Astraeus VIII: A new framework for Lyman-α emitters applied to different reionisation scenarios

Anne Hutter1,2,3*, Maxime Trebitsch1, Pratika Dayal1, Stefan Gottlöber4, Gustavo Yepes5,6, Laurent Legrand1

1 Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands
2 Cosmic Dawn Center (DAWN)
3 Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200, Copenhagen N, Denmark
4 Leibniz-Institut für Astrophysik, An der Sternwarte 16, 14482 Potsdam, Germany
5 Departamento de Física Teórica, Modulo 8, Facultad de Ciencias, Universidad Autonoma de Madrid, 28049 Madrid, Spain
6 CIAFF, Facultad de Ciencias, Universidad Autonoma de Madrid, 28049 Madrid, Spain

Accepted 2023 July 18. Received 2023 July 18; in original form 2022 September 30

ABSTRACT
We use the Astraeus framework to investigate how the visibility and spatial distribution of Lyman-α (Lyα) emitters (LAEs) during reionisation is sensitive to a halo mass-dependent fraction of ionising radiation escaping from the galactic environment (\(f_{\text{esc}}\)) and the ionisation topology. To this end, we consider the two physically plausible bracketing scenarios of \(f_{\text{esc}}\) increasing and decreasing with rising halo mass. We derive the corresponding observed Lyα luminosities of galaxies for three different analytic Lyα line profiles and associated Lyα escape fraction (\(f_{\text{esc}}^{\text{Lyα}}\)) models: importantly, we introduce two novel analytic Lyα line profile models that describe the surrounding interstellar medium (ISM) as dusty gas clumps. They are based on parameterising results from radiative transfer simulations, with one of them relating \(f_{\text{Lyα}}\) to \(f_{\text{esc}}\) by assuming the ISM of being interspersed with low-density tunnels. Our key findings are: (i) for dusty gas clumps, the Lyα line profile develops from a central to double peak dominated profile as a galaxy’s halo mass increases; (ii) LAEs are galaxies with \(M_\text{h} \gtrsim 10^{10} M_\odot\) located in overdense and highly ionised regions; (iii) for this reason, the spatial distribution of LAEs is primarily sensitive to the global ionisation fraction and only weakly in second-order to the ionisation topology or a halo mass-dependent \(f_{\text{esc}}\); (iv) furthermore, as the observed Lyα luminosity functions reflect the Lyα emission from more massive galaxies, there is a degeneracy between the \(f_{\text{esc}}\)-dependent intrinsic Lyα luminosity and the Lyα attenuation by dust in the ISM if \(f_{\text{esc}}\) does not exceed ~ 50%.

Key words: galaxies: high-redshift - intergalactic medium - dark ages, reionisation, first stars - methods: numerical

1 INTRODUCTION
The Epoch of Reionisation (EoR) marks the second major phase transition in the Universe. With the emergence of the first galaxies, ultraviolet (UV) radiation gradually ionises the neutral hydrogen (H I) in the intergalactic medium (IGM) until the Universe is reionised by \(z \approx 5.3\) (Fan et al. 2006; Keating et al. 2020; Zhu et al. 2021; Bosman et al. 2022). However, as only the brighter galaxies during the EoR are observed to date, key questions detailing the reionisation process remain outstanding: Did the few bright and more massive or the numerous faint and low-mass galaxies contribute more to reionisation? Feedback processes, such as heating by supernovae (SN) and photoionisation, suppress star formation in low-mass galaxies (Gnedin 2000; Gnedin & Kaurov 2014; Ocvirk et al. 2016, 2020; Hutter et al. 2021a), and reduce the contribution of very low-mass galaxies to reionisation. An even more critical quantity that regulates the ionising radiation (with energies \(E > 13.6\) eV) escaping from galaxies and thus the galaxy population driving the reionisation of the IGM is the fraction of ionising photons \(f_{\text{esc}}\) that escape from galaxies into the IGM (e.g. Kim et al. 2013; Seiler et al. 2019; Hutter et al. 2021b; Garaldi et al. 2022).

While the presence of H I in the IGM during the EoR impedes direct measurements of \(f_{\text{esc}}\), different theoretical models and simulations have investigated the physical processes determining and dependencies of \(f_{\text{esc}}\) (e.g. Ferrara & Loeb 2013; Wise et al. 2014; Kimm & Cen 2014; Kimm et al. 2019). Cosmological radiation hydrodynamical simulations suggest that \(f_{\text{esc}}\) decreases towards deeper gravitational potential (e.g. Wise et al. 2014; Kimm & Cen 2014; Kimm et al. 2017, 2019; Xu et al. 2016; Anderson et al. 2017; Lewis et al. 2020). High-resolution simulations of the ISM indicate that \(f_{\text{esc}}\) is dominated by the escape from star-forming clouds. The ionising radiation of massive stars and their explosions as SN ionise, heat and destroy the star-forming clouds clearing the way for the ionising radiation to escape (Howard et al. 2018; Kim et al. 2019; He et al. 2020; Kimm et al. 2022). The complex dependency of \(f_{\text{esc}}\) on the underlying gravitational potential, the gas distribution and stellar populations in the ISM leaves marks not only in the radiation emitted by galaxies but also in the ionisation topology, the time and spatial distribution of the ionised regions around galaxies.

Current and forthcoming observations of galaxies and the ionisa-
tion state of the IGM have the potential to constrain galactic properties, such as $f_{esc}$, and the reionisation process. On the one hand, detecting the 21cm signal from H I in the IGM with forthcoming large radio interferometers (e.g. Square Kilometre Array) will measure the ionisation topology, which provides constraints on the dependence of $f_{esc}$ on galaxy mass (Kim et al. 2013; Seiler et al. 2019; Hutter et al. 2020). On the other hand, being extremely sensitive to the attenuation by H I in the IGM, the observable Lyman-α (Lyα) radiation at 1216Å from high-redshift galaxies has gained popularity in probing reionisation for the following reason: A $z \gtrsim 6$ galaxy only exhibits detectable Lyα emission when: (i) it is surrounded by an ionised region that is large and ionised (i.e. low residual H I fraction) enough to allow a sufficient fraction of its emerging Lyα line to traverse the IGM, or (ii) it is gas-rich enough (corresponding to a high H I column density) such that the red part of the Lyα line emerging from the galaxy is redshifted out of absorption, or (iii) it has strong outflows that redshift the emerging Lyα line out of absorption, or it is a combination of all three. The first criterion suggest that more massive galaxies able to retain more gas might be the most likely to show observable Lyα emission during the EoR: their higher rates of forming stars emitting ionising photons lead to an increased production of Lyα radiation in the ISM and the growth of large ionised regions around them. The latter is accelerated by their ionised regions merging earlier with those of the surrounding lower mass objects attracted by their deeper gravitational potentials (Chardin et al. 2012; Furlanetto & Oh 2016; Chen et al. 2019). As reionisation progresses and the ionised regions grow, increasingly lower mass galaxies become visible as Lyα emitters (LAEs), which leads not only to a higher fraction of galaxies showing Lyα emission but also to a reduced clustering of LAEs (McQuinn et al. 2007; Jensen et al. 2013; Hutter et al. 2015; Sobacchi & Mesinger 2015).

