 2015, MNRAS, 449, 2156
Alexandroff, R. M., Heckman, T. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 810, 104
Atek, H., Richard, J., Jauzac, M., et al. 2015, ApJ, 814, 69
Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, Nature, 553, 473
Badnell, N. R. 2006, ApJS, 167, 334
Barrow, K. S. S., Wise, J. H., Norman, M. L., O'Shea, B. W., & Xu, H. 2017, MNRAS, 469, 4863
Barrow, K. S. S., Wise, J. H., Aykutalp, A., et al. 2018, MNRAS, 474, 2617
Bassett, R., Ryan-Weber, E. V., Cooke, J., et al. 2019, MNRAS, 483, 5223
Bergvall, N., Zackrisson, E., Andersson, B. G., et al. 2006, A&A, 448, 513
Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. 2014, Science, 346, 216
Bosman, S. E. I., Fan, X., Jiang, L., et al. 2018, MNRAS, 479, 1055
Carr, C., Scarlata, C., Panagia, N., & Henry, A. 2018, ApJ, 860, 143
Cen, R., & Kimm, T. 2015, ApJ, 801, L25
Chisholm, J., Orlitová, I., Schaerer, D., et al. 2017, A&A, 605, A67
Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018, A&A, 616, A30
Corlies, L., Peeples, M. S., Tumlinson, J., et al. 2020, ApJ, 896, 125
De Cia, A. 2018, A&A, 613, L2
Dere, K. P. 2007, A&A, 466, 771
Dessauges-Zavadsky, M., D'Odorico, S., Schaerer, D., et al. 2010, A&A, 510, A26
Dijkstra, M. 2014, PASA, 31, e040
Dijkstra, M. 2017, ArXiv e-prints [arXiv:1704.03416]
Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, ApJ, 828, 71
Eldridge, J. J., Izzard, R. G., & Tout, C. A. 2008, MNRAS, 384, 1109
Erb, D. K. 2015, Nature, 523, 169
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Feltre, A., Bacon, R., Tresse, L., et al. 2018, A&A, 617, A62

Feltre, A., Maseda, M. V., Bacon, R., et al. 2020, A&A, 641, A118
Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, Rev. Mex. Astron. Astrofis., 53, 385
Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, ApJ, 879, 36
Finley, H., Bouché, N., Contini, T., et al. 2017, A&A, 608, A7
Gazagnes, S., Chisholm, J., Schaerer, D., et al. 2018, A&A, 616, A29
Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, A&A, 639, A85
Giallongo, E., Grazian, A., Fiore, F., et al. 2019, ApJ, 884, 19
Gnedin, N. Y. 2016, ApJ, 825, L17
Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, ApJ, 672, 765
Grassi, T., Bovino, S., Schleicher, D. R. G., et al. 2014, MNRAS, 439, 2386
Grevesse, N., Asplund, M., Sauval, A. J., & Scott, P. 2010, Ap&SS, 328, 179
Grimes, J. P., Heckman, T., Aloisi, A., et al. 2009, ApJS, 181, 272
Grissom, R. L., Ballantyne, D. R., & Wise, J. H. 2014, A&A, 561, A90
Gronke, M., Ocvirk, P., Mason, C., et al. 2020, ArXiv e-prints [arXiv:2004.14496]
Guillet, T., & Teyssier, R. 2011, J. Comput. Phys., 230, 4756
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Habing, H. J. 1968, Bull. Astron. Inst. Neth., 19, 421
Hahn, O., & Abel, T. 2013, Astrophysics Source Code Library [record ascl:1311.011]
Heckman, T. M., Sembach, K. R., Meurer, G. R., et al. 2001, ApJ, 558, 56
Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, ApJ, 730, 5
Heckman, T. M., Alexandroff, R. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 809, 147
Henry, A., Berg, D. A., Scarlata, C., Verhamme, A., & Erb, D. 2018, ApJ, 855, 96
Henyey, L. G., & Greenstein, J. L. 1941, ApJ, 93, 70
Hummels, C. B., Smith, B. D., & Silvia, D. W. 2017, ApJ, 847, 59
Iliev, I. T., Mellema, G., Ahn, K., et al. 2014, MNRAS, 439, 725
Inoue, A. K., Hasegawa, K., Ishiyama, T., et al. 2018, PASJ, 70, 55
Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016a, Nature, 529, 178
Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, MNRAS, 461, 3683
Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018, MNRAS, 478, 4851
Jaskot, A. E., & Oey, M. S. 2013, ApJ, 766, 91
Jaskot, A. E., & Oey, M. S. 2014, ApJ, 791, L19
Jaskot, A. E., Dowd, T., Oey, M. S., Scarlata, C., & McKinney, J. 2019, ApJ, 885, 96
Jenkins, E. B. 2009, ApJ, 700, 1299
Jones, T., Stark, D. P., & Ellis, R. S. 2012, ApJ, 751, 51
Jones, T. A., Ellis, R. S., Schenker, M. A., & Stark, D. P. 2013, ApJ, 779, 52
Katz, H., Galligan, T. P., Kimm, T., et al. 2019, MNRAS, 487, 5902
Katz, H., Ďurovčíková, D., Kimm, T., et al. 2020, MNRAS, 498, 164
Kimm, T., Slyz, A., Devriendt, J., & Pichon, C. 2011, MNRAS, 413, L51
Kimm, T., Katz, H., Haehnelt, M., et al. 2017, MNRAS, 466, 4826
Kimm, T., Blaizot, J., Garel, T., et al. 2019, MNRAS, 486, 2215
Kretschmer, M., & Teyssier, R. 2020, MNRAS, 492, 1385
Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, MNRAS, 485, L24
Laursen, P., Sommer-Larsen, J., & Andersen, A. C. 2009, ApJ, 704, 1640
Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, A&A, 553, A106
Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, ApJ, 823, 64
Levermore, C. D. 1984, J. Quant. Spectrosc. Radiat. Transf., 31, 149
Li, A., & Draine, B. T. 2001, ApJ, 554, 778
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Luridiana, V., Morisset, C., & Shaw, R. A. 2015, A&A, 573, A42
Ma, X., Quataert, E., Wetzel, A., et al. 2020, MNRAS, 498, 2001
Madau, P., & Haardt, F. 2015, ApJ, 813, L8
Mason, C. A., Treu, T., de Barros, S., et al. 2018, ApJ, 857, L11
Michel-Dansac, L., Blaizot, J., Garel, T., et al. 2020, A&A, 635, A154
Mitchell, P. D., Blaizot, J., Cadiou, C., et al. 2021, MNRAS, 501, 5757
Nakajima, K., & Ouchi, M. 2014, MNRAS, 442, 900
Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, MNRAS, 463, 1462
Ouchi, M., Harikane, Y., Shibuya, T., et al. 2018, PASJ, 70, S13
Pallottini, A., Ferrara, A., Decataldo, D., et al. 2019, MNRAS, 487, 1689
Parsa, S., Dunlop, J. S., & McLure, R. J. 2018, MNRAS, 474, 2904
Patrício, V., Richard, J., Verhamme, A., et al. 2016, MNRAS, 456, 4191
Peeples, M. S., Corlies, L., Tumlinson, J., et al. 2019, ApJ, 873, 129
Prochaska, J. X., Kasen, D., & Rubin, K. 2011, ApJ, 734, 24
Puschnig, J., Hayes, M., Östlin, G., et al. 2017, MNRAS, 469, 3252
Reddy, N. A., Steidel, C. C., Pettini, M., Bogosavljević, M., & Shapley, A. E. 2016, ApJ, 828, 108
Rivera-Thorsen, T. E., Hayes, M., Östlin, G., et al. 2015, ApJ, 805, 14
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJ, 802, L19
Rosdahl, J., & Teyssier, R. 2015, MNRAS, 449, 4380
Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., & Teyssier, R. 2013, MNRAS, 436, 2188
Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, MNRAS, 479, 994
Scarlata, C., & Panagia, N. 2015, ApJ, 801, 43
Schroeder, J., Mesinger, A., & Haiman, Z. 2013, MNRAS, 428, 3058
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Smail, I., Swinbank, A. M., Richard, J., et al. 2007, ApJ, 654, L33
Smith, A., Safranek-Shrader, C., Bromm, V., & Milosavljević, M. 2015, MNRAS, 449, 4336
Sobacchi, E., & Mesinger, A. 2015, MNRAS, 453, 1843
Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, MNRAS, 456, 485
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, ApJ, 869, 123
Teyssier, R. 2002, A&A, 385, 337
Toro, E. F., Spruce, M., & Speares, W. 1994, Shock Waves, 4, 25
Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, MNRAS, 470, 224
Trebitsch, M., Dubois, Y., Volonteri, M., et al. 2020, A&A, submitted [arXiv:2002.04045]
Vasei, K., Siana, B., Shapley, A. E., et al. 2016, ApJ, 831, 38
Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, A&A, 578, A7
Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, A&A, 597, A13
Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, ApJ, 465, 487
Voronov, G. S. 1997, At. Data Nucl. Data Tables, 65, 1
Yoo, T., Kimm, T., & Rosdahl, J. 2020, MNRAS, 499, 5175

Appendix A: Comparison of ionization fractions between KROME and Cloudy

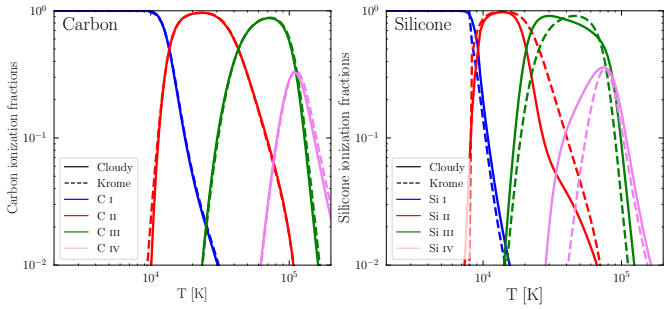

Fig. A.1. Ionization fractions of carbon and silicon as a function of temperature. KROME reactions are recombinations from Badnell (2006) and collisional ionization from Voronov (1997).

In Fig. A.1 we show a comparison of the ionization fractions of carbon and silicon as a function of temperature. We include here only recombinations and collisional ionizations, and no photo-reactions. The chemical network of KROME is chosen to match the results of Cloudy. We see that the results for carbon are coherent between the two codes. For silicon, the effects of charge transfer, which is included in Cloudy but not in KROME, are more important, and we do not recover the curves of Cloudy as well as with carbon. We still choose to omit those reactions in order to preserve the state of hydrogen and helium, as explained in Sect. 2.2.

Appendix B: Computing photoionization rates

In KROME, the usual method to include photo-reactions is in three steps:

- Specify the photo-reaction in the chemical network, for example $C^+ + \gamma \longrightarrow C^{++} + e$.
- Provide a file with a cross-section as a function of photon energy for this reaction.
- Before calling KROME in a cell of a simulation (on-the-fly or in post-processing), specify a discretized spectral energy distribution.

KROME then computes the rate of the photo-reaction based on the cross-section and the radiation field, and uses it to compute the chemical evolution. However, the photo-ionization rates of metallic ions can be easily computed from the RAMSES-RT output (along with the modeling of the radiation below 13.6 eV, see Sect. 2.2) without using KROME, and so we bypass the method above, giving the more simple procedure:

- Specify the photo-reaction in the chemical network.
- For every cell, compute the photoionization rate with the method below.
- Give this photoionization rate to KROME via a routine that we added in the KROME code.

The photoionization rates are products of cross-sections and fluxes of photons integrated over frequency. The flux is given in every cell directly by the simulation. For cross-sections, we have to apply the same method as RAMSES-RT, where the photoionization cross-sections are weighted by the SEDs of all stellar particles in the simulation.

To begin with, we have a stellar SED library, BPASS in our case, with an SED for a grid of stellar ages and metallicities. For all SEDs, at each stellar age and metallicity in the grid, we compute the mean cross-section of photoionization for each bin of radiation of RAMSES-RT, weighted by the number of photons, with this formula:

$$\bar{\sigma}^X_{i,\mathrm{age},Z} = \frac{\int_{\mathrm{bin}\,i} \frac{J_\nu^{\mathrm{age},Z}}{h\nu}\,\sigma_\nu^X\,\mathrm{d}\nu}{\int_{\mathrm{bin}\,i} \frac{J_\nu^{\mathrm{age},Z}}{h\nu}\,\mathrm{d}\nu}, \tag{B.1}$$

where bin i is the ith bin of radiation of the simulation, X indicates an ion for which we compute the rate, $J_\nu^{\mathrm{age},Z}$ is the luminosity as a function of frequency for a given stellar age and metallicity in BPASS, in units of erg s^{-1} Hz^{-1}, $h\nu$ is the energy of a photon at frequency ν, and σ_ν^X is the photoionization cross-section as a function of frequency for species X, given by Verner et al. (1996).

From those mean cross-sections we compute the cross-sections of a given stellar particle, $\bar{\sigma}^X_{i,j_{\mathrm{star}}}$, by interpolating on age and metallicity. In a given domain of the simulation, where we want to compute the photoionization rates in every cell, we average those cross-sections on all stellar particles, weighing by their luminosities:

$$\bar{\sigma}_i^X = \frac{\sum_{j_{\mathrm{star}}=1}^{N_{\mathrm{stars}}} \bar{\sigma}^X_{i,j_{\mathrm{star}}} \cdot L_i^{j_{\mathrm{star}}}}{\sum_{j_{\mathrm{star}}=1}^{N_{\mathrm{stars}}} L_i^{j_{\mathrm{star}}}}, \tag{B.2}$$

where $L_i^{j_{\mathrm{star}}}$ is the luminosity of the stellar particle j_{star} in the energy bin i. Finally, the photoionization rate in a given cell is:

$$\Gamma_X^{\mathrm{cell}}[\mathrm{s}^{-1}] = \sum_{i=1}^{N_{\mathrm{bins}}} \bar{\sigma}_i^X \cdot F_i^{\mathrm{cell}}, \tag{B.3}$$

where F_i^{cell} is the flux of photons in the cell, in units of s^{-1} cm^{-2}, for the ith bin of radiation.

4.2 On the covering fraction of neutral gas and the residual flux of LIS absorption lines

In Section 4.1 I have focused on the links between the residual flux of LIS absorption lines and the escape fraction of ionising photons. I also explained the notion of covering fraction of neutral gas, and how well it applies to realistic environments such as our simulated galaxy. Figure 17 shows that, contrary to the prediction of the picket-fence model, there is no one-to-one correspondence between the escape fraction of ionising photons and (one minus) the covering fraction of neutral hydrogen. This is because of the presence of gas with optical depth around 1, which is assumed to be absent in the picket-fence model.

Here, I would like to come back on the covering fraction, but directly in relation with the residual flux of the LIS lines, which I did not explicitly do in Section 4.1. It is often stated (e.g. Gazagnes et al., 2018) that the residual flux of LIS lines (either dust corrected or not, depending on the specific picket-fence that is adopted) is equal to the covering fraction of neutral hydrogen. To test this, I show the comparison between residual fluxes of Lyβ and C II λ1334, with and without dust correction, in Figure 4.1. We see that there is no correlation, except a scattered correlation in the lower right panel, when using the dust corrected residual flux of C II λ1334. The explanation is the same as for the relations between the residual flux and the escape fraction of ionising photons. The fact that some stars are covered by gas with an optical depth around one, that the residual flux of absorption lines are affected by infilling and large scale velocity gradients and that C^+ is not tracing H^0 perfectly, results in the scattered relations of Figure 4.1.

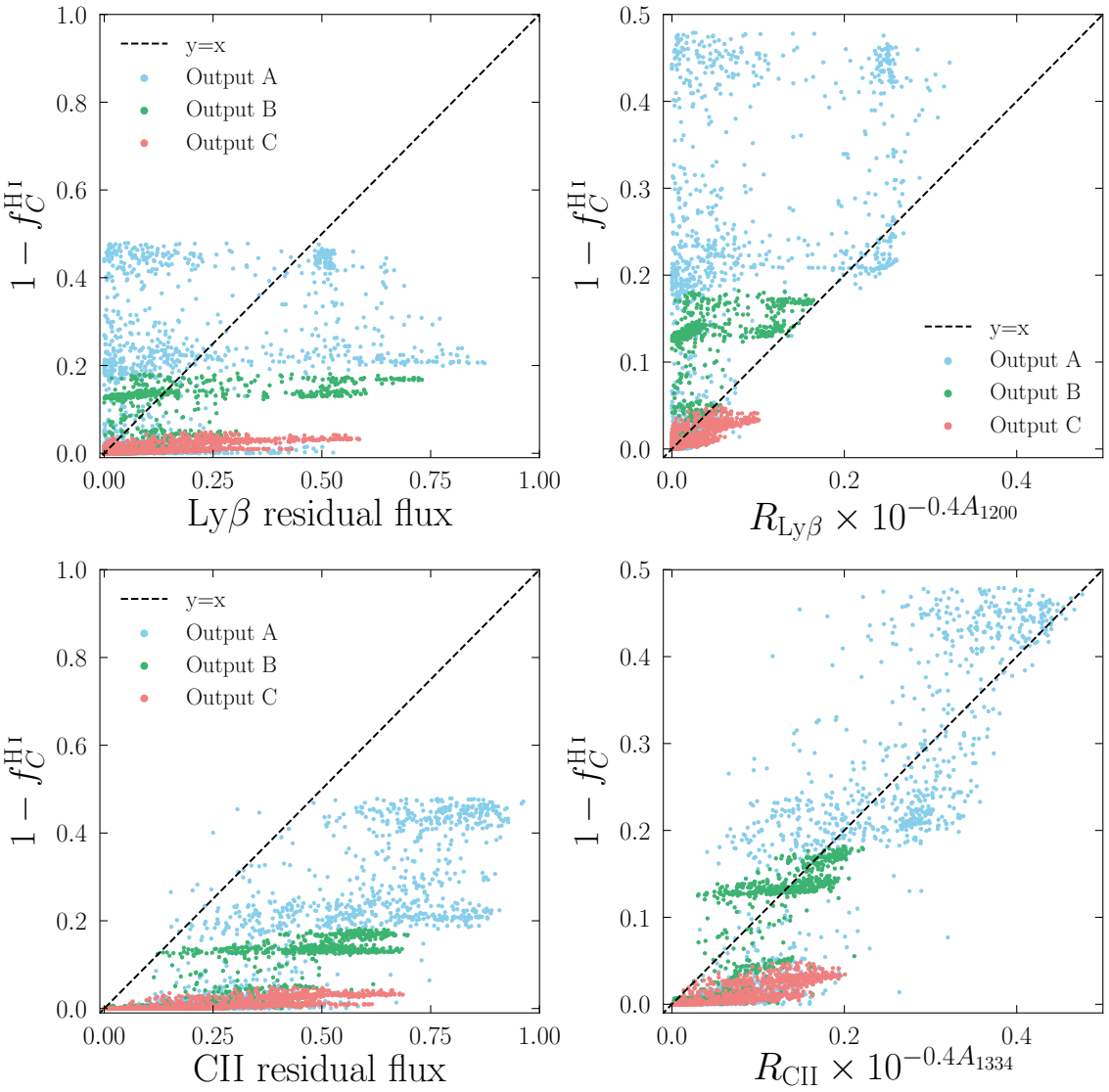

Figure 4.1: Comparison between residual flux of LIS lines and the covering fraction of neutral hydrogen. In each panel, the y-axis is one minus the covering fraction, as defined in Equation 7 of Section 4.1. Upper left panel: the x-axis is the residual flux of the Lyβ line. Upper right panel: the x-axis is the dust corrected residual flux of the Lyβ line, as defined in Section 5.2 of the paper. Lower left panel: the x-axis is the residual flux of the C II λ1334 line. Lower right panel: the x-axis is the dust corrected residual flux of C II λ1334, as defined in Section 5.2. For each panel, there are mock observations for three snapshots and 1728 directions of observations. The blue points are from output A, the green points from output B and the red points from output C. The black dashed lines indicate the one-to-one relation.

4.3 Update of the relation between $f_{\rm esc}$ and the residual flux of LIS absorption lines

One of the main results of Mauerhofer et al. (2021) is that after a dust correction of the continuum next to LIS absorption lines, low values of residual flux robustly indicate a high escape fraction of ionising photons. I did this same study on the massive halos of Sphinx20, and show the result in Figure 4.2, plotted with a 2d histogram instead of a scatter plot. When looking at the same range of values as Figure 22 of Chapter 4, there is no more correlation. Additionally, It does not hold anymore that a saturated absorption line indicates the absence of leakage of ionising photons. There are many points with a dust corrected residual flux below 0.05 and an escape fraction above 20%. Those points are coming from two galaxies, and more work has to be done to determine exactly which of their properties make them show this behaviour. However, it was already mentioned in Chapter 4 that large scale gas velocity gradients, scattering and dust affect LIS lines significantly, and that the exact ionisation balance between H^0 and LIS ions affect the relation between the residual flux of the lines and the escape fraction of ionising photons. Studying the galaxies of Sphinx20 confirms that the correspondence between the residual flux and $f_{\rm esc}$ is not universal. In general absorption lines are not tracing the escape fraction of ionising photons.