This picture is increasingly supported by observations of $z \gtrsim 6$ LAEs. Not only the fraction of Lyman Break Galaxies (LBGs) showing Lyα emission rises from $z \approx 8$ to $z \approx 6$ (Schenker et al. 2014; Pentericci et al. 2014, 2018; Fuller et al. 2020), but also the majority of Lyα emission at $z \gtrsim 6.5$ is detected in galaxies with a bright UV continuum (Oesch et al. 2015; Zitrin et al. 2015; Roberts-Borsani et al. 2016; Endsley et al. 2022; Endsley & Stark 2022). Moreover, the close proximity of UV-bright LAEs suggests that LAEs are located in over-dense regions (Vanzella et al. 2011; Castellano et al. 2016, 2018; Jung et al. 2020; Tilvi et al. 2020; Hu et al. 2021; Endsley & Stark 2022) that exhibit the first and largest ionised regions during the EoR. This hypothesis is also in line with the observed double-peaked Lyα profiles in $z \gtrsim 6$ galaxies (Songaila et al. 2018; Hu et al. 2016; Matthee et al. 2018; Meyer et al. 2021), indicating that the ionised regions surrounding them are so large that even the part bluewards the Lyα resonance redshifts out of resonance. Current theoretical predictions of the large-scale LAE distribution confirm this picture, suggesting that the LAEs we see during the EoR are more massive galaxies naturally located in over-dense regions (cf. Dayal et al. 2011; Jensen et al. 2013; Hutter et al. 2014; Mesinger et al. 2015; Weinberger et al. 2018; Qin et al. 2022).

Yet, all these LAE models effectively assume a constant $f_{esc}$ value across the entire galaxy population at a given redshift. This assumption remains highly uncertain as $f_{esc}$ is very sensitive to the ISM and the circumgalactic medium (CGM) of galaxies that again depend on the underlying gravitational potential of a galaxy. However, it is essential, since $f_{esc}$ defines the critical processes that shape the Lyα luminosities observed from galaxies. An $f_{esc}$ varying with galactic properties and the underlying gravitational potential might alter the galaxy population seen as LAEs for the following reasons: Firstly, within a galaxy, most Lyα radiation is produced by recombining hydrogen atoms (see e.g. Laursen et al. (2019) and Faucher-Giguère et al. (2010) for an estimate showing that Lyα cooling radiation is subdominant) and scales with the number of H i ionising photons absorbed within the galaxy ($\propto 1 - f_{esc}$). Secondly, a fraction of these Lyα photons undergoes only a few scattering events when they escape through the same low-density tunnels that facilitate the escape of H i ionising photons. In contrast, the other fraction that traverses optically thick clouds upon its escape is scattered and absorbed by hydrogen and dust, respectively (see, e.g. Verhamme et al. 2015; Dijkstra et al. 2016; Kimm et al. 2019; Kakiichi & Gronke 2021). These different escape mechanisms result not only in $f_{esc}$ posing a lower limit to the fraction of Lyα photons escaping from a galaxy but also determining the Lyα line profile that emerges from a galaxy. Detailed low-redshift galaxy observations increasingly supported the $f_{esc}$-sensitivity of these Lyα properties (Verhamme et al. 2017; Jaskot et al. 2019; Gazagnes et al. 2020). Thirdly, $f_{esc}$ shapes the IGM ionisation topology by determining the number of ionising photons available to ionise the IGM surrounding a galaxy. While a higher $f_{esc}$ value enlarges the ionised region surrounding a galaxy and enhances the transmission of Lyα radiation through the IGM (Dayal et al. 2011; Hutter et al. 2014), the corresponding Lyα line emerging from a galaxy will be more peaked around the Lyα resonance and raise the absorption by H I in the IGM. Given this complex $f_{esc}$-dependency of the observed Lyα luminosity, it remains unclear whether different dependencies of $f_{esc}$ with galaxy properties (e.g. increasing or decreasing with rising halo mass) would (i) identify the same galaxies as LAEs (exceeding a threshold Lyα luminosity) and/or (ii) lead to different spatial large-scale distribution of the LAEs’ Lyα luminosities. In other words, which of these $f_{esc}$-dependent Lyα processes dominates the observed Lyα luminosities? For example, is the $f_{esc}$-dependency of the intrinsic Lyα luminosity dominant, and we yield a weaker clustering of LAEs when $f_{esc}$ value decreases with rising halo mass? Or do they compensate each other once we reproduce the observed Lyα luminosity functions (Lyα LFs)?

To address these questions, we use our astræus framework that models galaxy evolution and reionisation self-consistently (Hutter et al. 2021a; Ucci et al. 2023), and simulate different reionisation scenarios that gauge the physically plausible range of $f_{esc}$ dependencies, i.e. $f_{esc}$ decreasing and increasing with rising halo mass. Moreover, we parameterise results from numerical Lyα radiative transfer (RT) simulations of clumpy media (Gronke 2017) and build an analytic model for the fraction of Lyα photons escaping and the corresponding Lyα line profile emerging from high-redshift galaxies. Importantly, we explore three different Lyα line profile models, including (i) a Gaussian profile around the Lyα resonance where the Lyα escape fraction is directly related to the dust attenuation of the UV continuum (used in previous LAE models outlined in Dayal et al. 2011; Hutter et al. 2014), (ii) a Lyα line profile emerging from a shell of dusty gas clumps, which we model by using the different Lyα escape regimes identified in Gronke (2017), and (iii) a Lyα line profile emerging from a shell of gas clumps with a fraction $f_{esc}$ of the solid angle interspersed by gas-free tunnels. The latter two give rise to various combinations of a central peak around the Lyα resonance (Lyα photons hardly scatter in an optically thin medium) and two peaks in the red and blue wings (Lyα photons are scattered in an optically thick medium). By deriving the observed Lyα luminosities of all simulated galaxies for all combinations of reionisation scenarios and Lyα line models, we address the following questions: Which $f_{esc}$-dependent Lyα process, i.e. intrinsic production, escape or transmission through the IGM of Lyα radiation, dominates the observed Lyα luminosity? Can the observed Lyα luminosities of galaxies inform us on their emerging Lyα line profile? Given the

MNRS 000, 1–25 (2022)
ionisation topology depends sensitively on the assumed dependency of $f_{\text{esc}}$ with halo mass, are the same or different galaxies identified as LAEs and do they differ in the spatial distribution of their Ly$\alpha$ luminosities?

This paper is organised as follows. In Section 2 we briefly describe the astraeus model, its implementation of dust and the different reionisation simulations. In Section 3 we introduce the different Ly$\alpha$ line profile models and their corresponding attenuation by dust. We then (Section 4) discuss how the Ly$\alpha$ line profiles depend on halo mass in our different reionisation scenarios, how free model parameters, such as the ISM clumpiness or size of the dust gas clumps, need to be adjusted to fit the observed Ly$\alpha$ LFs, and how the galaxy properties determining the observed Ly$\alpha$ luminosities depend on the halo mass of a galaxy. In Section 5 we identify the location of LAEs in the large-scale density and ionisation structure and assess whether the spatial distribution of LAEs differs for different $f_{\text{esc}}$-dependencies on halo mass/ionisation topologies. Finally, we briefly discuss which Lyman Break galaxies are preferentially identified as LAEs and do they differ in the spatial distribution of their Ly$\alpha$ line profiles?

2 THE MODEL AND SIMULATIONS

In this paper, we use the astraeus framework. This framework couples a semi-analytic galaxy evolution model (an enhanced version of delphi; Dayal et al. 2014) with a semi-numerical reionisation scheme (cifog; Hutter 2018) and runs the resulting model on the outputs of a dark matter (DM) only N-body simulation. In this Section, we provide a brief description of the physical processes implemented in astraeus (for more details, see Hutter et al. 2021a) and introduce the different reionisation simulations.

2.1 N-body simulation

As part of the Multidark simulation project, the underlying DM N-body simulation (very small multidark planck; vsmdpl) has been run with the gadget-2 tree+pm code (Springel 2005). In a box with a side length of 160$h^{-1}$Mpc, it follows the trajectories of 3840$^3$ DM particles. Each DM particle has a mass of 6$\times$10$^6 h^{-1}$M$_\odot$. For a total of 150 snapshots ranging from $z = 25$ to $z = 0$, the phase space rockstar halo finder (Behroozi et al. 2013a) has been used to identify all halos and subhalos down to 20 particles or a minimum halo mass of 1.24$\times$10$^8 h^{-1}$M$_\odot$. To obtain the local horizontal merger trees (sorted on a redshift-by-redshift basis within a tree) for galaxies at $z = 4.5$ that astraeus requires as input, we have used the pipeline internal cutresort scheme to cut and resort the vertical merger trees (sorted on a tree-branch by tree-branch basis within a tree) generated by consistent trees (Behroozi et al. 2013b). For the first 74 snapshots that range from $z = 25$ to $z = 4.5$, we have generated the DM density fields by mapping the DM particles onto 2048$^3$ grids and re-sampling these to 512$^3$ grids used as input for the astraeus pipeline.

2.2 Galaxy evolution

astraeus tracks key processes of early galaxy formation and reionisation by post-processing the DM merger trees extracted from the vsmdpl simulation. At each time step (i.e. snapshot of the N-body simulation) and for each galaxy, it tracks the amount of gas that is accreted, the gas and stellar mass merging, star formation and associated feedback from SNII and metal enrichment, as well as the large-scale reionisation process and its associated feedback on the gas content of early galaxies.

2.2.1 Gas and stars

In the beginning, when a galaxy starts forming stars in a halo with mass $M_h$, it has a gas mass of $M_g^i(z) = f_g (\Omega_b/\Omega_m) M_h(z)$, with $f_g$ being the gas fraction not evaporated by reionisation, i.e. $f_g = 1$ and $f_g < 1$ as the galaxy forms in a neutral and ionised region, respectively. In subsequent time steps a galaxy gains gas from its progenitors ($M_g^{\text{mer}}(z)$) and smooth accretion ($M_g^{\text{acc}}(z)$), while its total gas mass never exceeds the limit given by reionisation feedback:

$$M_g^i(z) = \min \left( M_g^{\text{mer}}(z) + M_g^{\text{acc}}(z), f_g (\Omega_b/\Omega_m) M_h(z) \right)$$

with

$$M_g^{\text{mer}}(z) = M_h(z) - \sum_{p=1}^{N_p} M_{h,p}(z + \Delta z)$$

$$M_g^{\text{acc}}(z) = \sum_{p=1}^{N_p} M_{h,p}(z + \Delta z),$$

where $N_p$ is the galaxy’s number of progenitors and $M_{h,p}$ the halo mass of each progenitor.

At each time step, a fraction of the merged (initial) gas mass is transformed into stellar mass, $M_g^{\text{new}}(z) = (f_\star^{\text{eff}} / \Delta t) M_g^i(z)$.\footnote{We note that this definition has been altered compared to the first version of astraeus in (Hutter et al. 2021a).}

Here $f_\star^{\text{eff}}$ represents the fraction of gas that forms stars over a time span $\Delta t$ and is limited by the minimum amount of stars that need to form to eject all gas from the galaxy, $f_\star^i$, and an upper limit, $f_\star^\text{max}$.

$f_\star^{\text{eff}}$ depends on the gravitational potential: more massive galaxies form stars at the constant rate $f_\star^i$, while low-mass galaxies form stars at the limited rate $f_\star^\text{max}$ due to SN and radiative feedback. While we account for radiative feedback from reionisation by modifying the initial gas mass reservoir with the factor $f_g$, $f_\star^{\text{eff}}$ incorporates the suppression of star formation in low-mass halos as gas is heated and ejected by SNII explosions. Our model incorporates a delayed SN feedback scheme, i.e. at each time step the effective star formation efficiency accounts for the SNII energy released from stars formed in the current and previous time steps, following the mass-dependent stellar lifetimes (Padovani & Matteucci 1993). In contrast to Hutter et al. (2021a), we have updated our model and do not assume stars to form in bursts to calculate the number of SNII exploding within a time step but $M_g^{\text{new}}(z)$ to form at a constant star formation over the entire time step (see Appendix B for a detailed calculation). The star formation efficiency in the SN feedback-limited regime is given by

$$f_\star^i(z) = \frac{v_c^2}{v_s^2 + f_a E_S I v_c} \left[ 1 - \frac{f_a E_S \sum_j v_j M_g^{\text{new}}(z,j)}{M_g^i(z) v_C^2} \right],$$

with $v_c$ being the rotational velocity of the halo, $E_S$ the energy released by a SNII, $f_a$ the fraction of SNII energy injected into the winds driving gas outflows, $M_g^{\text{new}}(z,j)$ the stellar mass formed during previous time steps $j$, and $v_j$ the fraction of stellar mass formed in previous time step $j$ that explodes in the current time step given the assumed IMF.
2.2.2 Metals and dust