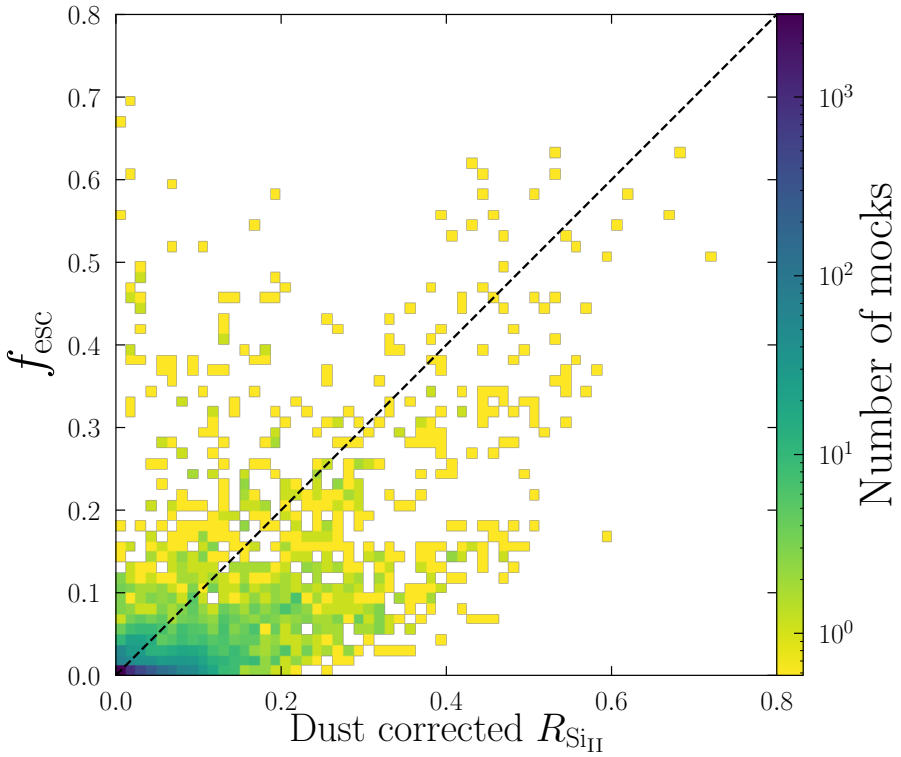

Figure 4.2: Relation between the escape fraction of ionising photons and the dust corrected residual flux of Si II λ1260 in the 100 most massive halos of Sphinx20. Each halo is observed from 110 orientations.

5 Dust attenuation in RAMSES-RT simulations

Now I move away from the question of the escape fraction of ionising photons to focus more on the role of dust in spectroscopy of high redshift galaxies, and whether our simulations and dust modelling reproduce observations faithfully. I begin by presenting and illustrating our dust models in more detail, in Section 5.1. Then, I study the realism of the simulations in Sections 5.2 and 5.3. In particular, I compare the UV properties such as the luminosity function, the β slope and the infrared excess with recent observational studies. Finally, in Section 5.4, I present observational ways of inferring the dust attenuation of high-redshift galaxies, and compare the resulting values with the attenuation obtained directly from the simulation and RASCAS post-processing.

5.1 Dust modelling

As a reminder of our dust model presented in Chapter 3, the equation for the optical depth of dust in cells of our simulations along a path dr is:

$$\mathrm{d}\tau_\lambda^{\mathrm{dust}} = \frac{Z}{0.005}(\mathrm{n_{H\,I}} + 0.01\mathrm{n_{H\,II}})\sigma_{\mathrm{dust}}(\lambda)\mathrm{d}r, \tag{5.1}$$

where Z is the metallicity of a cell and $\sigma_{\mathrm{dust}}(\lambda)$ is the cross-section of dust as a function of wavelength, which follows the SMC dust attenuation law and is represented in Figure 3.11. This model is taken from Laursen et al. (2009) and Gnedin et al. (2008), and is calibrated to reproduce realistic observations, in particular of Lyman alpha. In Section 5.2 I verify whether we reproduce dust related galaxy properties faithfully.

Figure 5.1, made from the last snapshot of the zoom simulation, at $z = 3.0$, shows images made to visualise the distribution of dust in the simulation, relative to sites of strong UV emission. The upper left panel shows the intrinsic FUV emission, at 1500 Å, with a range of three orders of magnitude in surface brightness. We see that the central region is very bright, due to the presence of many young stars. There are also bright spots scattered in a roughly 2 kpc radius around the center, indicating the presence of star formation. Interestingly, the right panel, representing the dust extinction in the same area, shows that the extinction is often collocated with the bright regions. This is a direct consequence of Equation 5.1, in which the opacity of dust scales with the density of neutral gas, which peaks where stars form. This is in agreement with Charlot & Fall (2000), where the authors develop a model of dust extinction with a stronger attenuation for young stars.

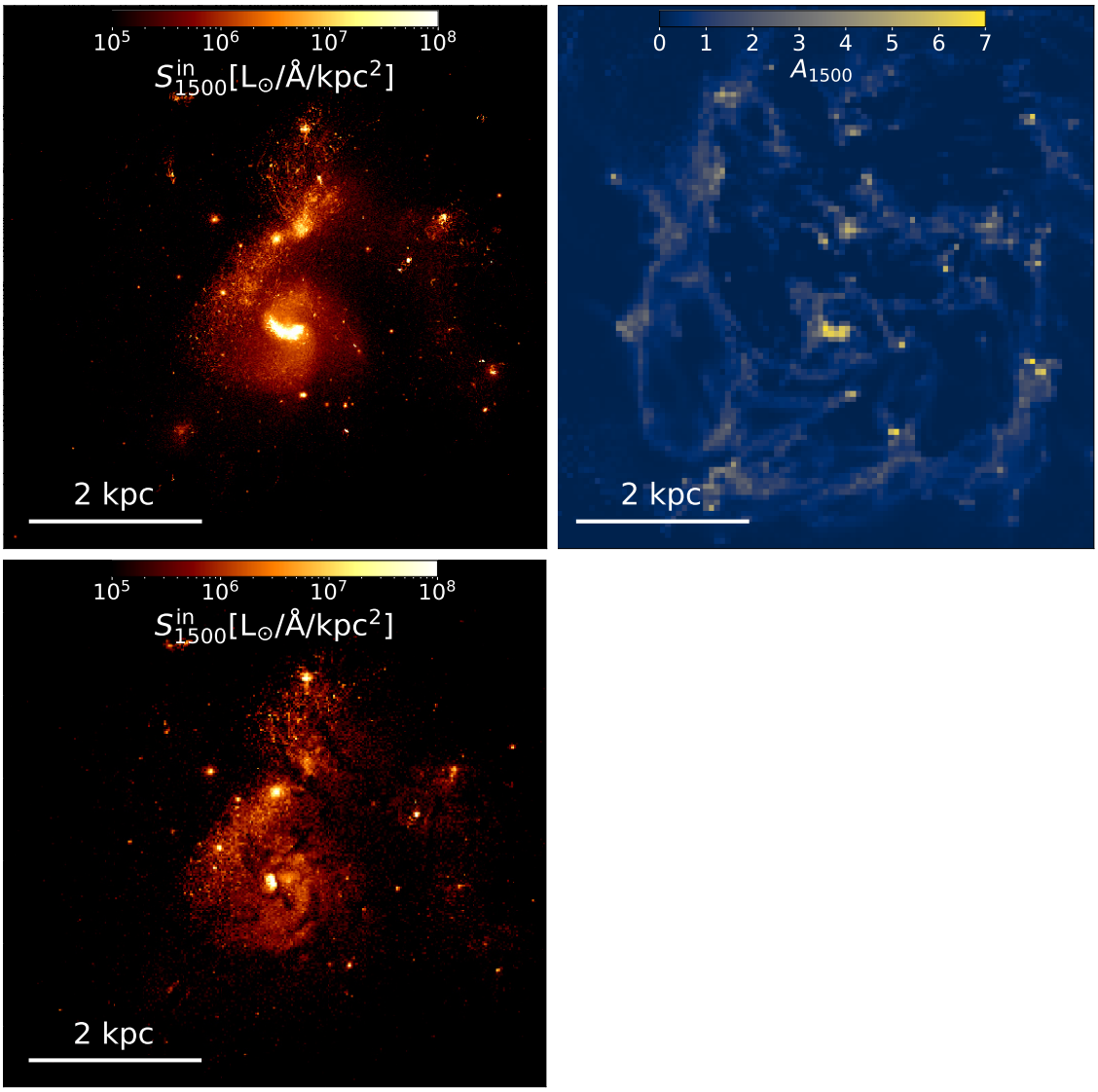

Figure 5.1: Upper left panel: Intrinsic surface brightness at 1500 Å of the last snapshot of the zoom simulation. Right panel: Resolved dust attenuation at 1500 Å, in magnitudes. For each pixel, I select all the stellar particles whose projected positions fall in it, and compute x: the ratio of unabsorbed 1500 Å over the intrinsic production. A_{1500} is equal to $-2.5\log(x)$. Lower left panel: Observed surface brightness at 1500 Å, including scattering on dust, obtained with the peeling-off algorithm of Rascas.

Combining the intrinsic emission and the dust absorption, I create a mock image of the galaxy with Rascas, shown in the lower left panel of Figure 5.1. We see that the bright central region is obscured, and that many of the small bright points have disappeared. There are dust lanes creating dark stripes, following the denser parts of the ISM. The UV light from this snapshot and direction of observation is intensely extinguished, and only 17% of the light comes out, which amounts to a global increase of the FUV magnitude of $A_{1500} = 1.88$.

There are other ways to model dust absorption in simulation. For example, Rascas also allows the user to use the LMC dust attenuation law instead of the SMC, which results in even stronger dust absorption. Also, studies of simulations that do not include transfer of ionising photons use different models than here, because they cannot easily know the density of neutral hydrogen. For example, Ma et al. (2020) assume a constant dust-to-metal ratio in the simulation, with the additional hypothesis that the dust is removed where the temperature is higher than 10^6 K, to take

into account the destruction of dust grains by sputtering (Barlow, 1978). Additionally, a new dust evolution model was implemented in Trebitsch et al. (2020), where they use Ramses-RT with an additional scalar variable that is not in our simulations, the dust-to-metal ratio. This allows them to model the destruction of dust near supernovae or the growth of dust grains in the cold ISM. Those two last dust modelling approaches could yield different results than ours, which I will test in future work.

5.2 Can our dust model reproduce observations?

To understand the level realism of my mock observations, in particular of the FUV continuum, I compare typical properties related to dust attenuation between our simulations and recent observations at high redshift. To do so, I use mock UV spectra produced with the code Rascas. This is done following the descriptions of Chapter 3, using a wide spectral range for the photon packets, between around 920 and 7000 Å, and with a spectral resolution of around 20 Å. I omit spectral lines for these mock spectra, so the only component interacting with photons here is dust. Both absorption and scattering of photons on dust grains are included. Some definitions of properties that I compare are:

- The luminosity function $\phi(M)$ [mag^{-1} Mpc^{-3}] is the measure of the distribution of galaxies as a function of magnitude. The integral $\int_{M_1}^{M_2} \phi(M)\mathrm{d}M$ gives the number of galaxies per unit volume that have magnitudes between M_1 and M_2. The (observed) magnitudes from the simulations are obtained directly from the mock FUV continuum spectra, converting luminosities to magnitude.
- The slope of the FUV continuum of a galaxy is called β. This quantity depends both on the intrinsic SED of stars in the galaxy and on the dust attenuation properties. I compute it by fitting a power-law to the continuum of my mock spectra between 1300Å and 2500Å.
- The infrared excess (IRX) is a measure of the intensity of infrared emission of dust compared to the observed UV luminosity. In principle, it is simply the ratio of the total infrared light re-emitted by dust, which by energy conservation is equal to the total absorbed UV luminosity, over the total observed UV light, IRX = $L_{\mathrm{IR}}/L_{\mathrm{UV}}$. However, since it is challenging to observe such large wavelength ranges, the definition of IRX in observational studies is in practice slightly different. L_{UV} is approximated by $1600[\text{Å}] \times L_{1600}[\mathrm{erg/s/Å}]$ (Meurer et al., 1999). For the infrared part, the available data is often observations from ALMA for one or a couple of wavelength bins (ALMA bands), and L_{IR} is extrapolated based on preliminary SED modelling to estimate the total infrared luminosity (e.g., Fudamoto et al., 2020a; Bouwens et al., 2020; Schouws et al., 2021). To get IRX from the simulations, I compute the denominator in the same way as for observations: $1600[\text{Å}] \times L_{1600}[\mathrm{erg/s/Å}]$. For the numerator L_{IR}, I do not transfer any infrared photons, rather I compute it by looking at the difference between the integrated intrinsic UV + optical light and the observed one, which is equal to the total infrared light re-emitted by dust. This might be slightly different than the observational approach for L_{IR}, which is more indirect.

In this section I use both the zoom simulation, described in Chapters 2 and 4, and the Sphinx20 simulation, also described in Chapter 2. Due to the large number of galaxies in Sphinx20, I analyse only one snapshot, at redshift 5.11. Hence I do not look at galaxies at different times as I did for the zoom simulation, but the presence of many galaxies in Sphinx20 is a great advantage to study a statistical set of galaxies at a given time, as it is done in observations. I selected the 100 most massive halos of the snapshot at $z = 5.11$ and made mock observations of them as seen from 12 directions of observations.

In Figure 5.2 we see the luminosity function computed from the Sphinx20 simulation at redshift 6, taken from Rosdahl et al. in prep. Our modelling faithfully reproduces the luminosity

function of Bouwens et al. (2017), derived from observations of the Hubble Frontier Fields at redshift 6, as well as the one of Livermore et al. (2017), where they study 167 strongly lensed galaxies in two clusters of the same fields. Figure 5.2 also shows that the dust attenuation of the brightest simulated galaxies is intense: the angle-averaged population-averaged A_{1500} reaches values of slightly more than 2 for galaxies with observed magnitude of -20.

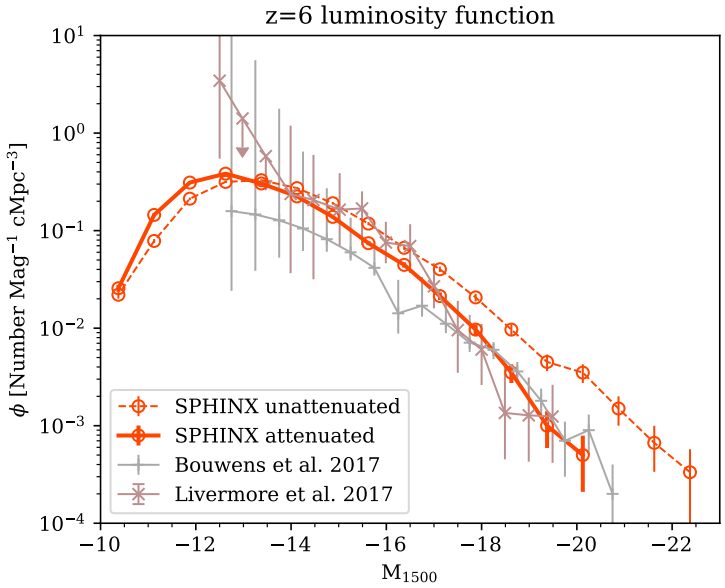

Figure 5.2: From Rosdahl et al in prep. Luminosity function at 1500Å in the Sphinx20 simulation at redshift 6, compared with data from Livermore et al. (2017) and Bouwens et al. (2017). The dashed red line shows intrinsic magnitudes and the continuous red line shows magnitudes after dust attenuation, following the same law as described in Section 3.5.

In Figure 5.3 I compare β slopes and stellar masses between simulations, namely the usual zoom simulation and the Sphinx20 simulation, and observations from Fudamoto et al. (2020b), which also includes a compilation of previous works. The galaxies of Fudamoto et al. (2020b) are observed with band 6 and band 7 of ALMA in the COSMOS field. The figure shows that there is an excellent agreement in the β slope - stellar mass relation between simulations and observations. The range of the zoom simulation, even if spanning 61 snapshots, is quite limited, but falls right at the same place as observations, for both McLure et al. (2018), who use ALMA data of redshift 2 to 3 galaxies from the Hubble Ultra Deep Field and Fudamoto et al. (2020b). The Sphinx20 data also follows the same relations, although the scatter at low-mass is larger than the one of the observations, probably due in part to the low number of observed low-mass galaxies. Additionally, there is only limited Sphinx20 data above $10^{10}\,M_\odot$, which explains the small range of values at this mass. We cannot exclude that the good agreement is due to a counterbalancing between high star-formation rates (our stellar masses are above what is expected, see Section 2.2) and a strong dust attenuation. To gain insight into this question, I indirectly look at infrared properties of the galaxy with the infrared excess.

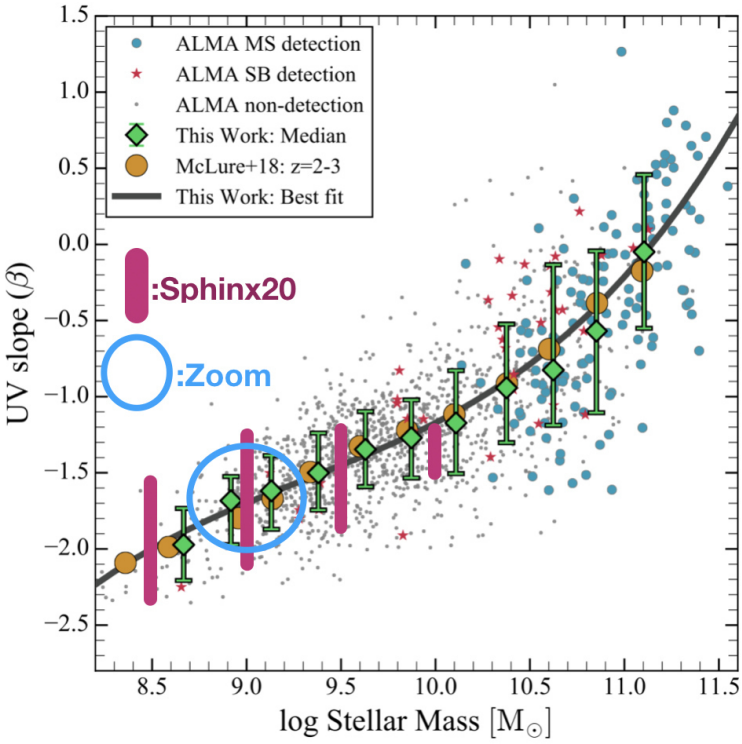

Figure 5.3: Comparison of β slope - stellar mass relation between observations and simulations. The data for the Zoom is taken from 61 snapshots between redshift 4.6 to 3, analysed from 108 directions of observation. Sphinx20 results are from the 100 most massive halos in the snapshot at redshift 5.11. The results from the simulations are overplotted on a figure from Fudamoto et al. (2020b). The green diamonds and error bars show new data of Fudamoto et al. (2020b), with the best fit curve in black. The yellow points are from McLure et al. (2018).

The IRX-beta relation taken from Fudamoto et al. (2020a) is displayed in Figure 5.4, along with the corresponding data from galaxies of the zoom simulation and of Sphinx20. The galaxies of Fudamoto et al. (2020a) are observed with ALMA, with the large program ALPINE, at redshifts between 4.4 and 5.8. The observational and theoretical publications display a wide range of values. The simulation data falls on the large-IRX side of the plot, relatively close to Meurer et al. (1999) and several of the compiled data, but in tension with Fudamoto et al. (2020a). This points out that the dust attenuation from our simulations is rather strong, while still being quite realistic, i.e. comparable to some of the observations. It is also plausible that the methods of measurements of IRX in the simulation differs slightly from the observational way, which might introduce a bias. There are other observational studies that are not in Figure 5.4 but that find an IRX-β relationship with which our simulation match better than with the results of Fudamoto et al. (2020a): Bouwens et al. (2016), where the data of the ASPECS survey (Walter et al., 2016; Aravena et al., 2016) or McLure et al. (2018). Additionally, there are recent studies of single star-forming galaxies at high redshift that find more dust attenuation than what Fudamoto et al. (2020a) predicts, as in Bakx et al. (2021). The fact that we find a realistic β slope compared to stellar mass, but IRX too large for a given β suggests that the intrinsic slope of the stars in the simulations are strongly negative. This is due to the BPASS models, which emit strong UV radiation, resulting in a highly negative intrinsic slope. Finally, studies of IRX have been done with simulations (e.g. Vijayan et al., 2021; Liang et al., 2021). In Vijayan et al. (2021), the simulations are performed with the SPH code EAGLE, and they also find values of IRX above the ones of Fudamoto et al. (2020a). However, in Liang et al. (2021), where they use simulated galaxies from the MassiveFIRE suite (Feldmann et al., 2016), they find IRX values more consistent with observations, suggesting that more work is needed to understand what causes these differences between simulations.