The current 	extsc{asteraues} model also incorporates the metal enrichment by stellar winds, SNII and SNIa explosions (for a detailed description see Ucci et al. 2023). At each time step, we assume that gas smoothly accreted has the average metallicity of the gas in the IGM, $Z_{\text{IGM}}$. Metals are produced through stellar winds, SNII and SNIa explosions. The amount of newly formed metals depends on the number of massive stars exploding as SN in the current time step according to Padovani & Matteucci (1993), Yates et al. (2013) and Maoz et al. (2012). For the corresponding stellar metal yields, 	extsc{asteraues} uses the latest yield tables from Kobayashi et al. (2020). We assume that gas and metals are perfectly mixed. Thus, the metals ejected from the galaxy are proportional to the ejected gas mass and the metallicity of the gas in the galaxy. This ejected metal mass contributes to $Z_{\text{IGM}}$.

In this work, we have extended the 	extsc{asteraues} model (Hutter et al. 2021a; Ucci et al. 2023) to follow the formation, growth, destruction, astration and destruction of dust in each galaxy (c.f. Dayal et al. 2022, for details). We note that we consider dust to be part of our metal reservoir (i.e. $M_{\text{dust}} \leq M_{\text{m}}$). At each time step, 	extsc{asteraues} computes the evolution of the dust mass $M_{\text{dust}}$ in a galaxy by solving the following differential equation

$$\frac{dM_{\text{dust}}}{dt} = M_{\text{prod}}^{\text{dust}} + M_{\text{grow}}^{\text{dust}} - M_{\text{dest}}^{\text{dust}} - M_{\text{astr}}^{\text{dust}} - M_{\text{ej}}^{\text{dust}}.$$  

The first term on the right hand side (RHS) of Eqn. 5 describes the dust grain growth through the accretion of heavy elements in dense molecular clouds in the ISM,

$$M_{\text{grow}}^{\text{dust}} = \left( Z' - \frac{M_{\text{dust}}}{M_{\text{g}}^*} \right) f_{\text{cold gas}} \frac{M_{\text{dust}}}{\tau_{\text{gg,0}} Z_{\odot}}$$  

where $Z'$ is the metallicity after accretion and star formation, $M_{\text{dust}}$ is the dust mass, $f_{\text{cold gas}}$ the fraction of cold and molecular gas, and $\tau_{\text{gg}} = \tau_{\text{gg,0}} / Z$ the accretion timescale adopted from Asano et al. (2013) (see also Triani et al. 2020). We assume $f_{\text{cold gas}} = 0.5$ and $\tau_{\text{gg,0}} = 30$Myrs. The third term in Eqn. 5 describes the destruction of dust by SN blastwaves, for which we adopt the analytic description outlined in McKee (1989)

$$M_{\text{dest}}^{\text{dust}} = \left( 1 - f_{\text{cold gas}} \right) \frac{M_{\text{dust}}}{M_{\text{g}}} \gamma_{\text{SN}} e M_{\text{SN,bw}},$$

with $e$ being the efficiency of dust destruction in a SN-shocked ISM and $M_{\text{SN,bw}}$ the mass accelerated to 100 km s$^{-1}$ by the SN blast wave. In line with McKee (1989) and Liskenfeld & Ferrara (1998) we adopt $e = 0.03$ and $M_{\text{SN,bw}} = 6.8 \times 10^3 M_{\odot}$. Finally, Eqn. 5 accounts also for the destruction of dust by astration as new stars form from the metal-enriched gas,

$$M_{\text{astr}}^{\text{dust}} = Z' \frac{M_{\text{new}}}{\Delta t},$$

and the ejection of metals through winds powered by the energy injected by SN,

$$M_{\text{ej}}^{\text{dust}} = Z' \frac{M_{\text{ej}}}{\Delta t}.$$  

The parameter values ($\gamma_{\text{SN}}, \tau_{\text{gg,0}}, e, M_{\text{SN,bw}}$) quoted reasonably reproduce the observed UV LFs when the UV is attenuated by dust as follows (please see Hutter et al. (2021a) for observational UV LFs data points included): From the dust mass, $M_d$, we obtain the total optical depth to UV continuum photons as (see e.g. Dayal et al. 2011)

$$\tau_{\text{UV,c}} = \frac{3\Sigma_d}{4\pi s},$$

with $\Sigma = M_d / (\pi r_d^2)$ being the dust surface mass density, $r_d$ the dust distribution radius, and $a = 0.03$ mm and $s = 2.25g$ cm$^{-3}$ the radius and material density of graphite/carbonaceous grains (Tohdini & Ferrara 2001). Since we assume that dust and gas are perfectly mixed, we equate the dust distribution radius, $r_d$, with the radius of the gas, $r_g = 4.5 \lambda_{\text{vir}} / [(1 + z) / 6]^{1.8}$. Here $\lambda$ is the spin parameter of the simulated halo, $r_{\text{vir}}$ the virial radius, and the third factor accounts for the redshift evolution of the compactness of galaxies and ensures that the observed UV LF at $z = 5 - 10$ are well reproduced. For a slab-like geometry, the escape fraction of UV continuum photons of a galaxy is then given by

$$f_{\text{esc}} = 1 - \exp\left(-\tau_{\text{UV,c}}\right) \frac{\tau_{\text{UV,c}}}{\tau_{\text{UV,c}}},$$

and its observed UV luminosity by

$$L_{\text{c}}^{\text{obs}} = f_{\text{esc}} L_{\text{c}},$$

with the intrinsic UV luminosity, $L_c$, being computed as outlined in Section 2.2.4 in Hutter et al. (2021a).