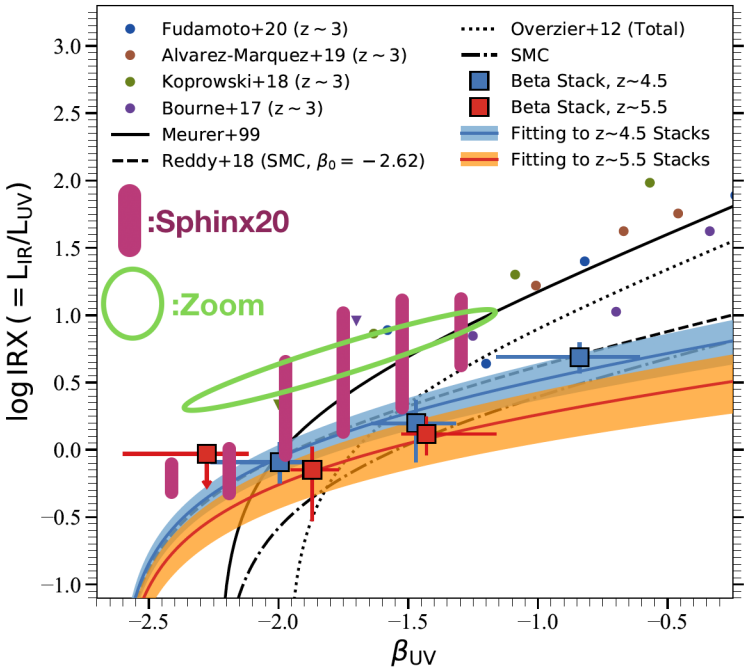

Figure 5.4: Comparison of the IRX - β relation between observations and simulations, for the same Sphinx20 and zoom simulation mock data as in Figure 5.3. The Sphinx20 relation is plotted with the purple vertical bars, and the zoom simulation results with the green oval. The main observational data is from Fudamoto et al. (2020a), in blue for redshifts around 4.5 and in orange for redshifts around 5.5. Other older observational data points are included with dots of different colours. The blue points are from Fudamoto et al. (2020b), the brown points from Álvarez-Márquez et al. (2019), the green points from Koprowski et al. (2018) and the purple points from Bourne et al. (2017). The black lines are different theoretical models: the solid line is from Meurer et al. (1999), the dashed line from Reddy et al. (2018) and the dotted line from Overzier et al. (2011).

We have seen in Chapter 4 that dust also affects down-the-barrel absorption lines significantly. Hence, for the next verification of the realism of the mock observation, we compare the absorption lines from Chapter 4 with new results from Saldaña-Lopez et al in prep., where they analyse low redshift analogs of high redshift galaxies, with low metallicity and high star-formation rates. This is shown in Figure 5.5. The distribution of values is remarkably similar between the observations and the simulation. The figure in the left panel also shows previous results from Reddy et al. (2016) at high redshift, over a smaller range of values, and it falls right on the overdensity part of the data from the simulation.

Those results are encouraging and show that our Ramses-RT simulations and post-processing methods reproduce realistic looking star-forming galaxies at high redshift, even though there are many young stars with extreme dust attenuation, which may need to be corrected by some new feedback recipes. Now that we are confident that our simulations are realistic, in the next section I explore the consequences of our inhomogeneous dust distribution on the estimation of dust attenuation properties from observations.

5.3 Another look at dust attenuation with GRB afterglows

As another comparison between observations and simulations, without doing a direct superposition of data, I would like to evoke the question of Gamma Ray Burst (GRBs) afterglows (see Dainotti & Del Vecchio, 2017 and references therein). There are two kinds of GRBs. The long GRBs are thought to be a kind of supernova explosion of young massive stars. The short-GRBs

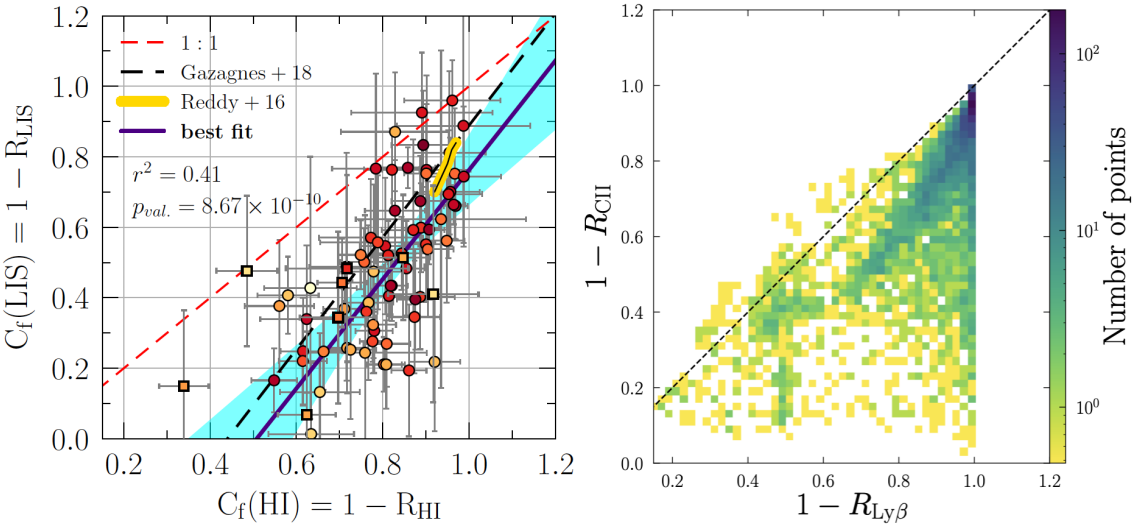

Figure 5.5: Comparison of neutral hydrogen and metal residual fluxes of absorption lines between observations and simulations. Left panel: the data is taken from Saldano-Lopez et al. in prep. The x-axis shows (one minus) the residual flux of H I lines, averaged between several lines of the Lyman series, and the y-axis is averaged over several LIS lines, Si II and C II. The black dashed line is from Gazagnes et al. (2018) and the yellow segment from Reddy et al. (2016). Right panel: the data is Mauerhofer et al. (2021), with the residual fluxes of C II and Lyβ for the three snapshots of the zoom simulation and 1728 directions of observation per snapshot. The data is presented in the form of a 2d-histogram, to better visualise where the majority of the points are.

likely emanate from the merger of a neutron star binary. Since we are interested in star-forming galaxies, long GRBs are more important, since they trace young stars, while neutron star binaries take hundreds of millions to billions of years to explode, and therefore they do not necessarily trace star-forming galaxies. Long GRBs emit a very energetic jet of gamma-rays and X-rays, which can be detected with space satellites such as SWIFT up to extremely high redshift. While there is still ongoing research on the exact mechanisms at play, these jets heat the ISM close to the explosion to extreme temperatures, thus emitting intense radiation at all wavelengths. This radiation is called the afterglow. Satellites that detect high energy photons send an alarm so that ground based instruments such as X-shooter or UVES on the VLT can observe the afterglow, which lasts from a few hours to a few days, progressively losing luminosity. GRB afterglows offer a probe into the ISM of the host galaxy, and a way to measure the dust extinction along the sightline of the afterglow.

Many studies have looked at the dust attenuation of GRB afterglows (e.g., Perley et al., 2016a,b; Zafar et al., 2018) and find that many of them, around a third, are compatible with no dust extinction. The majority are attenuated, but typically by about 0.5 to 3 magnitudes. According to Perley et al. (2016a), about 20% of GRBs are ”dark”, i.e. their afterglow is undetected. One of the explanations for these non-detections is the presence of strong dust attenuation in the host galaxy, making the UV and optical afterglow invisible. To compare with our simulations, Figure 5.6 shows the distribution of dust extinction for young stars in three snapshots of the zoom simulation that we also study in Chapter 4. I selected stars of ages between 3 and 15 Myr, which is the range of age at which long GRBs explode according to Tanvir et al. (2019). Here a weakness might arise in the realism of our simulations. From Figure 5.6 the vast majority of such young stars are embedded in extremely dusty clouds. More precisely, around 75% of young stars in output A and 85% in outputs B and C have $A_{1500} > 5$, which is enough to prevent detection of the afterglow. The gamma rays however are insensitive to dust. According to the SMC law, the cross-section of interaction of gamma rays with dust is at least seven orders of magnitude lower than the one of FUV photons. Thus, the GRBs of the young stars in the zoom simulation would be detected

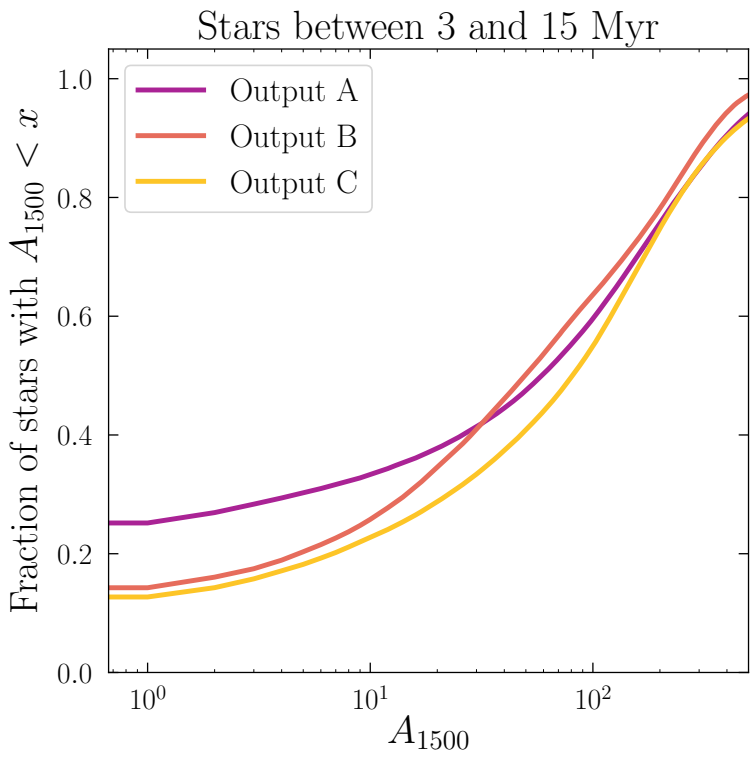

Figure 5.6: Cumulative distribution of dust attenuation for young stars between 3 and 15 Myr. The three colours correspond to different snapshots of the zoom simulation, the same three as in Chapter 4. Output A is at redshift 3.2 and contains 36'000 stellar particles in the chosen range, while output B is at redshift 3.1 with 21'000 selected stellar particles and output C at redshift 3.0 with 30'000 stellar particles. The distributions are averaged over 110 directions of observation. The dust attenuation is computed at 1500 Å and expressed in magnitudes.

in gamma rays but the afterglow would be invisible. That is the definition of dark GRBs, whose incidence in observations is only 20%, much lower than the ~ 80% predicted by our simulation. Does that mean that the simulations are unrealistic in terms of dust attenuation? It is indeed an additional clue, like the high IRX of Figure 5.4, that the extinction is too strong. However, several factors can alleviate this tension.

First, GRBs are extremely intense explosions that are believed to obliterate everything in a significant radius along the gamma ray jet, including dust. Hence it is possible that the GRB itself reduces the amount of dust attenuation its afterglow will be experiencing. This can be tested in the simulation by removing dust in the neighbourhood of stellar particles before computing their A_{1500}. Preliminary tests show that this is insufficient to reconcile the models with observations. Second, GRBs are thought to happen in low metallicity environments, which intrinsically have less dust. However this is probably not enough to explain such a difference of percentage of dark GRBs, and there might be subgrid physics and feedback processes that should be added in our simulations to decrease this percentage. One promising and relatively easy method is the inclusion of runaway stars (Kimm & Cen, 2014; Andersson et al., 2020, 2021). Cen & Kimm (2014) show that indeed the runaway stars result in many GRB explosions outside the dense star-formation clouds, reducing the typical dust attenuation of afterglows. In a future work I will analyse a new Ramses-RT zoom simulation including runaway stars, run by Maxime Rey, to study this question.

5.4 Consequences of non-homogeneous dust for deriving extinctions: SED-fitting and Balmer decrement

In this section I explain the methods commonly used to deduce the dust attenuation from spectroscopic observations of star-forming galaxies, and test those methods on the mock observations in order to compare with the actual extinction that we know from the simulation. I focus on the method of fitting the FUV continuum to a stellar library, thus deriving the parameter $E_{\mathrm{B-V}}$ that describes the dust attenuation, and the method of comparing Hα and Hβ to deduce the extinction of those nebular lines. Although we have seen in Sections 5.2 and 5.3 that the simulations might have slightly too strong dust absorption, it is insightful to test whether the observational methods derive the correct attenuation. What we learn still holds qualitatively for real observations even if there are differences between the simulations and the observations. This section builds on Chapter 4, since the knowledge of intrinsic emission of UV or nebular lines is important to derive escape fractions of ionising photons.

5.4.1 SED-fitting of the FUV continuum

Knowing the intrinsic FUV continuum is essential to get escape fractions, as they are usually computed by comparing the intrinsic ratio between the Lyman continuum emission and the UV continuum to the observed ratio. One method to derive the dust extinction of the stellar continuum is to fit the observed FUV continuum with a linear combination of spectra of given ages and metallicities from a stellar library, plus a global extinction terms. One can thus write the observed continuum F_λ as

$$F_\lambda = 10^{-0.4 E_{\mathrm{B-V}} k_\lambda} \sum_{\mathrm{age,Z}} F_\lambda^{\mathrm{library,age,Z}}, \tag{5.2}$$

where $E_{\mathrm{B-V}}$ is the parameter that describes the strength of the dust attenuation, k_λ is the dust extinction law that sets the dependence with wavelength of the attenuation and $F_\lambda^{\mathrm{library,age,Z}}$ are collections of spectra from an SED library such as Starburst99 (Leitherer et al., 2014) or BPASS (Eldridge et al., 2008; Stanway et al., 2016). Examples of recent studies using this procedure are Steidel et al. (2018); Chisholm et al. (2018a); Saxena et al. (2021), or Saldaño-Lopez et al. in prep. The product $E_{\mathrm{B-V}} k_\lambda$ is equal to the dust attenuation in magnitude, A_λ, that I have used in Figures 5.1 and 5.6. The main assumption behind Equation 5.2 is that dust forms a uniform screen in front of all the stars, which attenuates the light from all of them in the same way no matter the age. This contrasts with what I have been showing in this section, the young stars experience more dust attenuation than older stars.

To visualise the differences between the dust attenuation of the stellar continuum predicted by Equation 5.2[1] and the real attenuation of the photons in the simulation, I plot the two quantities in Figure 5.7. Interestingly, the points are almost parallel to the one-to-one relation, but shifted by about one magnitude to the right. This means that fitting the stellar continuum with Equation 5.2 can efficiently determine whether one galaxy has more dust attenuation than another, but it underestimates the dust absorption significantly. The large gap between the dashed line and the data points of Figure 5.7 is likely enhanced by the fact that the majority of young stars in the simulation have extreme dust attenuation. The gap might be reduced if using a runaway-stars algorithm, which kicks young stars out of their cloud of origin, thus lowering the dust attenuation. However, I would argue that the gap would still exist even for perfectly realistic simulations, since the presence of only a fraction of young stars with $A_{1500} > 5$ would be enough to hide part of the intrinsic UV.

This suggests that assuming a homogeneous dust screen when fitting observed spectra is not

[1] Thank you to Alberto Saldaño-Lopez for sharing his code to fit the stellar continuum.

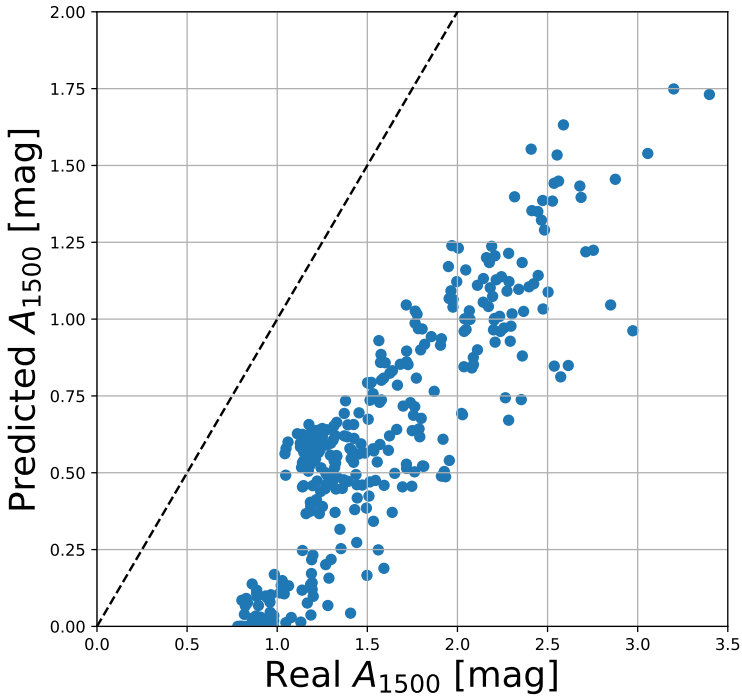

Figure 5.7: Comparison of the dust attenuation of 1500Å photons in the zoom simulation, on the x-axis, with the dust attenuation that would be derived from Equation 5.2, on the y-axis. The data comes from the same three snapshots of the zoom simulation that are analysed in Chapter 4, seen in 108 directions of observation for each.

accurate enough. A possible alternative is to use codes such as BEAGLE[2] (Chevallard & Charlot, 2016), used for example in Marchi et al. (2019), which fits observed spectra assuming a more complex dust geometry. In particular, BEAGLE assumes a stronger dust extinction for young stars, following e.g. Charlot & Fall (2000). In future work it will be interesting to redo Figure 5.7 with A_{1500} inferred with BEAGLE instead of Equation 5.2. Additionally, rest-frame FIR observations of galaxies are ideal to shed light on the question of dust attenuation, since the UV and optical light hidden by dust is re-emitted in the FIR. More observations of this kind will help to determine the exact dust extinction, which will be useful to infer escape fractions of ionising photons. Indeed, we have seen in Chapter 4 that one needs to know the intrinsic FUV continuum when inferring the escape fraction with LIS absorption lines.

5.4.2 Balmer lines

A different way to study dust attenuation is to look at the ratio of nebular lines, such as the Balmer series of hydrogen, which contains information about the nebular extinction. It is also relevant for studies of Lyman Continuum (LyC) escape at low redshift. Indeed, one way to measure the escape fraction of ionising photons when having access to an observed LyC flux is to estimate the intrinsic LyC flux with Balmer lines. The intrinsic LyC flux is the sum of the observed one and the one that was absorbed in the ISM. Assuming that only neutral hydrogen absorbs ionising photons, which is what I find in Chapter 4, the absorbed LyC flux can be derived from the intrinsic Hα or Hβ. Thus, correcting the observed Balmer lines for dust attenuation allows us to get the intrinsic lines and to deduce the escape fraction of ionising photons.

The method to derive the attenuation of, for example, Hα is to compare the flux ratio Hα/Hβ with its intrinsic ratio, fixed by atomic physics. According to Osterbrock & Ferland (2006), this

[2] http://www.jacopochevallard.org/beagle/

flux ratio is 2.87 at a temperature of 10′000K, and varies slightly from 3.3 to 2.76 when the temperature goes from 2'500 to 20'000K. Since dust absorbs photons with shorter wavelengths more efficiently, Hβ is more affected than Hα, leading to a larger observed ratio. If we call $R_{\rm intr}$ the intrinsic Hα/Hβ ratio, $R_{\rm obs}$ the observed ratio and $R_{\alpha\beta} \equiv k_{4863}/k_{6565}$ the ratio of the extinction law at the wavelengths of the two lines, a short analytical development gives the extinction $A_{\rm H\alpha}$, in magnitude, that the $H\alpha$ line is experiencing:

$$A_{\rm H\alpha} = \frac{2.5}{R_{\alpha\beta} - 1} \log\left(\frac{R_{\rm obs}}{R_{\rm intr}}\right). \tag{5.3}$$

Usually one assumes $R_{\rm intr} = 2.87$, and, using the SMC extinction law, one finds $R_{\alpha\beta} = 1.515$. Once again, there is a hidden assumption here that the dust is uniformly affecting all the regions emitting Balmer emission. If there are several regions with different emissivities and different dust attenuation, we cannot use a global $R_{\alpha\beta}$ anymore.

To compare the predicted $A_{\rm H\alpha}$ with the values from the simulation, I adapted RASCAS to transfer Balmer lines. The emission of Balmer lines is done by the gas rather than by stellar particles, which changes slightly the preparation of the initial conditions of photon packets compared to Chapter 3. I use the same procedure as for Lyman alpha emission, described in Michel-Dansac et al. (2020). To summarise, the Lyman alpha recombination emissivity of each gas cell is computed following Cantalupo et al. (2008). The emissivity of each Balmer line is a fraction of this Lyman alpha emissivity, following Osterbrock & Ferland (2006). The photon packets of RASCAS are then distributed to each gas cell based on their emissivity, just they are distributed to stellar particles when considering absorption lines or a continuum. The exact initial position of a photon packet in its starting cell is chosen uniformly at random. For Balmer emission, I omit the collisional emission, which is subdominant. The initial frequency of photon packets is chosen such that it follows a Gaussian emission line whose width depends on the Doppler parameter of the gas cell where the packet is emitted. Once the initial conditions of the photon packets are set, the transfer is performed, during which the only interaction that the photon packets have is with dust. With the peeling-off algorithm, this results in mock observations of Balmer emission lines.

After obtaining mock Hα and Hβ lines of the most massive halos of Sphinx20, I apply Equation 5.3 to obtain the inferred $A_{\rm H\alpha}$. The real attenuation is obtained by comparing the intrinsic Hα luminosity with the observed ones from the mock spectra. I show those two quantities in Figure 5.8. Again, it is striking that the dust attenuation is underestimated with the observational method, however, the shift between the one-to-one relation and the data points is not constant as in Figure 5.7. Rather, there are galaxies with low Hα extinction where the prediction of Equation 5.3 is exact, but there are also points that go to high extinction and the prediction stagnates below $A_{\rm H\alpha} = 1$. Again, the effect is likely increased by the strong dust absorption in our simulations, but this shows that assuming a homogeneous dust screen also prevents us from getting accurate dust attenuation of Balmer lines.

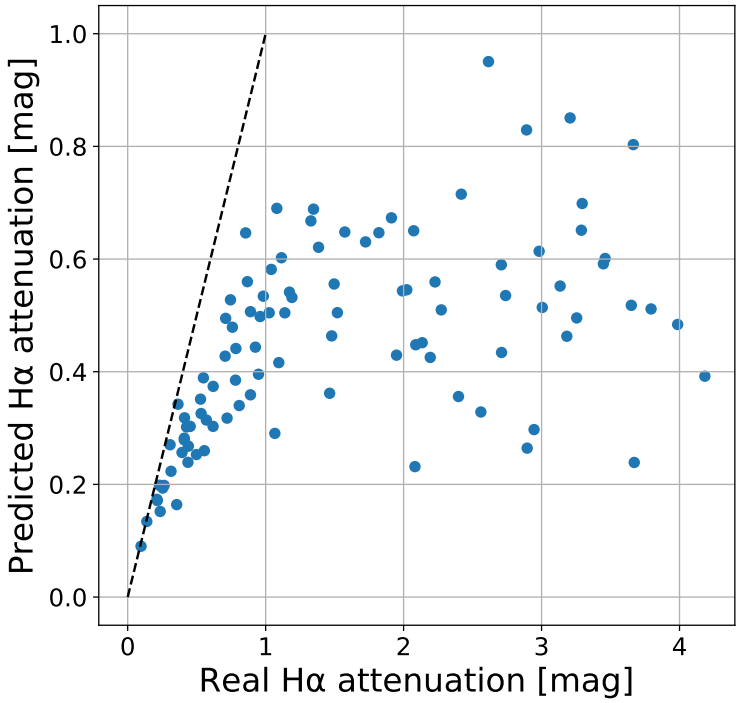

Figure 5.8: Comparison of the dust attenuation of Hα photons in the simulation, on the x-axis, with the dust attenuation that would be derived from the method described above for Balmer lines, on the y-axis. The mock observations are made on the 100 most massive halos of the Sphinx20 simulation, in one given arbitrary direction of observation.

6 Gas flows and CGM properties from absorption lines

In this chapter I move away from the question of escape of ionising photons to focus on the properties of the CGM of our simulations. The infall of gas which fuels star-forming galaxies and the energetic, metal rich outflows polluting the CGM due to stellar feedback, are the key processes for understanding galaxy evolution. Outflows are ubiquitously observed across redshifts, through blueshifted down-the-barrel absorption lines (Erb et al., 2012; Martin et al., 2012; Rubin et al., 2014; Zhu et al., 2015; Chisholm et al., 2015, 2016, 2017, 2018b). In this chapter I present the links between down-the-barrel absorption lines and gas flows in our simulations. In the first section, I describe the properties of neutral and ionised outflows and inflows at different radii around the galaxy of the zoom simulation described in Chapter 2. In the second section, I explore how well the centroid velocity of absorption lines correlates with gas flows. Then, in the third section I study the formation of absorption lines in terms of distance from the sources. I determine whether the ISM or the CGM contributes more to the lines. Finally, in the fourth section I examine the question of the ratio of observations of blueshifted compared to redshifted absorption lines. My mock spectra are more often redshifted than in observational studies, so I explore whether the simulations do not reproduce observations, or whether there is a bias in observations.

6.1 Overview of gas flows in the zoom simulation

The first step for conducting this new study is to compute the inflow and outflow rates around our simulated galaxies. I focus on the zoom simulation, presented in Chapter 2, which contains a galaxy evolving down to redshift 3, where it has a stellar mass of 2.3×10^9 $M_\odot$. Gas flows at high redshift are complex, due to intense infalls of pristine gas on young galaxies, accompanied by outflows due to supernova feedback. Additionally, for most of the simulation snapshots, the galaxy is irregular, which means that it did not have time to develop a disk structure yet. This means that the gas flows are highly inhomogeneous and anisotropic. Looking at different angles will result in different types and intensity of gas flows. Only the last few snapshots of the simulation, below a redshift of around 3.3, begin to develop a disk of stars and gas. This can be shown by looking at the evolution of the angular momentum of the stars and gas, which begins to stabilise at the end of the simulation, as can be seen in Figure 2.6. This indicates that the movement of stars and gas goes from a chaotic movement to a regular movement around an axis of rotation.

In this work I focus on angle averaged quantities, to have a global view of the properties of gas flows. I compute these gas flows at different radii around the galaxy, since the CGM is large (around 30 kpc in radius, while the ISM around 2 kpc). Additionally, I also split the gas flows into ionised and neutral gas (hydrogen). The analytical equation to compute net gas flows $F(R)$ through a sphere of radius R is

$$F(R)\,[\mathrm{M_\odot\,yr^{-1}}] = \int_{S_R} \vec{v}_{\mathrm{gas}} \cdot \vec{\mathrm{d}r}\, \rho_{\mathrm{gas}}\, \mathrm{d}S, \tag{6.1}$$

where S_R is the surface of the sphere, $\vec{v}_{\rm gas}$ is the gas velocity at the surface of the sphere, $\vec{dr}$ is the unit vector perpendicular to the surface and $\rho_{\rm gas}$ the mass density of the gas. This equation is the sum of outflows and inflows. To select only outflows (inflows) one can simply select the regions where $\vec{v}_{\rm gas} \cdot \vec{dr}$ is positive (negative). Similarly, to select neutral or ionised gas, I change $\rho_{\rm gas}$ to the density of the type of gas that I am interested in.

To apply Equation 6.1 to simulation data, it has to be discretised. Additionally, I average the gas flows on a shell with a thickness equal to 15% of its radius, such that it encompasses several layers of cells, to smooth the values of gas flows. Here is the equation for the discretised gas flow (e.g. Dubois & Teyssier, 2008):

$$F(R) = \frac{\sum_{\rm cells} m_{\rm gas} v_r}{\sum_{\rm cells} {\rm Vol}_{\rm cells}} S_R, \tag{6.2}$$

where $v_r = \vec{v}_{\rm gas} \cdot \vec{dr}$, for which positive and negative values can be selected for outflows and inflows, ${\rm Vol}_{\rm cells}$ is the volume of the cells and S_R the total area of the sphere at the middle of the shell.

However there is one more question to be addressed when computing gas flows, which is the definition of what we count as outflows or inflows. Generally we imagine that an outflow is gas that is accelerated away from the galaxy by feedback, and an inflow is gas falling from the IGM to feed the growth of the galaxy. There is one different kind of gas flow that we would be accounting for with Equation 6.2 which is the movement of satellite galaxies around the central object. There are numerous small galaxies orbiting our simulated galaxies, and when they cross the shell where we compute the gas flows, their ISM gas will be counted as outflow or inflow, which is not what we want. To remove the contribution of the satellites, I look for their position and radius in the simulation catalog containing all the information on halos and sub-halos (i.e. satellites), and remove all the cells that are within the virial radius of a satellite galaxy from Equation 6.2. I highlight the differences that it makes in the gas flows in Figure 6.2. More details on the different components of the gas, their dynamics, and the treatment of satellites can be seen in Mitchell et al. (2021). In this study the authors use a simulation similar to our zoom simulation, but including tracer particles, which follow the history of the gas displacement, allowing them to know at any time from where the gas is coming.

The Figures 6.1 to 6.4, are showing the evolution in time of the gas flows and star-formation rates for the zoom simulation. The neutral outflows are relatively small, as can be seen in Figure 6.1, with an average of about $0.75\,{\rm M}_\odot\,{\rm yr}^{-1}$ at $0.2\,R_{\rm vir}$. Only close to the galaxy, at $0.1\,R_{\rm vir}$, can the neutral outflows be large, with an average of $2.5\,{\rm M}_\odot\,{\rm yr}^{-1}$. The mass outflow rates at larger radii become smaller and smaller, even negligible at $1\,R_{\rm vir}$. The star-formation rate is larger than the H^0 mass outflow rate, with an average of around $3\,{\rm M}_\odot\,{\rm yr}^{-1}$ after redshift 5. There seems to be a small correlation between the H^0 mass outflow rate and the SFR, the former increasing shortly after bursts of star-formation, especially when focusing on the curve at $0.2\,R_{\rm vir}$. This is expected since it is the feedback of young stars that drives outflows.

The neutral outflows become occasionally much larger if satellite galaxies are taken into account. In Figure 6.2 I compare the neutral outflow rate at $0.2\,R_{\rm vir}$ with and without satellites. We see that the values of outflows oscillate strongly in time due to the neutral ISM of satellites passing through the region in which the mass outflow rate is computed. In this case the outflow does not represent gas that is ejected from the main galaxy, but movement of satellites.

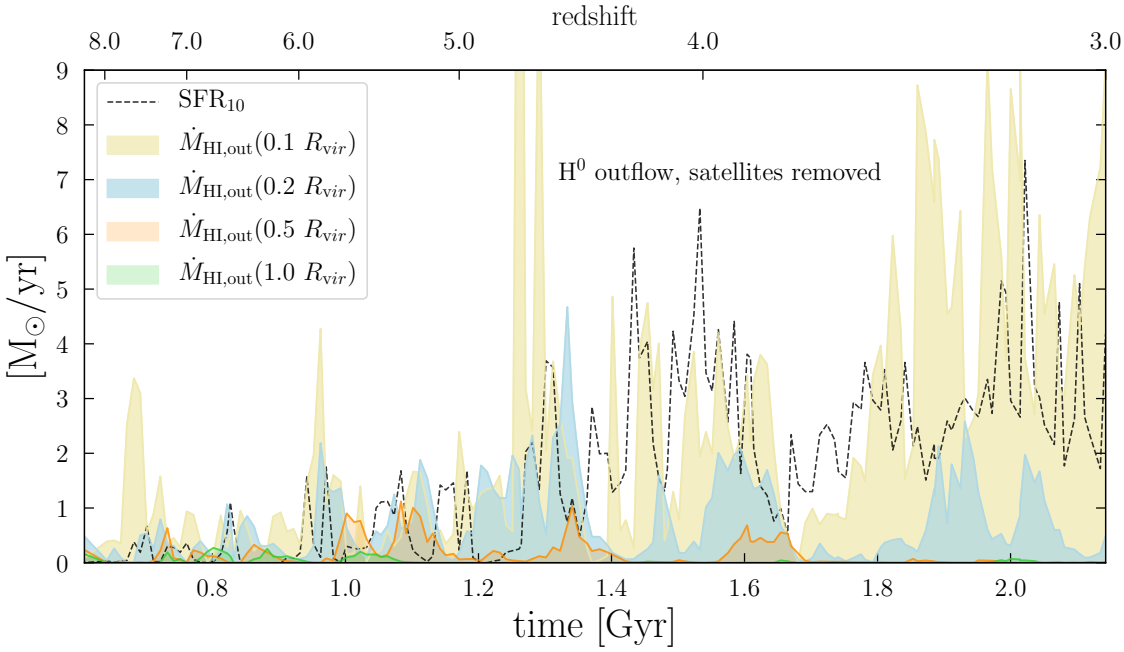

Figure 6.1: Mass outflow rate of neutral hydrogen as a function of time, integrated over a sphere of different radii. The gas in the virial radius of satellite galaxies has been removed. The black dashed line shows the star-formation rate averaged in the last 10 Myr.

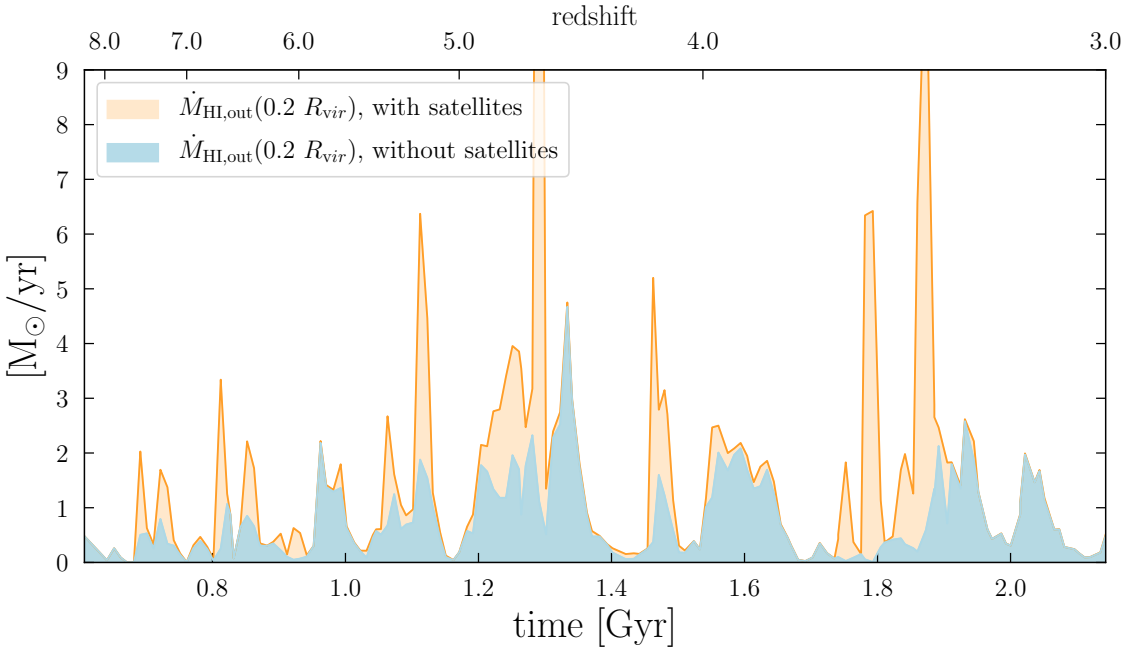

Figure 6.2: Mass outflow rate of neutral hydrogen as a function of time, at 20% of the virial radius. The blue curve is without satellite galaxies, as in Figure 6.1, while the red curve includes the contribution of satellite galaxies.

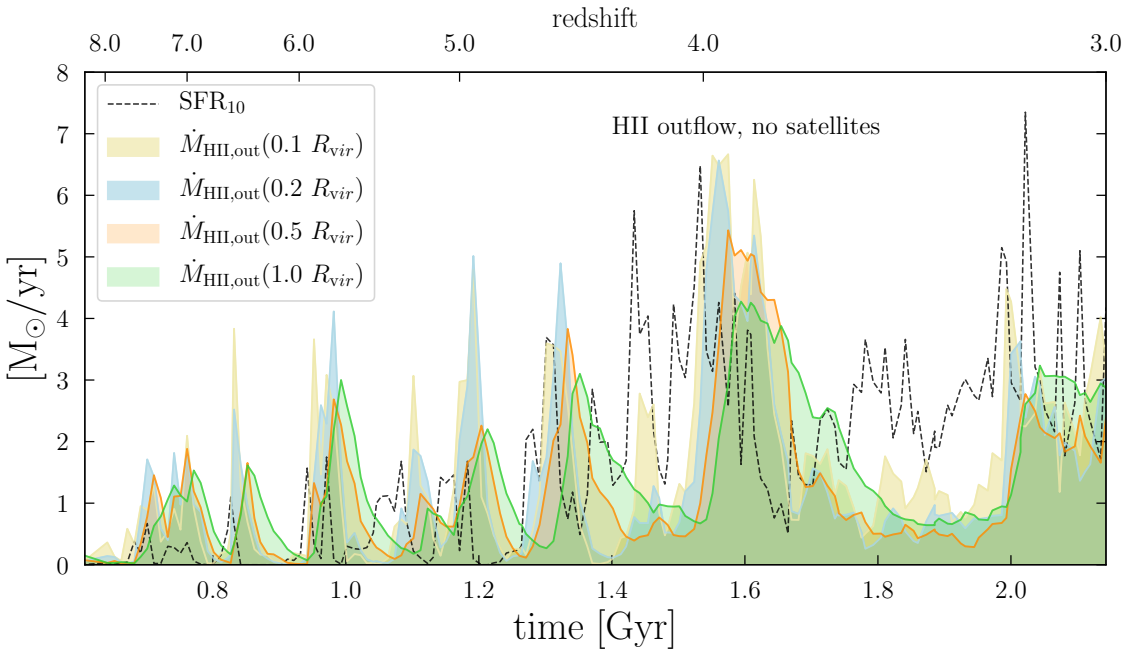

Figure 6.3: Same as Figure 6.1 but ionised hydrogen instead of neutral hydrogen. Curves including satellite galaxies are indistinguishable from the ones without satellites, therefore they are not shown.

The ionised outflows are much larger on average than the neutral ones, regardless of the presence of satellites, as can be seen in Figure 6.3, compared to Figure 6.1. The gas mass of satellite galaxies is by far dominated by neutral gas, so that the plot looks the same even if satellites are included (hence the version with satellites is not plotted here). This hints at feedback mechanisms in our simulations heating the gas to high temperatures when it accelerates out of the galaxy, resulting in ionised gas. I come back to discuss the realism of such ionised outflows in Section 6.4. For ionised gas it is even clearer that there is a correlation between the bursts of star-formation and the bursts of gas ejection. Only at 1.8 Gyr is there significant star-formation but weak outflows.

Figure 6.4 shows that there is more neutral inflow than neutral outflow at $0.2\,R_{\rm vir}$, which is what allows the galaxy to grow.

Now that we have a global view of the properties of outflows and inflows at different radii, also in comparison to the star-formation rate, let us dive into the question of the relation between those gas flows and the observables, in this case the absorption lines.

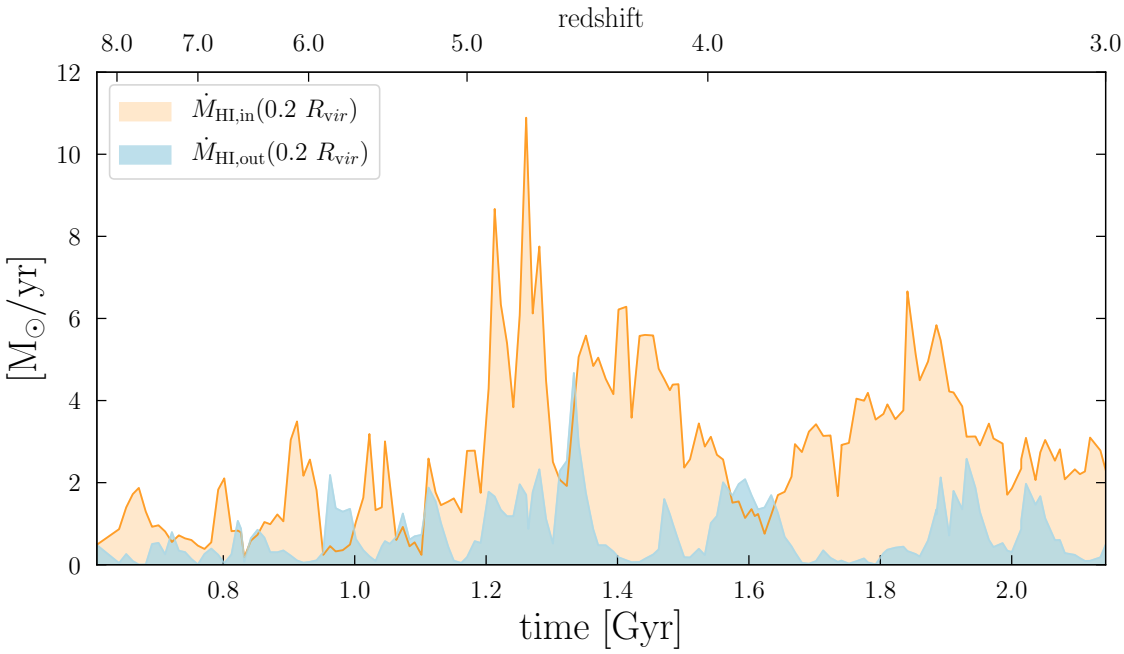

Figure 6.4: Mass outflow and inflow rate of neutral hydrogen as a function of time, at 20% of the virial radius. Satellite galaxies have been removed.

6.2 Tracing gas flows with absorption lines

The main motivation of this section is to assess how well we can measure the gas flows using down-the-barrel absorption lines. There has been extensive research on this topic (e.g. Martin et al., 2012; Chisholm et al., 2017; Carr et al., 2018), modelling the outflows with bi-conical shapes, having a characteristic opening angle, length, gas velocity, column density and covering fraction of the absorbant. I have shown in Chapter 4 the limitations of the concept of covering fraction, and of adopting a single column density, plus the effect of dust on absorption lines which is often not taken into account but affects the relation between lines and gas flows. High-resolution simulated galaxies are the perfect way to add all those elements of complexity and see how well absorption lines trace the gas flows. This is a preliminary study of the topic and will require more in-depth analyses in the future.

For this study I chose to use the Fe II λ2586 line, which is another low ionisation state (LIS) absorption line, with a longer wavelength than C II or Si II lines. Most studies of outflows and absorption lines are done with this line, sometimes along with Mg II $\lambda\lambda$2796, 2803, at redshifts between 0.5 and 2. This is because it is challenging to detect absorption lines of individual galaxies above redshift 2, due to the quickly increasing luminosity distance, which decreases the apparent brightness of galaxies. In addition, Fe II λ2586 is chosen over the usual Si II λ1260 and C II λ1334 because the latter are not observable from the ground for redshifts below 2.