\section{2.3 Reionisation}

At each time step 	extsc{asteraues} follows the time and spatial evolution of the ionised regions in the IGM. For this purpose, it derives the number

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Parameter & Value or reference & Description \\
\hline
$f^*$ & 0.025 & Maximum star-formation efficiency \\
$f_w$ & 0.2 & SN coupling efficiency \\
$\gamma$ & - & Photoionization Radiative feedback model \\
\text{IMF} & Salpeter (1955) & For stellar evolution, enrichment, SED \\
\text{SED} & Starburst99 & ionizing SED model \\
\hline
\end{tabular}
\caption{\textsc{asteraues} model parameters and chosen values in this work.}
\end{table}
of ionising photons produced in each galaxy, $\dot{Q}$, by convolving the galaxy’s star formation rate history with the spectra of a metal-poor ($Z = 0.05 Z_\odot$) stellar population. Spectra have been obtained from the stellar population synthesis model starburst99 (Leitherer et al. 1999). Again we assume that stars form continuously over a time step. Then the number of ionising photons that contribute to the ionisation of the IGM is then given by

$$N_{\text{ion}} = f_{\text{esc}} \dot{Q},$$

(16)

where $f_{\text{esc}}$ is the fraction of ionising photons that escape from the galaxy into the IGM. From the resulting ionising emissivity and gas density distributions astraus derives the spatial distribution of the ionised regions by comparing the cumulative number of ionising photons with the number of absorption events (see e.g., Hutter 2018, for details). Within ionised regions, it also derives the photoionisation rate and residual H I fraction in each grid cell. The ionisation and photoionisation fields obtained allow us then to determine on the fly whether the environment of a galaxy has been reionised and account for the corresponding radiative feedback by computing the gas mass the galaxy can hold on to ($f_g M_h^*$.)

### 2.4 Simulations

In the following we consider three different reionisation scenarios that explore the physically plausible space of the ionising escape fraction $f_{\text{esc}}$ (c.f. Fig. 1):

(i) \textbf{MHDEC}: $f_{\text{esc}}$ decreases with rising halo mass of a galaxy (red solid line)

$$f_{\text{esc}} = f_{\text{esc,low}} \left( \frac{f_{\text{esc,high}}}{f_{\text{esc,low}}} \right)^{\log_{10}(M_h/M_{h,\text{low}})/\log_{10}(M_{h,\text{high}}/M_{h,\text{low}})}$$

(17)

with $f_{\text{esc,low}} = 0.55$, $f_{\text{esc,high}} = 0.05$, $M_{h,\text{low}} = 2 \times 10^8 h^{-1} M_\odot$ and $M_{h,\text{high}} = 10^{10} h^{-1} M_\odot$.

(ii) \textbf{MHCONST}: $f_{\text{esc}} = 0.16$ for each galaxy (magenta dash-dotted line).

(iii) \textbf{MHINC}: $f_{\text{esc}}$ increases with rising halo mass of a galaxy (blue dotted line) following Eqn. 17 with $f_{\text{esc,low}} = 0.08$, $f_{\text{esc,high}} = 0.4$, $M_{h,\text{low}} = 10^9 h^{-1} M_\odot$ and $M_{h,\text{high}} = 10^{11} h^{-1} M_\odot$.

These three $f_{\text{esc}}$ prescriptions have been adjusted to reproduce the electron optical depth measured by Planck (Planck Collaboration et al. 2020) and fit the observational constraints from LAEs, quasar absorption spectra and gamma ray bursts (as depicted in the lower panel of Fig. 2). In addition, for \textbf{MHINC} the maximum $f_{\text{esc}}$ value of more massive galaxies is also limited by the observed Ly$\alpha$ LFs. Despite having very similar electron optical depths, these three $f_{\text{esc}}$ prescriptions lead to different ionisation histories and topologies (see Fig. 2 and 6). As $f_{\text{esc}}$ decreases with rising halo mass, reionisation is dominated by the low-mass galaxies ($M_h \lesssim 10^{10} M_\odot$), leading to on average smaller ionised regions and lower photoionisation rates. Since these low-mass galaxies appear earlier, reionisation begins earlier (see solid red line in Fig. 2); however, as shown in Hutter et al. (2021a) for the \textit{Photoionisation} model their overall star formation rate decreases around $z \approx 7$, resulting in the Universe being reionised at a later time and exhibiting a higher average residual H I fraction in ionised regions. In contrast, as $f_{\text{esc}}$ increases with rising halo mass, more massive galaxies ($M_h \gtrsim 10^{10} M_\odot$) drive reionisation. On average, ionised regions are larger and more clustered around more massive galaxies, and photoionisation rates within these ionised regions are higher. Reionisation begins later with the appearance of more massive galaxies and ends earlier as the abundance of these massive galaxies increases.
3 MODELLING $\text{Ly}\alpha$ EMITTERS

In this Section, we introduce the different models for the emergent $\text{Ly}\alpha$ line profiles (Section 3.1) and fractions of $\text{Ly}\alpha$ radiation escaping from a galaxy (Section 3.3), describe the attenuation of $\text{Ly}\alpha$ radiation by H I in the IGM, and the derivation of the observed $\text{Ly}\alpha$ luminosity of a galaxy (Section 3.4). We summarise the combinations of emergent $\text{Ly}\alpha$ line profile and dust attenuation models investigated in this paper in Section 3.3.3.

3.1 Emerging $\text{Ly}\alpha$ line profiles

We investigate three $\text{Ly}\alpha$ line profiles $J(x)$: (1) a thermally Doppler-broadened Gaussian centred at the $\text{Ly}\alpha$ resonance; (2) a single, double or triple-peaked profile that depends on the clumpiness and H I column density of the gas in a galaxy; (3) a single, double or triple-peaked profile that depends both on the ionising escape fraction $f_{\text{esc}}$ and the clumpiness and H I column density of the gas in a galaxy. While the first model represents a simple assumption used in previous works (e.g. Dayal et al. 2011; Hutter et al. 2014), the latter two models are inspired by observations and detailed $\text{Ly}\alpha$ radiative transfer simulations (e.g. Dijkstra et al. 2016; Gronke 2017). The $\text{Ly}\alpha$ line emerging from a galaxy is given by the intrinsic $\text{Ly}\alpha$ luminosity, $L_{\alpha}^{\text{gal}} = \frac{1}{2} Q(1 - f_{\text{esc}}) h\nu_{\alpha}$, the escape fraction $\text{Ly}\alpha$ photons from the galaxy, $f_{\text{esc}}^\text{Ly}\alpha$, and the line profile $J(x)$.

$$L_{\alpha}^\text{gal}(x) = L_{\alpha}^{\text{intr}} f_{\text{esc}}^\text{Ly}\alpha J(x) \quad (18)$$

Here we have expressed the frequency deviation from the $\text{Ly}\alpha$ resonance $\nu_{\alpha}$ in terms of the thermal line broadening $\sigma_{th} = (\nu_{th}/c)\nu_{\alpha}$ with $\nu_{th} = \sqrt{2kBT/m_H}$, yielding $x = \nu - \nu_{*\alpha}/\sigma_{th}$. $k_B$ is the Boltzmann constant, $m_H$ the mass of a hydrogen atom and $T$ the temperature of the H I gas. In the remainder of this Section, we detail our different models for the $\text{Ly}\alpha$ line profiles and escape fractions.