The procedure to make mock observations of absorption lines is described in Chapter 3. There, I show how to include the effect of fine-structure levels of the ground states, of which Fe II λ2586 has several. However, PyNeb computations show that the fine-structure levels of Fe^+ are negligibly populated, and the resulting absorption lines from a test Rascas run including them show that fine-structure levels have do not modify the spectra. Therefore, I do not use them for the results presented here.

The absorption line property that is most expected to correlate with gas flows is the centroid velocity $V_{\rm center}$. As a reminder, a spectrum can be expressed as a function of velocity instead of

wavelength, via the relation

$$V = \frac{\lambda - \lambda_{\rm ref}}{\lambda_{\rm ref}} c,$$

where $\lambda_{\rm ref}$ is the wavelength at which the velocity is defined to be 0, typically the wavelength of a spectral line of interest. This definition is useful to make a link between the Doppler effect and the change of wavelength of a line. If an absorbing cloud in front of a star has a velocity $V_{\rm cloud}$, an absorption line will be spectrally shifted by $V = V_{\rm cloud}$, no matter its wavelength $\lambda_{\rm ref}$. The centroid velocity $V_{\rm center}$ is the velocity that corresponds to the line wavelength where the area of the line is divided in two equal parts, as illustrated in Figure 3.12 of Chapter 3. In a galaxy there is not only one cloud absorbing light from one star, but a collection of many stars and clouds, each having different velocities. The centroid velocity then corresponds to the average velocity difference between sources and absorbants, weighted by the luminosity of the sources (e.g. stars with strong dust attenuation do not contribute to the line, so the velocities of the gas in front of them do not shape the absorption line). If $V_{\rm center}$ is positive, it indicates that on average the clouds are falling towards the sources, i.e. an infall of gas. More precisely, it means that the component of the velocity parallel to the line-of-sight is pointing from the gas to the sources. Inversely, If the centroid velocity is negative, the absorbing clouds are moving away from the sources, i.e. there is outflow of gas. This is why the centroid velocity is expected to correlate with gas flows. In this section, I compare the centroid velocity with outflows and inflows and neutral and ionised gas in the zoom simulation. As a first step, I compare two ways of determining centroid velocities in simulations.

6.2.1 Balmer lines and the determination of the centroid velocity

The definition of the centroid velocity is actually more subtle than I just presented. The intrinsic motion of the galaxy complicates the measurement of its centroid velocity. Let us assume that we observe a galaxy at redshift 0, so that its only movement is due to its peculiar velocity and not to the Hubble flow. This corresponds to the situation of the mock observations I am creating in this work: the galaxies are moving in the simulation boxes, but I do not apply any Hubble flow to the wavelength of my mock observations. The peculiar velocity of a galaxy is generally measured by comparing the observed wavelength of an emission line with its intrinsic wavelength, given by atomic physics. Emission lines are used mainly because they are easier to observe than absorption lines. But they also have the advantage of being emitted by all the gas of the galaxy, integrated on the entire volume, which averages the velocity of the emitting gas. On the contrary, absorption lines are created only by the gas between the stars and the observer, which means that their wavelength is shaped by the velocity of the gas only on one side of the galaxy. It is useful to measure gas flows, but not to measure the peculiar velocity of the galaxy.

Let us say that an emission line is observed at wavelength $\lambda_{\rm obs}$ (its peak wavelength) instead of its rest-frame value $\lambda^{\rm em}_{\rm rest}$. This means that the average velocity of emitting gas, projected along the line-of-sight, is

$$V_{\rm LOS} = \frac{\lambda_{\rm obs} - \lambda^{\rm em}_{\rm rest}}{\lambda^{\rm em}_{\rm rest}} c.$$

Now, let us look at an absorption line with rest-frame wavelength $\lambda^{\rm abs}_{\rm rest}$. Let us call $\lambda_{\rm center}$ the observed wavelength at which its area is separated in two equal parts. The centroid velocity is then defined to be

$$V_{\rm center} = \frac{\lambda_{\rm center} - \lambda_{\rm abs}}{\lambda_{\rm abs}} c + V_{\rm LOS}. \tag{6.3}$$

The $V_{\rm LOS}$ is a term to compensate the peculiar projected velocity of the galaxy, and the first term is the shift between the rest-frame absorption wavelength and the wavelength of the center of the line, which is due to the movement of the absorbing gas.

However, in Chapter 4, I did not use any emission line to determine $V_{\rm LOS}$ and the centroid velocity of absorption lines. Rather, I used the velocity of the center of mass of the galaxy directly, since it is known from the simulation. How do those two approaches compare? In other words, does the velocity shift of an emission line, for example Hα, accurately reflect the velocity of the center of mass of a galaxy?

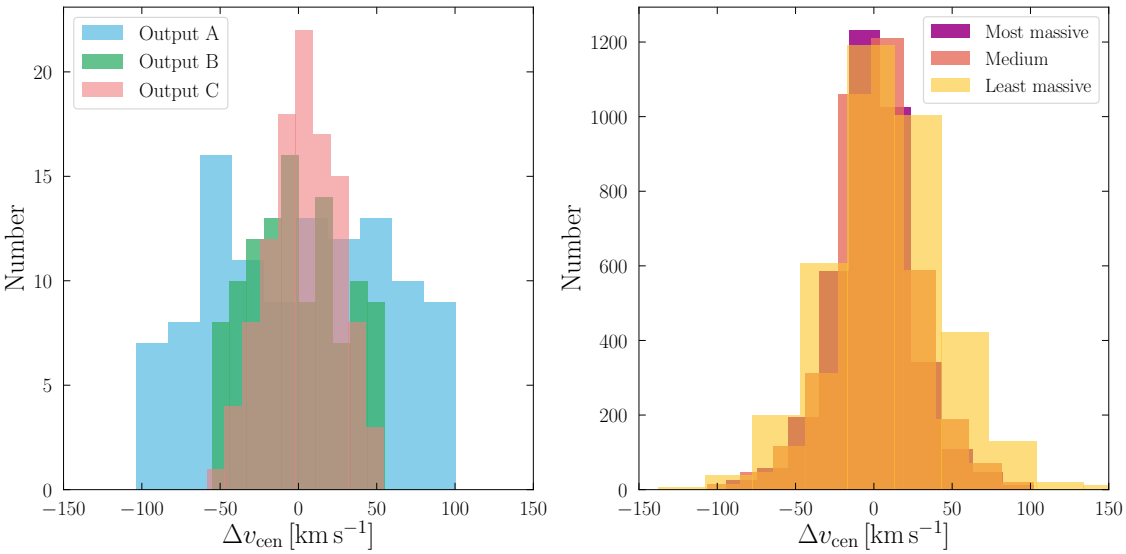

Figure 6.5: Histogram of the differences of centroid velocity between the method used in Mauerhofer et al. (2021), using the center of mass of the halos from the simulations, and the method using the Hα line, both described in this section. Left panel: The three colours correspond to three snapshots of the zoom simulation, the same as in Chapter 4. Each snapshot has 108 directions of observations. Right panel: Each colour contains 33 Halos from the usual snapshot of Sphinx20 at redshift 5.11, sorted by mass. The purple histogram contains Halos ranging from 1.9×10^{10} M$_\odot$ to 2.7×10^{10} M$_\odot$, the orange histogram from 2.7×10^{10} M$_\odot$ to 3.8×10^{10} M$_\odot$ and the yellow histogram from 3.1×10^{10} M$_\odot$ to 5.1×10^{11} M$_\odot$. Each halo is observed from 110 directions of observations, making a total of 3,666 data points per histogram.

To answer these questions, I use new mock Hα emission lines from the zoom and the Sphinx20 simulations. The procedure to produce mock Balmer emission lines was already introduced in Chapter 5. In summary, the emissivity of each gas cell is computed following Cantalupo et al. (2008) and Osterbrock & Ferland (2006). A number of photon packets are then distributed to each gas cell following Michel-Dansac et al. (2020), and their transfer in the galaxy is performed with Rascas, including absorption and scattering by dust grains. The peeling-off algorithm, described in Chapter 3, creates mock observations of the emission lines from multiple directions of observations. The differences between $V_{\rm center}$ obtained using Hα and using the velocity of the center of mass directly are plotted in Figure 6.5. We see that the difference can be large, up to 100 km s^{-1}. This highlights that the wavelength of the Hα line does not depend only on the velocity of the galaxy, but also on the gas movements, which change when looking from different directions. On the right panel, the results for Sphinx20 are separated by mass bins, but no significant differences can be noticed from the different histograms. On the left panel, the output A of the zoom simulation has a wider distribution than the other outputs. This gives us a hint to understanding the origin of the difference between the velocity of the center of mass and the velocity of Hα. Output A of the zoom simulation is dominated by one bright region. For stellar emission, this can be seen in Figure 3.13, which shows mock UV images of this output in several directions of observations. For Hα, the emission is also more localised in output A than in the outputs B and C. Since this site of emission is rotating, the wavelength of the Hα line depends on the angle between the velocity of the emission site and the line-of-sight. Dust also plays a role here, a part of the emission of Hα

is absorbed by dust, in a non-symmetric way.

This shows that there are uncertainties in the centroid velocity when it is measured based on the Hα wavelength. This has a limited impact for galaxies with a centroid velocity of several hundred km s^{-1}, but can change an interpretation from outflow to inflow if it is relatively close to 0, between -100 km s^{-1} and 100 km s^{-1}. The complexity of emission lines has been studied recently in Matthee et al. (2021), where the authors discuss Hα and other emission lines, showing that high resolution spectra contain emission from several HII-regions with different velocities, which complicates the measurement of the systemic redshift.

Using mock Balmer emission lines to determine $V_{\rm center}$ in simulations would be the most faithful way of reproducing observations, but I did not compute those mocks for all snapshots of the zoom simulation. Therefore, I keep using $V_{\rm center}$ based on the velocity of the center of mass. Additionally, using the velocity of the center of mass results in a centroid velocity with more physical meaning: the average velocity difference between the absorbing gas and the global velocity of the halo. In any case, using one method or the other would not affect the results of this section significantly, since I study angle-averaged centroid velocities (average of 108 directions of observations), and we see in Figure 6.5 that the histograms are centered on 0. This implies that, despite the dispersion in the figure, the two methods to derive $V_{\rm center}$ give on average the same results.

6.2.2 Comparing centroid velocities and gas flows

I now compare the centroid velocity of the Fe II λ2586 absorption line with the gas flows in the zoom simulation. This is a first step, where I compare only angle-averaged quantities. In future work I will look more specifically at selected directions of observations, to assess whether the centroid velocity of the line corresponds well with the gas velocity along the line-of-sight. It is however a complicated task, since it is not trivial to define the "velocity of the gas". At which radius and with which opening angle? Also, only the gas that is moving in front of stars, and not hidden by dust, can be probed by absorption lines, so one would have to take this into account for a more in depth study. This is why for now I restrict the study to angle-average quantities. To do so, for each snapshot of the simulation, I create mock Fe II λ2586 spectra in 108 directions of observations, and average them to get the mean centroid velocity. The gas flows are computed in spherical shells surrounding the galaxy, removing the contributions of satellites galaxies, as described in Section 6.1. In that section, we have also seen that the neutral outflows are almost zero at distances of 0.5 $R_{\rm vir}$ and $R_{\rm vir}$, which is why I focus on the gas flows at 0.2 $R_{\rm vir}$.

In Figure 6.6, we can see the evolution of the average Fe II λ2586 centroid velocity in the zoom simulation in comparison to the H I outflow rate. Until around 1.7 Gyr, the centroid velocity is negative when the outflows are large, and positive when there are no outflows. This is as expected, since the absorption lines get blueshifted when the absorbing gas is moving toward the observer. However it seems that the centroid velocity changes slightly before the outflow rate, which is a sign that the blueshift of the absorption line is not a direct consequence of the outflow of neutral gas, but rather an indirect correlation. Some data is missing between redshifts 3.7 and 3.3, as it was computationally expensive to get the Fe II λ2586 spectra for each snapshot of the simulation. The data is present for the last snapshots, starting at redshift 3.3. There, the centroid velocity is almost constant.

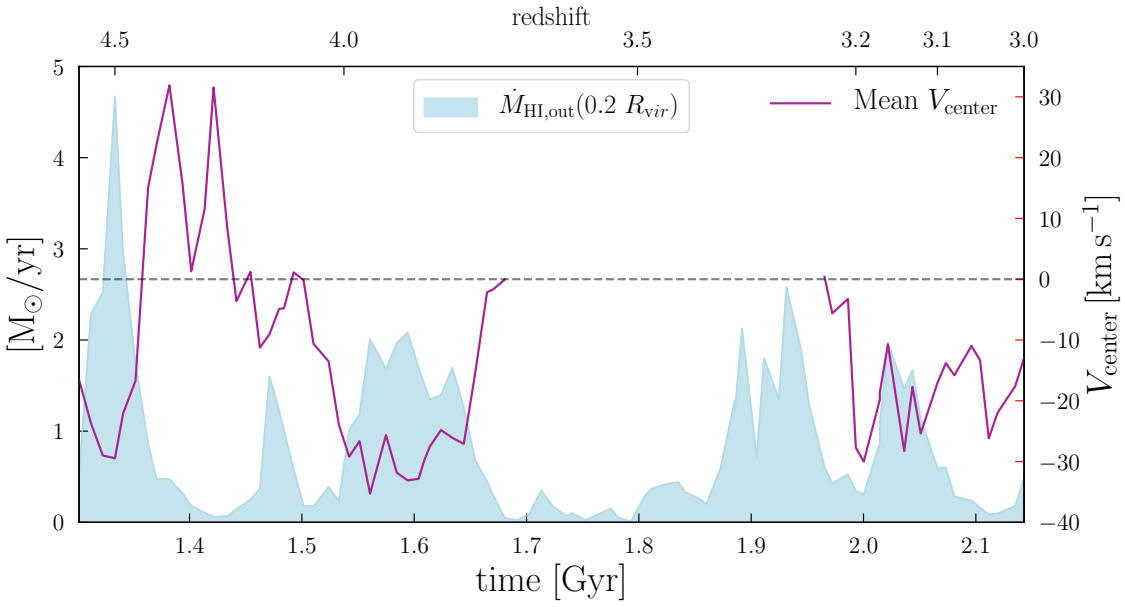

Figure 6.6: Evolution of the neutral hydrogen outflow rate at $0.2\,R_{\rm vir}$ in blue, along with the mean centroid velocity of Fe II λ2586 in magenta. Data taken from the zoom simulation, from redshift 4.6 to 3.0. The centroid velocities are averaged over 108 directions of observations for each snapshot. Fe II λ2586 data is missing from around 1.7 Gyr to 1.95 Gyr. The horizontal dashed line marks $V_{\rm center} = 0$, to guide the eye.

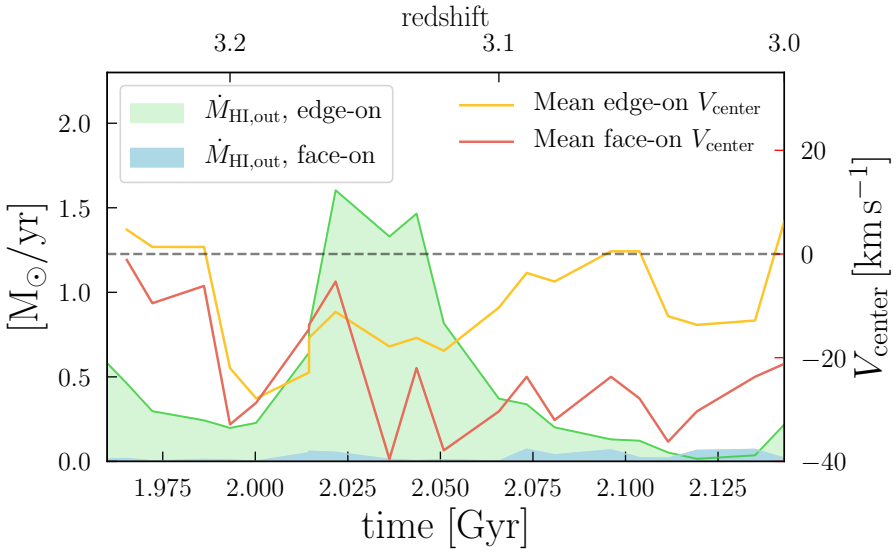

Figure 6.7: The filled curves show the late-time evolution of the neutral hydrogen outflows at $0.2\,R_{\rm vir}$, in edge-on directions of observations for the green and face-on directions for the blue. The solid lines show the centroid velocity of Fe II λ2586, in edge-on directions for the yellow and face-on for the red. Face-on directions are defined as being in a double-cone with an opening angle of 45 degrees around the axis of rotation of the galaxy. Edge-on directions are defined as having an angle of more than 70 degrees from the axis of rotation.

As I stated above, the galaxy of the zoom simulation begins to form a disk from around redshift 3.3, which can be seen through the stabilisation of the angular momentum in Figure 2.6. For lower redshifts, it is therefore possible to compare centroid velocities and gas flows in face-on and edge-on directions. For the face-on direction, I compute gas flows in a double-cone with an opening

angle of 45 degrees around the axis of rotation of the galaxy, and average the centroid velocity of Fe II λ2586 of around 30 directions of observations that are in this double-cone, per snapshot. For the edge-on direction, I do the same by taking the gas that is at an angle of more than 70 degrees from the axis of rotation. The results are shown in Figure 6.7. The data is limited, and more work needs to be done for better interpretation. However, two interesting properties can be noticed here. First, there is no neutral outflow at $0.2\,R_{\rm vir}$ in the face-on direction. This is contrary to common expectations (e.g. Schroetter et al., 2019) that outflows are stronger face-on. This is in part due to the simulation, which ionises all outflowing gas. There are intense face-on outflows around this galaxy, but they are made of extremely hot and low-density gas. This is a common issue in simulations (see also Mitchell et al., 2021), and is due to the feedback modelling. It might be puzzling that there is no outflow even though the face-on centroid velocity is negative, indicating an outflowing absorbing gas. This can be explained by the fact that the absorption actually happens at smaller distances than $0.2\,R_{\rm vir}$ (6-7 kpc), as I show in Section 6.3.

The second thing to notice in Figure 6.7 is that the centroid velocity is more negative for face-on directions than for edge-on directions. This is commonly interpreted (e.g. Martin et al., 2012) as gas falling onto the edges of galactic disks, which tends to redshift absorption lines seen from those directions, while in face-on directions there is more outflowing gas. It is not easy to disentangle the different possible mechanisms at play here and more work is necessary to come to definite conclusions.

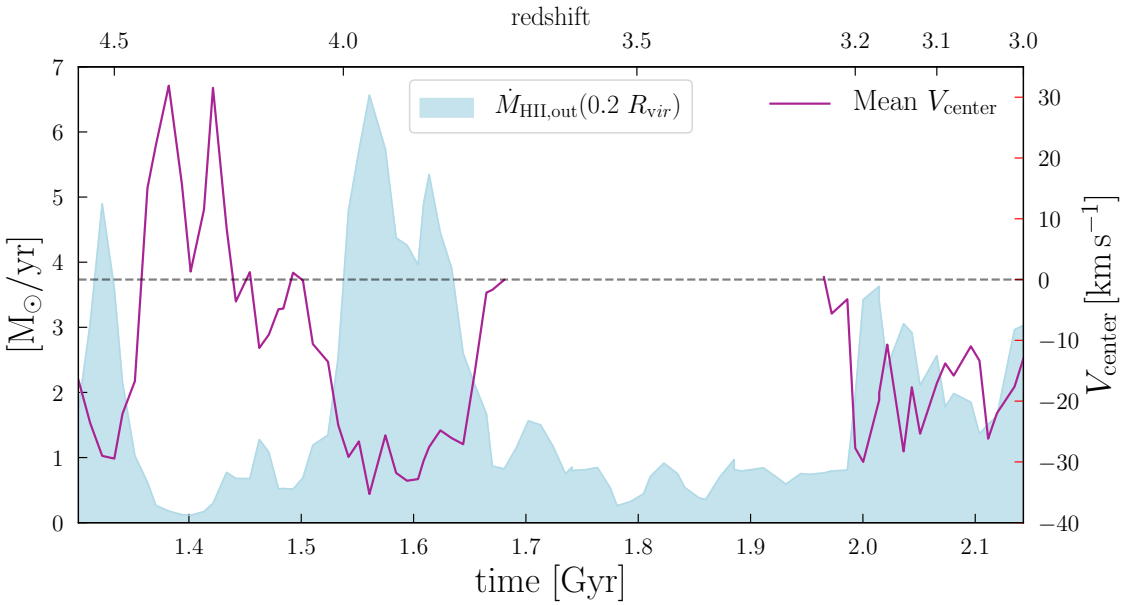

Figure 6.8: Same as Figure 6.6 but with ionised outflows instead of neutral outflows.

In Figure 6.8, I show the comparison between the centroid velocity and the *ionised* outflows. They show the same behaviour as neutral outflows. Globally the centroid velocity of the line is larger when outflows are weak and smaller (more negative) when there are bursts of outflows. This might be surprising because Fe^+ is a low-ionisation element, so it is not expected to trace ionised gas. However, I show in Figure 18 of Mauerhofer et al. (2021) (also in Chapter 4) that low-ionisation state ions can be present in both neutral and ionised (hydrogen) gas. Additionally, neutral and ionised outflows are correlated, one often happens with the other, which also explains the similarities between Figure 6.6 and Figure 6.8.