3.1.1 Central Gaussian

This model assumes that the emission sites of $\text{Ly}\alpha$ radiation, the hydrogen atoms within a galaxy, move at velocities that reflect the galaxy’s rotation. The corresponding Doppler-broadened $\text{Ly}\alpha$ line profile is then given by

$$J_{\text{centre}}(x) = \frac{1}{\sqrt{\pi}} \frac{\sigma_{th}}{\sigma_{r}} \exp\left[-x^2 \frac{\sigma_{th}^2}{\sigma_{r}^2}\right] \quad (19)$$

We note that since the $\sigma_{th}$-dependence of $x$ cancels any dependency of $J_{\text{Gaussian}}(\nu)$ on $\sigma_{th}$, the assumed gas temperature has no effect on the emerging $\text{Ly}\alpha$ line profile (we use $T = 10^4$ K in Fig. 3). $\sigma_{r} \approx (\nu_{r}/c)\nu_{\alpha}$ describes the Doppler broadening of the line due to the rotation of the galaxy. The rotation velocity of the galaxy $v_r$ is closely linked to the halo rotational velocity $v_c = (3\pi G M_{\text{halo}})^{1/3}/\Omega_m^{1/6}(1 + z)^{1/2} M_{\text{halo}}^{1/3}$, ranging between $v_r = v_c$ and $v_r = 2v_c$ (Mo et al. 1998; Cole et al. 2000). We assume $v_r = 1.5v_c$.

3.1.2 Single, double or triple-peak in a clumpy/homogeneous medium

This model describes the $\text{Ly}\alpha$ line profile emerging from a clumpy medium. It implements the regimes and characteristic escape frequencies identified in Gronke (2017). We consider a slab with a thickness of $2B$ and a total optical depth of $2\tau_0$. The source is located at the slab’s midplane and injects photons at the $\text{Ly}\alpha$ resonance $x = 0$. If the slab medium is homogeneous, Neufeld (1990) derived the emergent $\text{Ly}\alpha$ profile as

$$J_{\text{slab}}(T, \tau_0, x) = 4\pi \frac{\sqrt{6}}{24} \frac{\tau_0^2}{a(T)} \frac{1}{\cosh\left(\frac{\sqrt{54} |x|}{34 a(T) \tau_0}\right)} \quad (20)$$

for $a(T)\tau_0 \gtrsim 3^0$ with $a(T) = \frac{A_{\alpha}}{M_{\text{gal}}^0 a(T)}$ and $\int_{-\infty}^{\infty} J_{\text{slab}}(x, x) \, dx = 1$. $A_{\alpha}$ is the Einstein for the spontaneous emission of $\text{Ly}\alpha$ photons.

In the following we will revisit the regimes for $\text{Ly}\alpha$ escape in a clumpy medium that have been identified in Gronke (2017). The clumpy medium is characterised by the total optical depth of the clumps and the average number of clumps each $\text{Ly}\alpha$ photon escaping the slab scatters with. For a slab consisting of clumps with each having an optical depth $\tau_{0,cl}$ at the line centre, $\text{Ly}\alpha$ photons escaping the slab will encounter on average $f_c$ clumps and have a total optical depth of $\tau_0 = \frac{1}{2} f_c \tau_{0,cl}^2$ at the line centre. The emerging $\text{Ly}\alpha$ line profile depends sensitively on the total and clump optical depth at line centre, $\tau_0$ and $\tau_{0,cl}$, respectively, and the number of clumps the $\text{Ly}\alpha$ photons scatter with. Gronke (2017) identified the following regimes:

- **Free-streaming regime**: The clumpy medium is optically thin ($\tau_{0} < 1$), and $\text{Ly}\alpha$ photons can stream through. The emerging line profile peaks around $x = 0$.
- **Porous regime**: The clumps are optically thick to $\text{Ly}\alpha$ photons ($t_{cl} > 1$), but only a fraction $1 - \exp(-\tau_{0,cl})$ of the $\text{Ly}\alpha$ photons scatter with a clump. The emerging line profile is again peaked around $x = 0$.
- **Random walk regime**: The clumps are optically thick to $\text{Ly}\alpha$ ($t_{cl} > 1$), and each $\text{Ly}\alpha$ photon encounters $N_{\text{cl}} \propto f_{c,0}^2$ scattering events (Hansen & Oh 2006). However, the number of scattering events is too low for the $\text{Ly}\alpha$ photons to scatter in frequency space far enough into the wings to escape through excursions. Hence, the emerging line profile peaks also around $x = 0$.
- **Homogeneous regime**: The clumps are optically thin ($t_{cl} \leq 1$) and $\text{Ly}\alpha$ photons scatter $\sim \tau_0$ times ($N_{\text{sc}} \propto f_c$) and escape via excursions: they follow a random walk in space and frequency and escape as they are scattered into the wings where the clumps become optically thin. The emerging line profile is a double-peak with the two peaks being located at

$$x_{\text{esc}}(\tau_0) = \begin{cases} \pm \frac{(k\nu\tau_0/\sqrt{\pi})^{1/3}}{2\nu_{*\alpha}^2} - \frac{\sqrt{\pi} x_{\alpha}^2}{k\nu} \leq \tau_0, & \tau_0 < \frac{\sqrt{\pi} x_{\alpha}^2}{k\nu} \\ + x_{*}, & \text{else} \end{cases} \quad (21)$$

for an injection frequency $x = 0$ (c.f. Adams 1975; Gronke 2017). Here $x_{*}$ is the frequency where the $\text{Ly}\alpha$ absorption profile transitions from the Gaussian core to the Lorentzian wings, and $\tau_0 = \frac{\sqrt{\pi} x_{\alpha}^2}{k\nu}$ marks the transition where the slab becomes optically thin at the escape frequency $x_{\text{esc}}$.

Gronke (2017) derived the boundary criteria between these regimes for a static clumpy medium, which we briefly revisit here. To derive the critical number of clumps, $f_c$, separating the regimes, we first consider the time and distances covered that it takes a $\text{Ly}\alpha$ to traverse the slab.

**Excursion**: As $\text{Ly}\alpha$ photons traverse or escape the slab, they scatter with H I many times. This alters their direction and frequency $x$, and they essentially perform a random walk. However, as the $\text{Ly}\alpha$ cross section is higher close to the line centre, most scatterings will occur close to the line centre and remain spatially close.