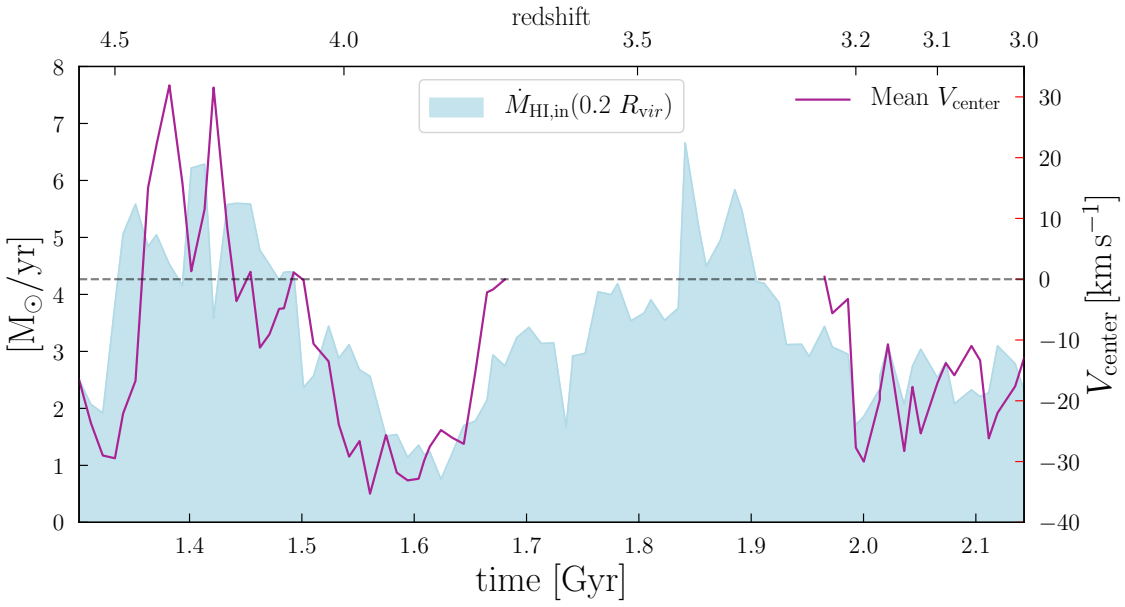

Figure 6.9: Same as Figure 6.6 but with neutral inflows instead of neutral outflows.

Finally, let us look at the comparison between inflows of neutral gas and the centroid velocity of Fe II λ2586. We see in Figure 6.9 that $V_{\rm center}$ and the inflow rate are correlated. Periods of strong inflows are accompanied by positive centroid velocities.

There is still a big step from here to reliably measuring gas flows with absorption lines, as what we show here is merely that there exists a correlation. More work has to be done on more simulated galaxies, for example the ones in Sphinx20, and also from other simulations using different feedback recipes, to test the robustness of those results by varying the subgrid models. Additionally, we saw extensively in the two previous chapters that dust has subtle and complex effects, hiding part of what is happening in the galaxy, which alters the measurements of gas flows.

6.3 Is the gas in the ISM or in the CGM responsible for absorption lines?

To better understand the link between absorption lines and gas flows, one has to know where the absorption is produced. There is a debate on whether the down-the-barrel absorption lines are produced in the ISM or the CGM. The ISM scenario is favored in Chisholm et al. (2017), while e.g. Carr et al. (2018); Schroetter et al. (2019); Prusinski et al. (2021) argue for models where the galactic outflows in the CGM shape the absorption lines. To address this topic, I developed a new technique and implemented it as a new feature in Rascas, to stop the propagation of photons after they reach a given distance from their emission point. This is useful to be able to compute the evolution of the line properties, such as equivalent width, with the radius of propagation. In other words, to see which part of the galaxy contributes the most to the formation of absorption lines. This required a change in the code concerning the photon propagation: In Rascas, photon packets travel in a given domain of propagation, as explained in Section 3.7. Usually, each photon packet has the same domain, also called the computational domain of the Rascas run. To stop each packet after a given distance, I added a new parameter, the radius of propagation $r_{\rm prop}$. Then, I modified the part of the code that assigns the computation domain to each photon packet, to replace it with an assignment to a different domain of propagation, centered on the emission point of the photon packet, and with a radius equal to $r_{\rm prop}$. The photon packet will then stop travelling as soon as it reaches this radius, in the same way as when photon packets reach the virial radius in normal

Rascas runs.

6.3.1 LIS absorption lines

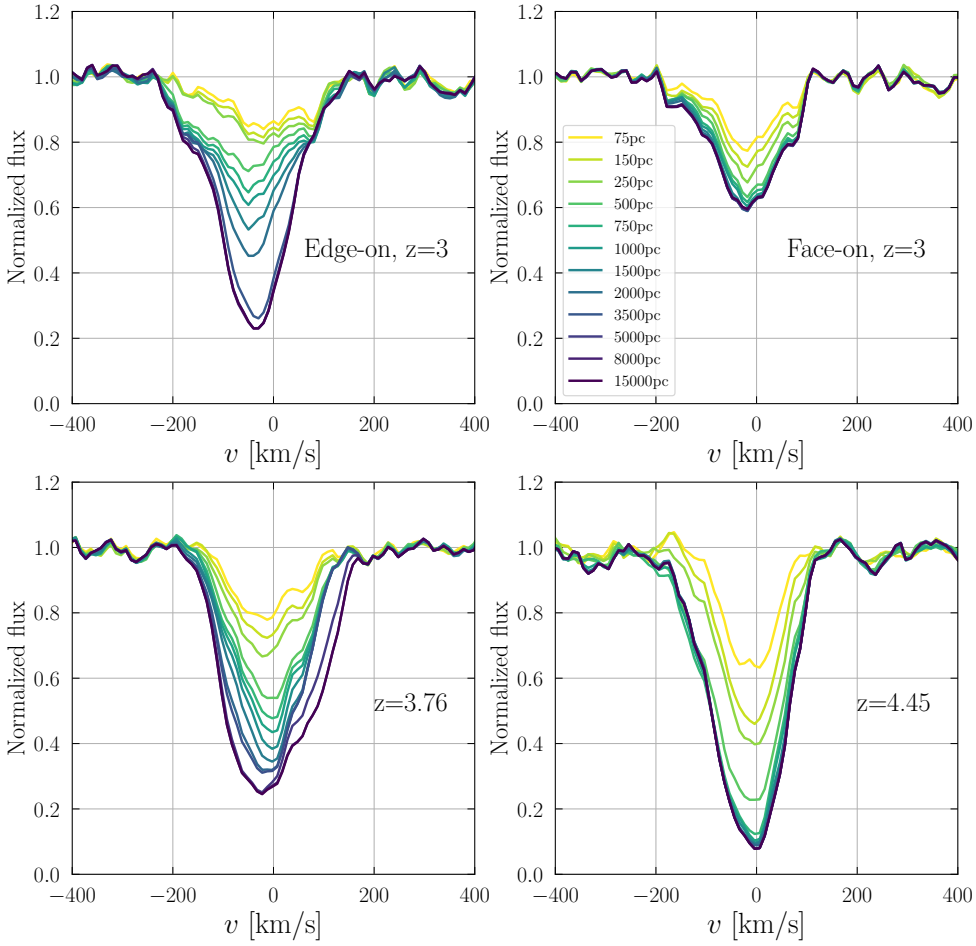

Figure 6.10: Fe II λ2586 spectra of the galaxy from the zoom simulation, including the evolution of the line with radius. The radius of propagation of photons around their source is increased from 75 pc to the full virial radius, as in fiducial spectra. The top two panels show spectra of the last snapshot, at redshift 3; the left one with a direction of observation parallel to the disk and the right one perpendicular. The bottom two panels are spectra at higher redshifts. The bottom left is chosen to have an equivalent width and centroid velocity as close to the median values as possible. The bottom right one was selected randomly over 71 snapshots with 110 directions of observations for each.

As a first visual way of learning where the absorption is produced, I show in Figure 6.10 the evolution of the Fe II λ2586 profile when increasing the radius of propagation of photons around their sources, for four mock observations. The top two panels are spectra from the last snapshot of the zoom simulation, at redshift 3, at an edge-on angle in the left panel and face-on angle on the right. It is striking that the edge-on absorption line is wider and deeper than the face-on one. Concerning the evolution with radius, both top panels show a steady increase of equivalent width until a distance of around 500 pc, after which the face-on direction of observation seems to almost stop evolving, while the edge-on continues to grow. This can be explained by the fact that the width of the ISM disk (around 750 pc) is smaller than its radius (around 2.5 kpc), therefore the dense gas reaches less far in the face-on direction. The edge-on mock Fe II λ2586 absorption

line, in the top left panel of Figure 6.10, stops growing after a radius of propagation of 8 kpc. This distance is large enough to call this region the CGM, but is also much smaller than the virial radius (29 kpc). Therefore, in the zoom simulation, absorption lines are not affected by gas at the scale of the virial radius.

The two bottom panels of Figure 6.10 show spectra at higher redshifts. In those snapshots the galaxy has not formed a disk-like structure yet, hence the concepts of face-on and edge-on do not apply. The bottom left panel is a spectrum with an equivalent width and a centroid velocity equal to the median values of those quantities over all snapshots above redshift 3.5 and all directions of observations. The centroid velocity is blueshifted, but only slightly, in contrast to many studies that find significantly blueshifted absorption lines at high redshift (e.g. Shapley et al., 2003; Erb, 2015; Rivera-Thorsen et al., 2015; Trainor et al., 2016). Like for the edge-on spectrum in the top left panel, the equivalent width evolves regularly until a radius of 8 kpc. The bottom right panel was done by randomly selecting a snapshot and angle of observation. Here the equivalent width stops evolving after only 750 pc, due to the smaller size of the galaxy at this redshift ($R_{\rm vir}$ = 19 kpc). The equivalent width is however quite large, slightly larger than the top and bottom left panels. This shows that in general the size of a galaxy does not impact the equivalent width of absorption, which is also discussed in Prusinski et al. (2021) or Du et al. (2021). On the contrary, small galaxies might have a more uniform coverage of the stars by gas, which produces stronger absorption lines (this does not apply when the galaxy is so small and star-forming that intense feedback ejects all the gas, leaving no absorption lines).

To have a more statistical view of the evolution of the equivalent width with radius, I plot in Figure 6.11 three curves of evolution of the equivalent width in the zoom simulation. These curves are made from mock C II λ1334 of the galaxy from the usual zoom simulation, as I do not yet have the data for the Fe II λ2586 line, which I have been using in this section. However, those lines are both from LIS ions, so their behaviour is similar, and one would get the same qualitative results with Fe II λ2586 as with C II λ1334. Each curve of Figure 6.11 is a median of hundreds of curves, each obtained by making mock C II λ1334 observations with different radii of propagation for the photon packets. The blue curve is made from face-on mock observations of the three usual snapshots of the zoom simulation, the same as in Chapter 4. For each snapshot, I make 10 mock observations from different directions of observation, all within an angle of 30 degrees from the axis of rotation of the galaxy. The blue curves show that the face-on absorption lines grow relatively fast, reaching half their equivalent width at around 300 pc. They reach 90% of their equivalent width at around 1500 pc, which is larger than the 500 pc of the top right panel of Figure 6.10, which shows that this panel does not represent the statistical average. The yellow curve of Figure 6.11 is made with edge-on directions of observations of the same three snapshots, i.e. it is the median of observations from an angle of more than 60 degrees from the axis of rotation. The evolution is slower than for the face-on case, since it reaches half its equivalent width at around 900 pc, and 90% of its equivalent width at around 3 kpc. Finally, the green curve of Figure 6.11 is made from more than four thousand mock observations of C II λ1334 for snapshots of the zoom simulations at redshifts above 3.75, when the galaxy had no disk structure. Interestingly, the curve falls in between the two others, with a radius of around 450 pc where the line reaches half its equivalent width and 2 kpc where it reaches 90% of its equivalent width.

In summary, the entirety of the area of absorption lines in our simulation is produced within a radius of 8 kpc. Both the ISM and the CGM contribute to the formation of LIS lines, with a slightly larger contribution from the ISM, and no contribution from the distant CGM, at scales of around the virial radius. For snapshots where the galaxy has formed a disk, the absorption lines seen from an edge-on direction keep evolving on a longer distance, due to the presence of more gas parallel to the disk than perpendicular.

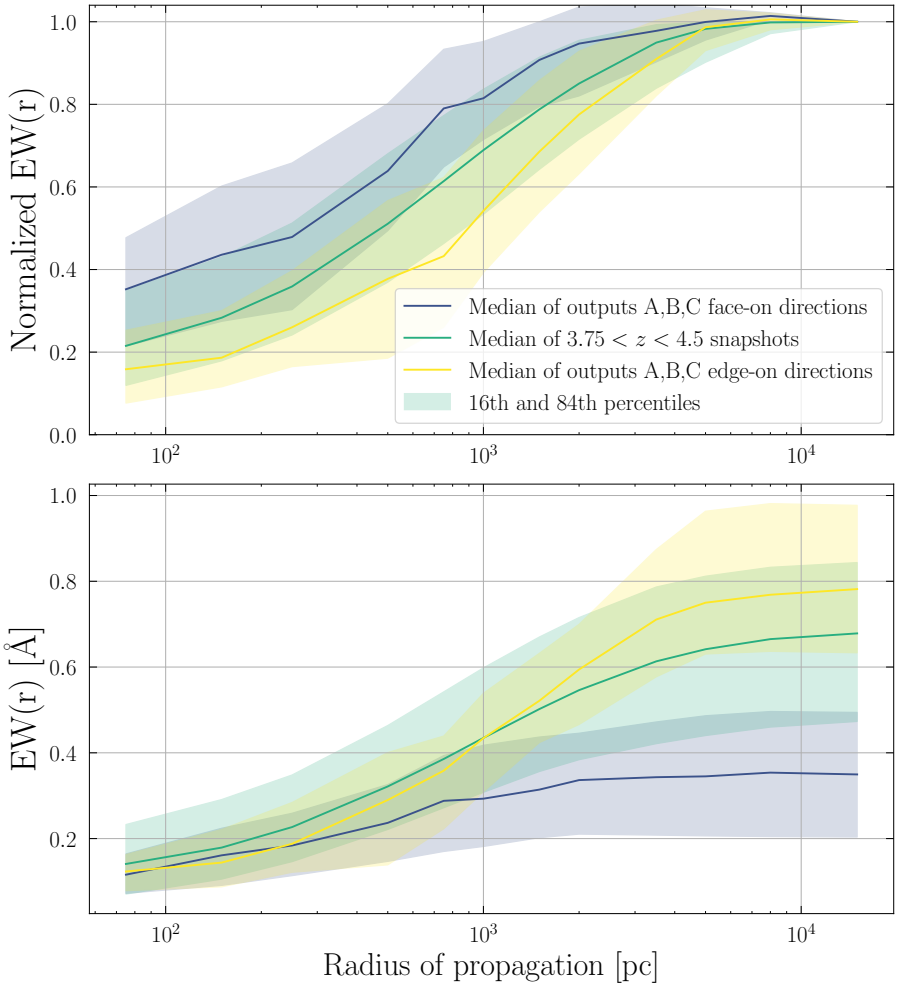

Figure 6.11: Evolution of the EW of C II λ1334 as a function of radius of propagation from the sources. In the upper panel, the curves are normalised to the equivalent width at 15 kpc, while they are not normalised in the lower panel. The dark-blue curve shows the median from the face-on mock absorption lines of the same three snapshots of the zoom simulation as in Chapter 4. The green curve is the median of 108 directions of observations from 41 snapshots between redshifts 3.75 and 4.5, and the yellow curve is the median from edge-on directions of the three usual snapshots. The shaded regions show the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentile of the equivalent width distributions.

6.3.2 High ionisation state absorption lines

Until now we have focused more on LIS absorption lines, which are the most important in terms of escape of ionising photons. However it is worth mentioning that high ionisation state lines have an intriguing behaviour in our simulations. The C IV $\lambda\lambda$1548, 1550 line is a doublet that is purely resonant, i.e. all absorption events result in the reemission of the absorbed photon, which differs from many LIS absorption lines that can have fluorescent decay. This leads to strong infilling effects, and can even transform the absorption line into an emission line in some directions of observation.

The transfer of C IV $\lambda\lambda$1548, 1550 is done in the same way as other absorption lines presented in this work, following Chapter 3. However, stellar spectra have a C IV $\lambda\lambda$1548, 1550 absorption feature, which we could in principle include in the modelling of the line, in the initial conditions of the photon packets. But since we are interested in properties of absorption by ISM and CGM gas, I smooth the stellar absorption features by fitting a power-law to the intrinsic spectra. Furthermore, another reason to do this is that stellar models have different predictions for the C IV $\lambda\lambda$1548, 1550

absorption features, as it is shown in Figure 6.12.

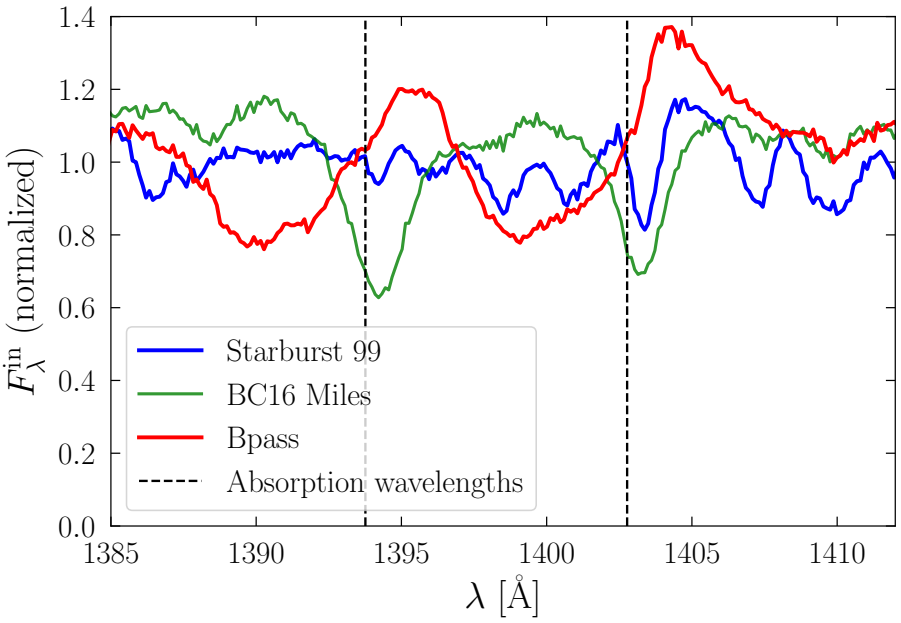

Figure 6.12: Intrinsic stellar continuum emission around the wavelength of the Si IV $\lambda\lambda$1394, 1403 doublet, according to different stellar libraries. The Starburst 99 models are from Leitherer et al. (2014), the BC16 models from Vazdekis et al. (2016) and the BPASS models from Eldridge et al. (2008); Stanway et al. (2016). The vertical dashed lines highlight the intrinsic wavelength of the absorption.

The curves of evolution of equivalent width of C IV $\lambda\lambda$1548, 1550 as a function of the maximum radius of propagation, made in the same way as for Figure 6.11, are shown in Figure 6.13. The behaviour of face-on directions are similar for C IV $\lambda\lambda$1548, 1550 as for C II λ1334. Half the total equivalent width is reached at around 600 pc, and 90% at around 2 kpc. However, the edge-on directions are surprising. There are mocks for which the equivalent width is reduced when letting photons propagate further, due to the infilling effects of resonance. So, when normalising to the observed equivalent width (when stopping photons only at the virial radius), the curve goes above unity, and sometimes much higher.

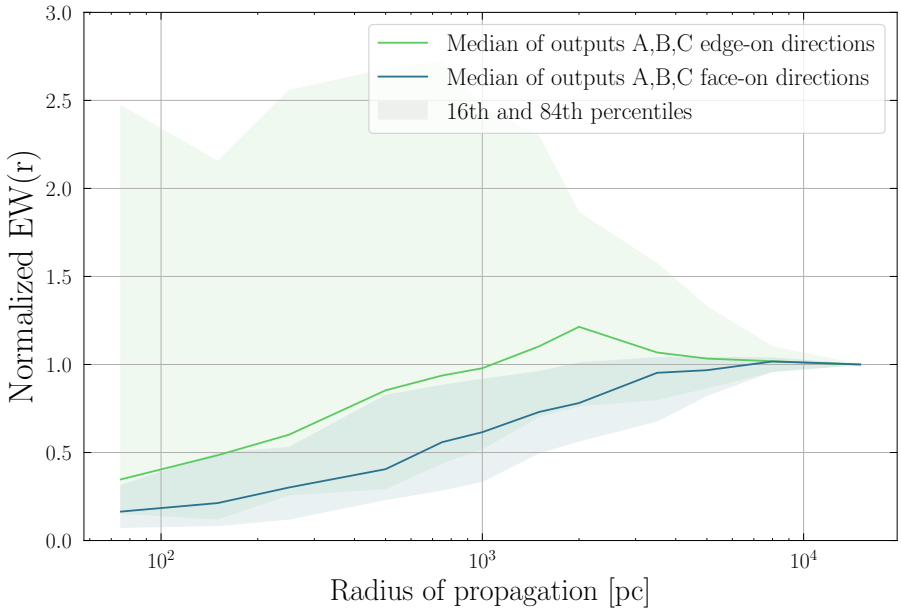

Figure 6.13: Same as Figure 6.11 but for C IV $\lambda\lambda$1548, 1550.