$^2$ The factor $4/3$ arises from the mean path length through a sphere.
Only as the Ly\(\alpha\) photons are scattered into the wings of the Ly\(\alpha\) absorption profile their mean free paths become larger, allowing them to escape the slab (Adams 1975). The series of these so-called wing scatterings that allow Ly\(\alpha\) photons to escape are referred to as excursion. We can estimate the mean displacement and time spent in such an excursion event: a Ly\(\alpha\) photon with frequency \(x\) will scatter on average \(x^2\) times before it returns to the core. For a slab of thickness \(B\), its average mean free path is \(\lambda_{\text{mfp,exc}}(x) = \frac{B}{\tau_B(k_B\sigma_T(x))} = B/(k_B\tau_B H(a,x))\) using the wing approximation of the Ly\(\alpha\) cross section and \(k\) \(\approx \frac{1}{3}\) being a geometrical factor that accounts for slant paths in a plane-parallel medium and was determined in Adams (1975). This and the random walk nature of the Ly\(\alpha\) escape imply an average displacement of

\[
d_{\text{exc}} = \sqrt{N_{\text{exc},\text{mfp,exc}}(x)} = \frac{\sqrt{N_{\text{exc},\text{mfp,exc}}B}}{k_B\tau_B H(a,x)} = \frac{xB}{k_B\tau_B H(a,x)}
\]

and time spent in the excursion of

\[
t_{\text{exc}} = N_{\text{exc,exc}} \frac{\lambda_{\text{mfp,exc}}(x)}{c} = \frac{N_{\text{exc,exc}}B}{ck_B\tau_B H(a,x)} = \frac{x^2B}{ck_B\tau_B H(a,x)},
\]

with \(H_B(a,x) = \frac{a}{\sqrt{x^2 + a^2}}\) being the effective line absorption profile in the wings.

**Random Walk**: As the clumps become optically thick at corresponding escape frequencies \(x_{\text{esc}}(\tau_0)\), the Ly\(\alpha\) photons do not escape the slab via excursion anymore but by random walking: the number of scattering events is smaller than required for excursion and scale with the square of the number of clumps, \(N_{\text{exc,exc}} \propto f^2\) (Hansen & Oh 2006). With the mean free path given by the average clump separation \(\lambda_{\text{mfp,exc}} = kB/\alpha\), the average displacement and time are then

\[
d_{\text{rw}} = \sqrt{N_{\text{exc,exc}}\lambda_{\text{mfp,exc}}} = \sqrt{N_{\text{exc,exc}}kB} = kB
\]

and

\[
t_{\text{rw}} = \sqrt{N_{\text{exc,exc}}\lambda_{\text{mfp,exc}}} = \sqrt{N_{\text{exc,exc}}kB} = kB/\alpha
\]

(i) **Division between random-walk and homogeneous regime in optically thick medium**: For a given total optical depth at the line centre, \(\tau_0\), we can derive the critical number of clumps along a line of sight that marks the transition from the random (clumps are optically thick) to the homogeneous regime (clumps become optically thick at the excursion frequency). We estimate this transition to arise when both regimes contribute equally to the flux of escaping Ly\(\alpha\) photons.

\[
\frac{F_{\text{rw}}}{F_{\text{exc}}} = \frac{t_{\text{exc}}}{t_{\text{rw}}} = \frac{x^2}{k_B^2\tau_B H(a,x)c} = \frac{\sqrt{\pi}x^4}{k_B^2\alpha^2\tau_{0,c}\alpha} = 1
\]

With \(\tau_0 = 4/3f_c\tau_{0,\text{cl}}\), the critical number of clumps for Ly\(\alpha\) photons escaping at frequency \(x\) then yields as

\[
f_c = \sqrt{2\sqrt{\pi}x}\frac{1}{k_B\tau_{0,\text{cl}}}.
\]

As long as the wings remain optically thick, the majority of Ly\(\alpha\) photons (with injection frequency \(x = 0\)) will escape at \(x_{\text{esc}} = \frac{\sqrt{\pi}x}{k_B\tau_{0,\text{cl}}}\).

3 The geometrical factor \(k\) is often not explicitly included in the literature when displacements and excursion times are discussed.

This \(f_{c,\text{crit}}\) value marks the transition from the random walk to the escape frequency. We can understand its increase with the clump optical depth \(\tau_{0,\text{cl}}\) as follows: for optically thicker clumps to become optically thin at the escape frequency \(x_{\text{esc}}\), a higher escaping frequency and thus a higher effective total optical depth are required. This can only be achieved by interacting with more clumps (higher \(f_{c,\text{crit}}\)).

Because the transition described by \(f_{c,\text{crit}}\) is not sharp, we model the Ly\(\alpha\) line profile emerging from the moving slab by superposing the Ly\(\alpha\) radiation escaping in the homogeneous \((J_{\text{slab}})\) and random walk regimes \((J_{\text{centre}})\).

\[
J_{\text{slab}}(\tau_0, x) = (1 - f_{\text{rw}})J_{\text{slab}}(T, \tau_0, x) + f_{\text{rw}}J_{\text{centre}}(T, x)
\]

Here we assume \(J_{\text{centre}}\) is given by Eqn. 19 with \(\sigma_T = \sigma_T\) and

\[
f_{\text{rw}} = \frac{F_{\text{rw}}}{F_{\text{exc}}} = \frac{J_{\text{slab}}(T, \tau_0, x) + J_{\text{slab}}(T, \tau_0, x_{\text{esc}})}{J_{\text{slab}}(T, \tau_0, x_{\text{esc}})}
\]

We derive the corresponding ratio \(f_{\text{rw}}\) by assuming that the Ly\(\alpha\) flux escapes predominantly where the Ly\(\alpha\) profiles peak,

\[
f_{\text{rw}} = \frac{F_{\text{rw}}}{F_{\text{exc}}} = \frac{f_{\text{cen}}}{f_{\text{cl}}}
\]

We note that \(J_{\text{slab}}\) reproduces the results of Ly\(\alpha\) radiative transfer simulations down to \(\tau_0 \approx 10^5\). While the line profiles for lower \(\tau_0\) values start deviating, we will see in Section 3.2 that the galaxies considered here exceed this threshold. Importantly, we find that the assumed \(J_{\text{slab}}, J_{\text{centre}}\) and \(f_{\text{rw}}\) reproduce the Ly\(\alpha\) line profiles for resting clumps, fixed \(\tau\) values, and varying \(f_c\) values in Gronke (2017).