6.4 A lack of observed redshifted absorption lines: Are simulations missing neutral outflows?

We have seen with all the figures in this chapter that, even though the centroid velocities of our absorption lines are on average negative, hinting at outflowing gas, the value of this velocity is still low compared to observations (e.g. Shapley et al., 2003; Erb, 2015; Rivera-Thorsen et al., 2015; Trainor et al., 2016). It seems that there is not enough neutral outflowing gas in our simulated galaxy. In addition, a substantial fraction of my mock absorption lines display a redshifted absorption, while that is rarely observed. I discuss here whether the simulations are failing to reproduce outflows, and whether observations miss a population of galaxies with redshifted absorption lines.

I have found that mock observations with redshifted absorption lines tend to have lower magnitudes. To prove that for the case of the zoom simulation, I show in Figure 6.14 the median values of centroid velocity of the Fe II λ2586 in bins of FUV magnitude. There are three curves, to separate snapshots at low and high redshift, which are all similar, showing that there is no evolution of this trend with redshift. We see that there is indeed a correlation between magnitudes and centroid velocities, as the brightest observations have the most blueshifted absorption, i.e. the most outflows, and vice versa. This suggests that high redshift galaxy surveys, which are often magnitude limited, i.e. they detect the continuum and absorption lines only of the brightest galaxies, might miss a population of galaxies with gas inflows, displaying redshifted absorption lines.

In order to determine whether this correlation also arises for a galaxy at one given time, when looking at different directions of observation, I plot in Figure 6.15 the same data as in the previous figure but with the magnitude normalised to the average magnitude of the snapshot at which the mocks have been made. Except for small dips at the extremities, where there are only a fraction of percent of the data points, and which thus suffers from low number statistics, the curve is essentially flat, contrary to Figure 6.14. This means that the correlation between the magnitude and the centroid velocity is not due to variations of direction of observation, but rather to variations in time, from snapshot to snapshot.

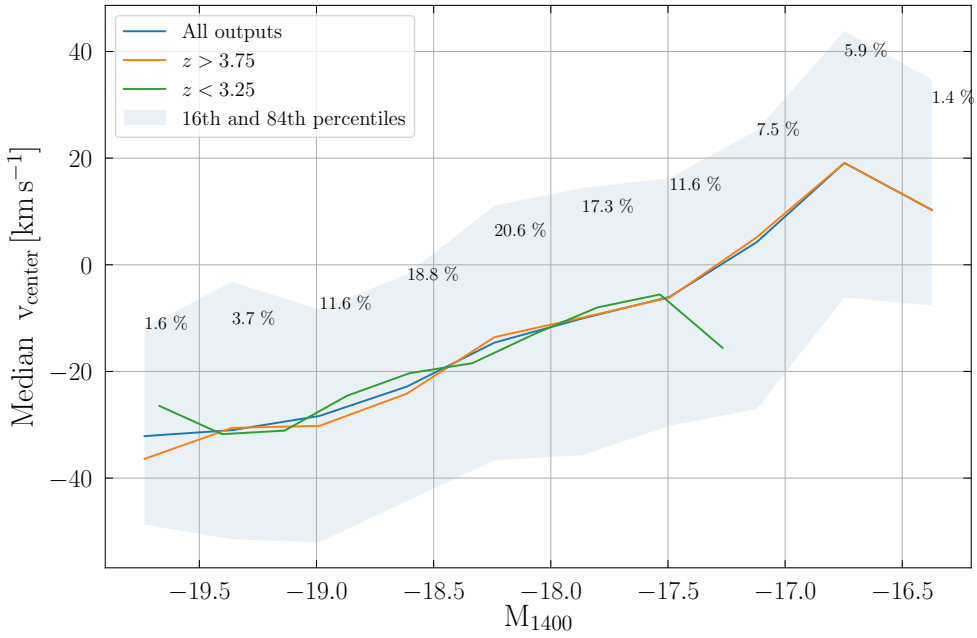

Figure 6.14: Median of the Fe II λ2586 centroid velocity in different bins of FUV magnitude. The data comes from 41 snapshots of the zoom simulation, at redshift above 3.75, along with 20 snapshots at redshift below 3.25, in red and green, respectively, each seen from 110 directions of observations. The FUV magnitude is computed from the flux at 1400 Å. The percentages indicated above the curves are the fraction of mock observations that fall in the corresponding magnitude bins, among the 6710 total mock observations.

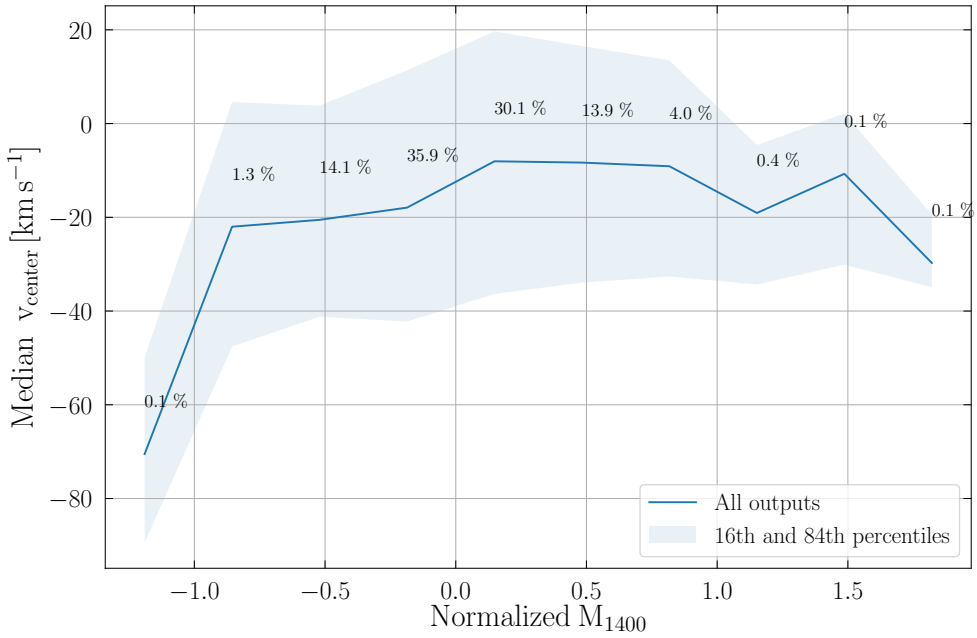

Figure 6.15: Same as Figure 6.14 but for each mock observation, I subtract the angle-averaged magnitude of the galaxy at the corresponding snapshot.

The results from the zoom simulation are shown at different times and angles for a single galaxy. This raises the question of whether there is also a bias in this kind of study, compared to looking at a collection of galaxies with different properties. To answer this, I started analysing the Sphinx20

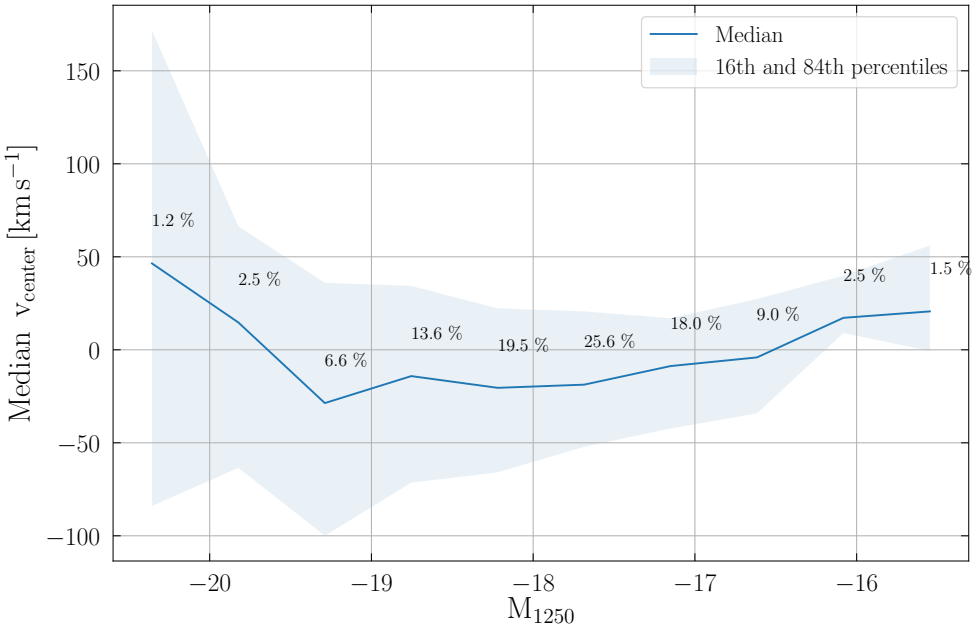

Figure 6.16: Same as Figure 6.14 but for the 100 most massive halos of Sphinx20 at redshift 5.11 instead of the zoom simulation.

simulation. The first result I got was the same as Figure 6.14 but for the 100 most massive halos of Sphinx20, which I show in Figure 6.16. When looking at the same range of FUV magnitude as for the zoom simulation, the results are similar, only with a wider spread. However for the massive galaxies of Sphinx20, which constitute the great advantage of this simulation, the trend disappears. We see that those bright galaxies experience intense inflows of gas, although the curve is extremely spread and some of them also have strong outflows. On the one hand it seems natural to have strong inflows at high redshift, since the galaxies are growing intensely. On the other hand, a wide range of observations show that absorption lines are most commonly blueshifted, indicating that there must be intense neutral gas movements in the ISM, going out of the galaxy. The feedback model from our Ramses-RT simulations tends to ionise all the gas that is outflowing, hence not reproducing enough blueshifted absorption.

7 Conclusion and Outlook

This thesis gave insights into the physical processes that shape down-the-barrel absorption lines and the escape fractions of ionising photons. The main goal was to develop a method that produces accurate absorption lines from radiation-hydrodynamic simulations, and to study the links between the line properties and the escape of Lyman continuum. This provided clues on how to determine the history of Reionisation and identify Lyman Continuum leaking galaxies

7.1 Summary

The first step of the thesis was to present the simulation code and the resulting RHD simulations that are used in this thesis, in Chapter 2. Those simulations have a high-resolution and include the out-of-equilibrium photo-chemistry of the gas, star-formation as well as supernova and radiative feedback. The first one is a zoom-in simulation, following a galaxy of LAE size evolving until redshift 3. The second simulation (Sphinx20) is a large cosmological volume containing thousands of galaxies, of which I study the 100 most massive at redshift 5.11.

In Chapter 3, I explained my method to post-process Ramses-RT simulations in order to make mock observations of metallic absorption lines. The first sections are about the extension of the simulation data, adding a new energy bin of radiation, the opacity of dust, the density of heavy elements, their ionisation fractions and the population fractions of the fine-structure levels of the ground state. Then, I described how to model the emission of photons from the stellar particles using SED libraries. The next part was to explain the functioning of radiative transfer code Rascas, and how I expanded it to produce mock absorption lines from different directions of observation. Finally, I gave an example of a metallic absorption line, Si II λ1260, from a galaxy of Sphinx20, and described its different properties such as the centroid velocity and the residual flux.

Chapter 4 contains the main scientific results of the thesis, relating absorption lines and the escape fraction of ionising photons. It also contains an extensive analysis of the different processes affecting down-the-barrel absorption lines. The presence of inhomogeneous dust and the large scale velocity gradients of the absorbing gas are what affect the lines the most. One of the main results is the variability of observables when observing the galaxy from different angles. Both the escape fraction of ionising photons and the shapes of the line change significantly. Then, I explored the correlations between the residual flux of LIS absorption lines with the escape fraction, and showed that without dust absorption they correlate poorly. When correcting the residual flux of metallic LIS lines by the attenuation of the UV continuum, the correlation is better, and low values of such dust corrected residual fluxes indicate a high escape fraction, in the zoom simulation. There is no correlation when using absorption lines of the Lyman series. In the last section, I explained these results, by invoking the complex distribution of LIS elements with respect to H^0, and the fact that the gas column densities in front of the sources are not separated in two distinct groups, optically thin and optically thick, thus making the concept of covering fraction inapplicable. This also resulted in the fact that the residual flux of absorption lines is not equal to the covering fraction of neutral hydrogen. Finally, I showed in an addendum that, when applying the same developments to a statistical sample of galaxies, from the Sphinx20 simulation, the residual

flux of LIS absorption lines does not correlate with the escape fraction of ionising photons, even with dust correction. This suggests that in general LIS absorption lines cannot predict the escape of Lyman continuum.

Dust is the ingredient that affects UV spectra the most, which is why I studied the dust properties of our simulations in more detail in Chapter 5. I began by giving the details of the modelling, showing that star-forming regions are dusty and hide a significant fraction of the UV light from young stars. In a following section, I assessed the realism of the UV properties of the simulations, finding that they match well with observations, in terms of the luminosity function at 1500 Å, the β slope and the infrared excess, although the latter is larger than in some observational studies. I continued by testing observational methods of deriving the dust attenuation of galaxies, finding that those methods underestimate the dust extinction.

I studied the gas flow properties of our simulated galaxies in Chapter 6. The first section was about the definition of mass outflow and inflow rates, as well as the comparison of those quantities at different radii around the galaxy from the zoom simulation. I showed that there are almost no neutral outflows at 50% of the virial radius and above. However, the ionised outflows are intense, due to our feedback modelling which tends to ionise the gas accelerated by supernovae. In the second section, I discussed the notion of centroid velocity of absorption lines and how to measure it. Then I showed that there are correlations between the gas flows and the centroid velocity, even when taking an angle-average of those quantities. In the next section, I assessed at which spatial scales the absorption lines are formed in the zoom simulation. I found that the ISM affects the equivalent width the most, although the CGM contributes by about a third. In our simulations, the absorption lines are not affected by gas further than 8 kpc from the sources. Finally, I studied whether galaxies with redshifted absorption lines (indicating inflows of gas) are fainter in the UV, which would explain why few of them are detected in observations. I found that, in the zoom simulation, directions of observations with redshifted absorption lines are indeed fainter than ones with blueshifted absorption lines. However, this does not hold when looking at the galaxies from the Sphinx20 simulation, leaving the question unanswered.

7.2 Outlook

Here I give an overview of the future work that I plan to carry out in order to extend the present thesis, and of the improvements of the techniques and simulations that I used.

Simulations

Thanks to the analysis I have conducted, there are now new indications on how to design future simulations that will optimise the post-processing and yield even more realistic galaxies. First, it is easy to reduce the number of free parameters in the post-processing pipeline by adding new ingredients in the simulations, without significantly increasing their computational cost. By adding an additional bin of radiation between around 6 eV and 13.6 eV, one would not need to add this bin in post-processing, making it much more accurate. Additionally, such a non-ionising bin is also important for a more precise determination of the ISM temperature (e.g. Pallottini et al., 2019). Including a model for dust creation and destruction is another interesting implementation, as done for example in Trebitsch et al. (2020). This would allow us to have a more realistic distribution of dust, to be compared with the model I used in this thesis. Then, runaway stars have proven to improve several aspects of simulations. The dust attenuation factor of young stars is reduced to more realistic values (Cen & Kimm, 2014) and the galactic outflows are boosted (Andersson et al., 2020), which might alleviate tensions that I discussed in Chapters 5 and 6. Therefore, I plan to include runaway stars in a future zoom simulation.

Post-processing

Regarding the post-processing of simulations, the new implementations mentioned above will make it more simple. One new procedure that I plan to use is the local turbulent velocity used in the computation of the criteria of star formation (see Section 2.1.6). This is directly accessible without new implementations in the simulations, although it requires some developments. This turbulent velocity will then replace the free parameter that I presented in Chapter 4, resulting in a more self-consistent modelling. Finally, I plan to update the peeling-off implementation in Rascas in order to design mock observations with multiple resolutions. The goal is to build data cubes with a higher spatial resolution in selected regions, typically in the ISM of galaxies, and a higher spectral resolution around the wavelength of spectral lines. This will allow us to make mock observations with large wavelength and spatial ranges in only one Rascas run.

Escape of Lyman continuum

In this thesis I have shown and explained why down-the-barrel absorption lines are poor tracers of the escape fraction of ionising photons. It is however possible that connecting several observables together, such as absorption lines, FIR emission lines and Lyman alpha properties, one would recover a more faithful tracer of escape fractions. The links between FIR lines and Lyman continuum escape have been studied in e.g. Katz et al. (2020) or Barrow et al. (2020), while correlations between Lyman alpha and Lyman continuum is presented in Maji et al. in prep. and Chuniaud et al. in prep. In the near future, we plan to assemble data from those papers and apply statistical tools similar as in Maji et al. in prep. to find robust multivariable tracers of the escape of ionising photons.

Gas flows in the CGM

To better understand the links between absorption lines and outflows of gas, I will pursue the work of Chapter 6 by computing the gas flows in selected directions of observations, and compare them with observations from the same directions. This will be more closely related to what is done in observations. Additionally, I will study in more detail the gas flows around the galaxies of Sphinx20, and of the future state-of-the art simulations I plan to perform.

8 Bibliography

Álvarez-Márquez, J., Burgarella, D., Buat, V., Ilbert, O., & Pérez-González, P. G. 2019, *A*stronomy and Astrophysics, 630, A153

Andersson, E. P., Agertz, O., & Renaud, F. 2020, *M*onthly Notices of the RAS, 494, 3328

Andersson, E. P., Renaud, F., & Agertz, O. 2021, *M*onthly Notices of the RAS, 502, L29

Aravena, M., Decarli, R., Walter, F., et al. 2016, *A*strophysical Journal, 833, 68

Asplund, M., Amarsi, A. M., & Grevesse, N. 2021, arXiv e-prints, arXiv:2105.01661

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A*stronomy and Astrophysics, 558, A33

Aubert, D. & Teyssier, R. 2008, *M*onthly Notices of the RAS, 387, 295

Badnell, N. R. 2006, *A*strophysical Journal, Supplement, 167, 334

Bakx, T. J. L. C., Sommovigo, L., Carniani, S., et al. 2021, arXiv e-prints, arXiv:2108.13479

Barlow, M. J. 1978, *M*onthly Notices of the RAS, 183, 367

Barrow, K. S. S., Robertson, B. E., Ellis, R. S., et al. 2020, *A*strophysical Journal, Letters, 902, L39

Barrow, K. S. S., Wise, J. H., Aykutalp, A., et al. 2018, *M*onthly Notices of the RAS, 474, 2617

Barrow, K. S. S., Wise, J. H., Norman, M. L., O'Shea, B. W., & Xu, H. 2017, *M*onthly Notices of the RAS, 469, 4863

Becker, R. H., Fan, X., White, R. L., et al. 2001, *A*stronomical Journal, 122, 2850

Behroozi, P. S. & Silk, J. 2015, *A*strophysical Journal, 799, 32

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *P*hysics Reports, 367, 1

Bode, P., Bahcall, N. A., Ford, E. B., & Ostriker, J. P. 2001, *A*strophysical Journal, 551, 15

Bourne, N., Dunlop, J. S., Merlin, E., et al. 2017, *M*onthly Notices of the RAS, 467, 1360

Bouwens, R., González-López, J., Aravena, M., et al. 2020, *A*strophysical Journal, 902, 112

Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, *A*strophysical Journal, 843, 129

Bouwens, R. J., Smit, R., Labbé, I., et al. 2016, *A*strophysical Journal, 831, 176

Bruzual, G. & Charlot, S. 2003, *M*onthly Notices of the RAS, 344, 1000

Cantalupo, S., Porciani, C., & Lilly, S. J. 2008, *A*strophysical Journal, 672, 48

Carr, C., Scarlata, C., Panagia, N., & Henry, A. 2018, *A*strophysical Journal, 860, 143

Cen, R. 1992, *A*strophysical Journal, Supplement, 78, 341

Cen, R. & Kimm, T. 2014, *A*strophysical Journal, 794, 50

Charlot, S. & Fall, S. M. 2000, *A*strophysical Journal, 539, 718

Chevallard, J. & Charlot, S. 2016, *M*onthly Notices of the RAS, 462, 1415

Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018a, *A*stronomy and Astrophysics, 616, A30

Chisholm, J., Tremonti, C., & Leitherer, C. 2018b, *M*onthly Notices of the RAS, 481, 1690

Chisholm, J., Tremonti, C. A., Leitherer, C., & Chen, Y. 2017, *M*onthly Notices of the RAS, 469, 4831

Chisholm, J., Tremonti, C. A., Leitherer, C., Chen, Y., & Wofford, A. 2016, *M*onthly Notices of the RAS, 457, 3133

Chisholm, J., Tremonti, C. A., Leitherer, C., et al. 2015, *A*strophysical Journal, 811, 149

Conselice, C. J., Wilkinson, A., Duncan, K., & Mortlock, A. 2016, *A*strophysical Journal, 830, 83

Courant, R., Friedrichs, K., & Lewy, H. 1928, Mathematische Annalen, 100, 32
Dainotti, M. G. & Del Vecchio, R. 2017, *N*ew Astronomy Review, 77, 23
De Cia, A. 2018, *A*stronomy and Astrophysics, 613, L2
De Cia, A., Ledoux, C., Mattsson, L., et al. 2016, *A*stronomy and Astrophysics, 596, A97
Dekel, A. & Silk, J. 1986, *A*strophysical Journal, 303, 39
Dere, K. P. 2007, *A*stronomy and Astrophysics, 466, 771
Dijkstra, M. 2017, arXiv e-prints, arXiv:1704.03416
Dijkstra, M. & Loeb, A. 2008, *M*onthly Notices of the RAS, 391, 457
Dopita, M. A. & Sutherland, R. S. 2003, Astrophysics of the diffuse universe
Du, X., Shapley, A. E., Topping, M. W., et al. 2021, arXiv e-prints, arXiv:2103.15824
Dubois, Y. & Teyssier, R. 2008, *A*stronomy and Astrophysics, 477, 79
Duncan, K. & Conselice, C. J. 2015, *M*onthly Notices of the RAS, 451, 2030
Eldridge, J. J., Izzard, R. G., & Tout, C. A. 2008, *M*onthly Notices of the RAS, 384, 1109
Erb, D. K. 2015, *N*ature, 523, 169
Erb, D. K., Quider, A. M., Henry, A. L., & Martin, C. L. 2012, *A*strophysical Journal, 759, 26
Fan, X., Carilli, C. L., & Keating, B. 2006, *A*nnual Review of Astron and Astrophys, 44, 415
Feldmann, R., Hopkins, P. F., Quataert, E., Faucher-Giguère, C.-A., & Kereš, D. 2016, *M*onthly Notices of the RAS, 458, L14
Feltre, A., Bacon, R., Tresse, L., et al. 2018, *A*stronomy and Astrophysics, 617, A62
Feltre, A., Maseda, M. V., Bacon, R., et al. 2020, arXiv e-prints, arXiv:2007.01878
Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *R*evista Mexicana de Astronomia y Astrofisica, 53, 385
Few, C. G., Courty, S., Gibson, B. K., et al. 2012, *M*onthly Notices of the RAS, 424, L11
Fudamoto, Y., Oesch, P. A., Faisst, A., et al. 2020a, *A*stronomy and Astrophysics, 643, A4
Fudamoto, Y., Oesch, P. A., Magnelli, B., et al. 2020b, *M*onthly Notices of the RAS, 491, 4724
Fujita, T., Mauerhofer, V., Senatore, L., Vlah, Z., & Angulo, R. 2020, *J*ournal of Cosmology and Astroparticle Physics, 2020, 009
Gallego, S. G., Cantalupo, S., Sarpas, S., et al. 2021, *M*onthly Notices of the RAS, 504, 16
Gatkine, P., Veilleux, S., & Cucchiara, A. 2019, *A*strophysical Journal, 884, 66
Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A*stronomy and Astrophysics, 639, A85
Gazagnes, S., Chisholm, J., Schaerer, D., et al. 2018, *A*stronomy and Astrophysics, 616, A29
Geen, S., Bieri, R., Rosdahl, J., & de Koter, A. 2021, *M*onthly Notices of the RAS, 501, 1352
Geen, S., Rosdahl, J., Blaizot, J., Devriendt, J., & Slyz, A. 2015, *M*onthly Notices of the RAS, 448, 3248
Giallongo, E., Grazian, A., Fiore, F., et al. 2019, *A*strophysical Journal, 884, 19
Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, *A*stronomy and Astrophysics, 634, A135
Girichidis, P., Naab, T., Walch, S., et al. 2016, *A*strophysical Journal, Letters, 816, L19
Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, *A*strophysical Journal, 672, 765
Godunov, S., Friedrichs, K., & Lewy, H. 1959, Mathematische Annalen, 47, 271
Grassi, T., Bovino, S., Schleicher, D. R. G., et al. 2014, *M*onthly Notices of the RAS, 439, 2386
Grond, J. J., Woods, R. M., Wadsley, J. W., & Couchman, H. M. P. 2019, *M*onthly Notices of the RAS, 485, 3681
Haardt, F. & Madau, P. 2012, *A*strophysical Journal, 746, 125
Hahn, O. & Abel, T. 2013, MUSIC: MUlti-Scale Initial Conditions
Hassan, S., Davé, R., Mitra, S., et al. 2018, *M*onthly Notices of the RAS, 473, 227
Heckman, T. M., Sembach, K. R., Meurer, G. R., et al. 2001, *A*strophysical Journal, 558, 56
Henyey, L. G. & Greenstein, J. L. 1941, *A*strophysical Journal, 93, 70
Hockney, R. W. & Eastwood, J. W. 1981, Computer Simulation Using Particles
Hopkins, P. F. 2015, *M*onthly Notices of the RAS, 450, 53
Hopkins, P. F., Chan, T. K., Squire, J., et al. 2021, *M*onthly Notices of the RAS, 501, 3663

Hubble, E. P. 1925, *A*strophysical Journal, 62, 409
Hubble, E. P. 1953, *M*onthly Notices of the RAS, 113, 658
Hui, L. & Gnedin, N. Y. 1997, *M*onthly Notices of the RAS, 292, 27
Humason, M. L., Mayall, N. U., & Sandage, A. R. 1956, *A*stronomical Journal, 61, 97
Humlícek, J. 1982, *J*ournal of Quantitiative Spectroscopy and Radiative Transfer, 27, 437
Hummels, C. B., Smith, B. D., & Silvia, D. W. 2017, *A*strophysical Journal, 847, 59
Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
Jaskot, A. E., Dowd, T., Oey, M. S., Scarlata, C., & McKinney, J. 2019, *A*strophysical Journal, 885, 96
Jenkins, E. B. 2009, *A*strophysical Journal, 700, 1299
Katsianis, A., Xu, H., Yang, X., et al. 2021, *M*onthly Notices of the RAS, 500, 2036
Katz, H., Galligan, T. P., Kimm, T., et al. 2019, *M*onthly Notices of the RAS, 487, 5902
Katz, H., Kimm, T., Sijacki, D., & Haehnelt, M. G. 2017, *M*onthly Notices of the RAS, 468, 4831
Katz, H., Rosdahl, J., Kimm, T., et al. 2021, arXiv e-prints, arXiv:2108.01074
Katz, H., Ďurovčíková, D., Kimm, T., et al. 2020, *M*onthly Notices of the RAS, 498, 164
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, *A*nnual Review of Astron and Astrophys, 57, 511
Khokhlov, A. 1998, Journal of Computational Physics, 143, 519
Kimm, T. & Cen, R. 2014, *A*strophysical Journal, 788, 121
Kimm, T., Haehnelt, M., Blaizot, J., et al. 2018, *M*onthly Notices of the RAS, 475, 4617
Kimm, T., Katz, H., Haehnelt, M., et al. 2017, *M*onthly Notices of the RAS, 466, 4826
Kinney, A. L., Bohlin, R. C., Calzetti, D., Panagia, N., & Wyse, R. F. G. 1993, *A*strophysical Journal, Supplement, 86, 5
Koprowski, M. P., Coppin, K. E. K., Geach, J. E., et al. 2018, *M*onthly Notices of the RAS, 479, 4355
Kroupa, P. 2001, *M*onthly Notices of the RAS, 322, 231
Laursen, P., Sommer-Larsen, J., & Andersen, A. C. 2009, *A*strophysical Journal, 704, 1640
Lebouteiller, V., Heap, S., Hubeny, I., & Kunth, D. 2013, *A*stronomy and Astrophysics, 553, A16
Leitherer, C. 1994, Reviews in Modern Astronomy, 7, 73
Leitherer, C., Ekström, S., Meynet, G., et al. 2014, *A*strophysical Journal, Supplement, 212, 14
Lemaître, G. 1927, Annales de la Société Scientifique de Bruxelles, 47, 49
Levermore, C. D. 1984, *J*ournal of Quantitiative Spectroscopy and Radiative Transfer, 31, 149
Li, A. & Draine, B. T. 2001, *A*strophysical Journal, 554, 778
Liang, L., Feldmann, R., Hayward, C. C., et al. 2021, *M*onthly Notices of the RAS, 502, 3210
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *A*strophysical Journal, 835, 113
Luridiana, V., Morisset, C., & Shaw, R. A. 2015, *A*stronomy and Astrophysics, 573, A42
Ma, X., Hopkins, P. F., Garrison-Kimmel, S., et al. 2018, *M*onthly Notices of the RAS, 478, 1694
Ma, X., Quataert, E., Wetzel, A., et al. 2020, arXiv e-prints, arXiv:2003.05945
Madau, P. & Dickinson, M. 2014, *A*nnual Review of Astron and Astrophys, 52, 415
Madau, P. & Haardt, F. 2015, *A*strophysical Journal, Letters, 813, L8
Maio, U., Dolag, K., Ciardi, B., & Tornatore, L. 2007, *M*onthly Notices of the RAS, 379, 963
Marchi, F., Pentericci, L., Guaita, L., et al. 2019, arXiv e-prints, arXiv:1903.08593
Martin, C. L., Shapley, A. E., Coil, A. L., et al. 2012, *A*strophysical Journal, 760, 127
Matthee, J., Sobral, D., Hayes, M., et al. 2021, *M*onthly Notices of the RAS, 505, 1382
Mauerhofer, V., Verhamme, A., Blaizot, J., et al. 2021, *A*stronomy and Astrophysics, 646, A80
McLure, R. J., Dunlop, J. S., Cullen, F., et al. 2018, *M*onthly Notices of the RAS, 476, 3991
Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *A*strophysical Journal, 521, 64
Michel-Dansac, L., Blaizot, J., Garel, T., et al. 2020, *A*stronomy and Astrophysics, 635, A154
Mitchell, P. D., Blaizot, J., Cadiou, C., et al. 2021, *M*onthly Notices of the RAS, 501, 5757
Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, *A*strophysical Journal, 892, 109
Narayanan, A., Sameer, Muzahid, S., et al. 2021, *M*onthly Notices of the RAS, 505, 738

Ocvirk, P., Aubert, D., Sorce, J. G., et al. 2020, *M*onthly Notices of the RAS, 496, 4087
Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, *M*onthly Notices of the RAS, 463, 1462
Oh, S. P. & Benson, A. J. 2003, *M*onthly Notices of the RAS, 342, 664
O'Shea, B. W., Bryan, G., Bordner, J., et al. 2004, arXiv e-prints, astro
Osterbrock, D. E. & Ferland, G. J. 2006, Astrophysics of Gaseous Nebulae and Actice Galactic Nuclei (University Science Books)
Overzier, R. A., Heckman, T. M., Wang, J., et al. 2011, *A*strophysical Journal, Letters, 726, L7
Pallottini, A., Ferrara, A., Decataldo, D., et al. 2019, *M*onthly Notices of the RAS, 487, 1689
Parsa, S., Dunlop, J. S., & McLure, R. J. 2018, *M*onthly Notices of the RAS, 474, 2904
Pentericci, L., Fan, X., Rix, H.-W., et al. 2002, *A*stronomical Journal, 123, 2151
Penzias, A. A. & Wilson, R. W. 1965, *A*strophysical Journal, 142, 419
Perley, D. A., Krühler, T., Schulze, S., et al. 2016a, *A*strophysical Journal, 817, 7
Perley, D. A., Tanvir, N. R., Hjorth, J., et al. 2016b, *A*strophysical Journal, 817, 8
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A*stronomy and Astrophysics, 641, A1
Prusinski, N. Z., Erb, D. K., & Martin, C. L. 2021, arXiv e-prints, arXiv:2102.10187
Reddy, N. A., Shapley, A. E., Sanders, R. L., et al. 2018, *A*strophysical Journal, 869, 92
Reddy, N. A., Steidel, C. C., Pettini, M., Bogosavljević, M., & Shapley, A. E. 2016, *A*strophysical Journal, 828, 108
Rivera-Thorsen, T. E., Hayes, M., Östlin, G., et al. 2015, *A*strophysical Journal, 805, 14
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *A*strophysical Journal, Letters, 802, L19
Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., & Teyssier, R. 2013, *M*onthly Notices of the RAS, 436, 2188
Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, *M*onthly Notices of the RAS, 479, 994
Rosdahl, J. & Teyssier, R. 2015, *M*onthly Notices of the RAS, 449, 4380
Rubin, K. H. R., Prochaska, J. X., Koo, D. C., et al. 2014, *A*strophysical Journal, 794, 156
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, *A*strophysical Journal, 799, 138
Sasaki, S. & Takahara, F. 1993, *P*ublications of the ASJ, 45, 655
Saxena, A., Pentericci, L., Ellis, R. S., et al. 2021, arXiv e-prints, arXiv:2109.03662
Schmidt, K. B., Kerutt, J., Wisotzki, L., et al. 2021, arXiv e-prints, arXiv:2108.01713
Schouws, S., Stefanon, M., Bouwens, R. J., et al. 2021, arXiv e-prints, arXiv:2105.12133
Schroetter, I., Bouché, N. F., Zabl, J., et al. 2019, *M*onthly Notices of the RAS, 490, 4368
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *A*strophysical Journal, 588, 65
Smith, A., Bromm, V., & Loeb, A. 2017, *M*onthly Notices of the RAS, 464, 2963
Smith, A., Safranek-Shrader, C., Bromm, V., & Milosavljević, M. 2015, *M*onthly Notices of the RAS, 449, 4336
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *N*ature, 435, 629
Springel, V., Yoshida, N., & White, S. D. M. 2001, *N*ew Astronomy, 6, 79
Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, *M*onthly Notices of the RAS, 456, 485
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, *A*strophysical Journal, 869, 123
Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996a, *A*stronomical Journal, 112, 352
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996b, *A*strophysical Journal, Letters, 462, L17
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *A*strophysical Journal, 795, 165
Sutherland, R., Dopita, M., Binette, L., & Groves, B. 2018, MAPPINGS V: Astrophysical plasma modeling code
Tanvir, N. R., Fynbo, J. P. U., de Ugarte Postigo, A., et al. 2019, *M*onthly Notices of the RAS, 483, 5380

Tasitsiomi, A. 2006, *A*strophysical Journal, 645, 792
Teyssier, R. 2002, *A*stronomy and Astrophysics, 385, 337
Teyssier, R. 2015, *A*nnual Review of Astron and Astrophys, 53, 325
Toro, E. 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction (Springer Verlag)
Toro, E. F., Spruce, M., & Speares, W. 1994, Shock Waves, 4, 25
Trainor, R. F., Strom, A. L., Steidel, C. C., & Rudie, G. C. 2016, *A*strophysical Journal, 832, 171
Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, *M*onthly Notices of the RAS, 470, 224
Trebitsch, M., Dubois, Y., Volonteri, M., et al. 2020, arXiv e-prints, arXiv:2002.04045
Vazdekis, A., Koleva, M., Ricciardelli, E., Röck, B., & Falcón-Barroso, J. 2016, *M*onthly Notices of the RAS, 463, 3409
Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, *A*strophysical Journal, 465, 487
Vijayan, A. P., Wilkins, S. M., Lovell, C. C., et al. 2021, arXiv e-prints, arXiv:2108.00830
Volonteri, M. & Gnedin, N. Y. 2009, *A*strophysical Journal, 703, 2113
Voronov, G. S. 1997, Atomic Data and Nuclear Data Tables, 65, 1
Walter, F., Decarli, R., Aravena, M., et al. 2016, *A*strophysical Journal, 833, 67
Wise, J. H., Turk, M. J., Norman, M. L., & Abel, T. 2012, *A*strophysical Journal, 745, 50
Wolfe, A. M., Prochaska, J. X., & Gawiser, E. 2003, *A*strophysical Journal, 593, 215
Wünsch, R., Walch, S., Dinnbier, F., et al. 2021, *M*onthly Notices of the RAS, 505, 3730
Zafar, T., Watson, D., Møller, P., et al. 2018, *M*onthly Notices of the RAS, 479, 1542
Zhu, G. B., Comparat, J., Kneib, J.-P., et al. 2015, *A*strophysical Journal, 815, 48

9 List of Publications

The following is a list of papers in which I am a first or co-author, in chronological order. Most of them are from collaborations such as the Rascas-Sphinx group, the MUSE consortium or the low-z Lyman continuum Survey. The first one is from my Master's thesis, during which I studied the Effective Field Theory of Large-scale structures in Stanford University under the supervision of Professor Mikhail Shaposhnikov and Professor Leonardo Senatore.

Published papers

- Fujita et al., 2020: *Very massive tracers and higher derivative biases*
  T. Fujita, **V. Mauerhofer**, L. Senatore, Z. Vlah, and R. Angulo.
  Journal of Cosmology and Astroparticle Physics (2020), 2020:009, doi: 10.1088/1475-7516/2020/01/009

- Mauerhofer et al., 2021: *UV absorption lines and their potential for tracing the Lyman continuum escape fraction*
  V. Mauerhofer, A. Verhamme, J. Blaizot, T. Garel, T. Kimm, L. Michel=Dansac, and J. Rosdahl.
  Astronomy & Astrophysics (2021), 646, A80, doi: 10.1051/0004-6361/202039449

- Gallego et al., 2021: *Constraining the cosmic UV background at $z > 3$ with MUSE Lyman-α emission observations*
  S. G. Gallego, S. Cantalupo, S. Sarpas, B. Duboeuf, S. Lilly, G. Pezzulli, R. A. Marino, J. Matthee, L. Wisotzki, J. Schaye, J. Richard, H. Kusakabe and **V. Mauerhofer**.
  Monthly Notices of the RAS (2021), 504:16-32, doi: 10.1093/mnras/stab796

- Narayanan et al., 2021: *A partial Lyman limit system tracing intragroup gas at $z = 0.8$ towards HE 1003 + 0149*
  A. Narayanan, Sameer, S. Muzahid, S. D. Johnson, P. Udhwani, J. C. Charlton, **V. Mauerhofer**, J. Schaye and M. Yadav.
  Monthly Notices of the RAS (2021), 505:738-754, doi: 10.1093/mnras/stab1315

- Schmidt et al., 2021: *Recovery and analysis of rest-frame UV emission lines in 2052 galaxies observed with MUSE at $1.5 < z < 6.4$*
  K. B. Schmidt, J. Kerutt, L. Wisotzki, T. Urrutia, A. Feltre, M. V. Maseda, T. Nanayakkara, R. Bacon, L. A. Boogaard, S. Conseil, T. Contini, E. C. Herenz, W. Kollatschny, M. Krumpe, F. Leclercq, G. Mahler, J. Matthee, **V. Mauerhofer**, J. Richard, and J. Schaye.
  Astronomy & Astrophysics (2021), 654, A80, doi: 10.1051/0004-6361/202140876

Submitted papers

- Kerutt et al. submitted: *Equivalent Widths of Lyman α Emitters in MUSE-Wide and MUSE-Deep*
  J. Kerutt, L. Wisotzki, A. Verhamme, K. B. Schmidt, F. Leclercq, E. C. Herenz, T. Urrutia, T. Garel, T. Hashimoto, M. Maseda, J. Matthee, H. Kusakabe, J. Schaye, J. Richard, B. Guiderdoni, **V. Mauerhofer**, T. Nanayakkara, and E. Vitte

- Saldana-Lopez et al. submitted: *The Low-Redshift Lyman Continuum Survey Unveiling the ISM properties of low-z Lyman continuum emitters*
  A. Salda na-Lopez, D. Schaerer, J. Chisholm, S. R. Flury, A. E. Jaskot, G. Worseck, K. Makan, S. Gazagnes, **V. Mauerhofer**, A. Verhamme, R. O. Amorín, H. C. Ferguson, M. Giavalisco, A. Grazian, M. J. Hayes, T. M. Heckman, A. Henry, Z. Ji, R. Marques-Chaves, S. R. McCandliss, M. S. Oey, G. Östlin, L. Pentericci, T. X. Thuan, M. Trebitsch, E. Vanzella, and X. Xu.

- Flury et al. submitted: *The Low-Redshift Lyman Continuum Survey II: New Insights into LyC Diagnostics*
  S. R. Flury, A. E. Jaskot, H. C. Ferguson, G. Worseck, K. Makan, J. Chisholm, A. Salda na-Lopez, D. Schaerer, S. McCandliss, B. Wang, N. Ford, T. Heckman, Z. Ji, M. Giavalisco, R. Amorin, H. Atek, J. Blaizot, S. Borthakur, C. Carr, M. Castellano, S. De Barros, M. Dickinson, S. Finkelstein, B. Fleming, F. Fontanot, T. Garel, A. Grazian, M. Hayes, A. Henry, J. Kollmeier, **V. Mauerhofer**, G. Micheva, M. S. Oey, I. Orlitova, G. Ostlin, C. Papovich, L. Pentericci, S. Ravindranath, J. Rosdahl, M. Rutkowski, P. Santini, C. Scarlata, H. Teplitz, T. Thuan, M. Trebitsch, E. Vanzella, A. Verhamme, and I. Wold.

- Maji et al. submitted.: *Predicting LyC emission of galaxies using their physical and LyA emission properties*
  M. Maji, A. Verhamme, J. Rosdahl, T. Garel, J. Blaizot, **V. Mauerhofer**, M. Pittavino, M. V. Feser, M. Chuniaud, T. Kimm, H.Katz, and M. Haehnelt.

- Katz et al. in prep.: Mg II *in the JWST Era: a Probe of Lyman Continuum Escape?*
  H. Katz, T. Garel, J. Rosdahl, **V. Mauerhofer**, T. Kimm, J. Blaizot, L. Michel-Dansac, J. Devriendt, A. Slyz, and M. Haehnelt.