(ii) **Division between porous and homogeneous regime in optically thin medium**: As the medium becomes optically thinner, Ly\(\alpha\) photons that scatter into the wings can escape the slab before completing their excursion. This transition occurs as the wings become optically thin, i.e. \(k\tau(x) \leq 1\) translating to \(k\tau_{0,\text{cl}} \leq \sqrt{\pi}x\), and Ly\(\alpha\) photons escape at \(x_{\text{esc}} = x\). While the slab is optically thin at \(x\), depending on whether the clumps are optically thin or thick at \(x_{\text{esc}}\), the escape of Ly\(\alpha\) photons is described by the homogeneous and porous regime, respectively. Again we estimate the transition to arise when both regimes contribute equally to the flux of escaping Ly\(\alpha\) photons.

\[
\frac{F_{\text{por}}}{F_{\text{hom}}} = \frac{t_{\text{exc}}}{t_{\text{rw}}} = \frac{\sqrt{\pi}x^4}{k_B^2\alpha^2\tau_{0,\text{cl}}} = 1
\]

We yield the critical number of clumps that mark the transition from the porous to the homogeneous regime as

\[
f_c = \frac{x}{k(1 - e^{-x_{\text{esc}}})}
\]

We note that if clumps are optically thin at line centre \((\tau_{0,\text{cl}} < 1)\), not every clump encounter leads to a scattering event; this reduces the number of clumps encountered by a factor \(1 - e^{-x_{\text{esc}}\tau_{0,\text{cl}}}\). The emerging Ly\(\alpha\) line profile accounts again for Ly\(\alpha\) photons escaping in homogeneous \((J_{\text{slab}},\) see Eqn. 30) and porous regime \((J_{\text{centre}},\) see Eqn. 33).
\[ J_{\text{ph}}(\tau_0, x) = (1 - \phi_{\text{esc}}) J_{\text{slab}}(T, \tau_0, x) + \phi_{\text{esc}} J_{\text{centre}}(T, x) \]  \hspace{1cm} (34)

The ratio between the two different escape regimes is then again given by assuming that most Ly\(\alpha\) photons escape at the peak frequencies,

\[ \phi_{\text{esc}} = \frac{F_{\text{esc}}/F_{\text{hom}}}{F_{\text{esc}}/F_{\text{hom}} + \pi^2 J_{\text{slab}}(\tau_0, \phi_{\text{esc}} J_{\text{centre}}(\tau_0, x))}. \]  \hspace{1cm} (35)

(iii) Division between porous and random-walk regime:
As the optical depth of the already optically thick clumps, \(\tau_0, \text{cl}\), exceeds the optical depth of the slab, \(\tau_0\), a fraction of the Ly\(\alpha\) photons traverse the slab without scattering, leaving the random and entering the porous regime. This transition occurs as the number of clumps encountered by the Ly\(\alpha\) photons becomes less than unity, \(f_c < 1\).

3.1.3 Ionising escape fraction dependent in a clumpy/homogeneous medium
For this Ly\(\alpha\) line profile model, we assume a model similar to the so-called picket fence model (Heckman et al. 2011). Here a fraction \(f_{\text{esc}}\) of the ionising radiation escapes through low-density channels, while the other fraction of ionising photons is absorbed by the dense shell. Corresponsingly, the Ly\(\alpha\) photons escaping through the channels scatter only a few times, while those escaping through the shell encounter many scattering events. For optically thin channels, the former gives rise to a single-peaked Ly\(\alpha\) line centred around \(x = 0\), while the latter creates a broader double-peaked Ly\(\alpha\) line (assuming a homogeneous slab model with peaks at \(x_{\text{esc}}\)).

Here we assume the channels to be fully ionised and a fraction \(f_{\text{esc}}\) of the Ly\(\alpha\) photons to escape through them without scattering. As a consequence, the other fraction of Ly\(\alpha\) photons encounter the shell with an optical depth of

\[ \tau_{\text{shell}} = \frac{\tau_0}{1 - f_{\text{esc}}}, \]  \hspace{1cm} (36)

at their first scattering. \(\tau_0\) is the optical depth as derived in Section 3.2. We make the simplifying assumption that those photons traverse the shell without being scattered into the channels. In reality a fraction of these photons encounter a lower optical depth when traversing the empty channels on their scattering path out of the slab, leading to a profile closer centered around \(x = 0\). Thus, the emerging Ly\(\alpha\) line profile, which we assume to be

\[ J(\tau, x) = f_{\text{esc}} J_{\text{slab}}(T, 0, x) + (1 - f_{\text{esc}}) J_{\text{shell}}(T, \tau_{\text{shell}}, x) \]  \hspace{1cm} (37)

\[ J_{\text{shell}} = f_{\text{shell}} J_{\text{slab}}(T, 0, x) + (1 - f_{\text{shell}}) J_{\text{slab}}(T, \tau_{\text{shell}}, x) \]  \hspace{1cm} (38)

represents a lower limit to the fraction of Ly\(\alpha\) photons escaping close to the resonance. We derive \(f_{\text{shell}}\) by choosing a clump optical depth \(\tau_{0, \text{cl}}\) and use Eqn 31 and 35 for the random-homogeneous \((k a\tau_0 > \sqrt{\pi} x_*)\) or porous-homogeneous \((k a\tau_0 < \sqrt{\pi} x_*)\) transitions.

3.2 Optical depth and \(\text{H}\!\text{I}\) column density
To derive the Ly\(\alpha\) line profile emerging from a simulated galaxy for the models described in Sections 3.1.2 (Clumpy model) and 3.1.3 (Porous model), we yield the optical depth at the Ly\(\alpha\) line centre from our simulated galaxies as

\[ \tau_0 = \frac{4}{\pi} f_c \tau_{0, \text{cl}} = N_{\text{HI}} \sigma_{\text{HI}}. \]  \hspace{1cm} (39)

\(\tau_{0, \text{cl}}\) is a free parameter in our model and reflects the optical depth of a dense clump in the ISM. We will use this parameter to calibrate our model to the observed Ly\(\alpha\) LFs at \(z = 6.6 - 8\) in Section 4. To obtain a rough estimate of \(\tau_{0, \text{cl}}\), we consider the median mass \((M_{\text{cl}})\) and size \((r_{\text{cl}})\) of molecular clouds \((M_{\text{cl}} \approx 10^3 M_\odot\) and \(r_{\text{cl}} = 20\) pc) resembling the mass and size of possible dense structures in the ISM,

\[ \tau_{0, \text{cl}} = \sigma_{\text{HI}} M_{\text{cl}} / r_{\text{cl}} \]  \hspace{1cm} (40)

To obtain the optical depth \(\tau_0\), we derive the neutral hydrogen column density \(N_{\text{HI}}\) from the initial gas mass, \(M_g\) as

\[ N_{\text{HI}} = \frac{\xi}{4 \pi r_g^2 m_H} = \frac{3X_c (1 - Y) M_g^2}{4 \pi (4.5 \lambda_{\text{vir}})^3 m_H} = \xi \frac{3 f_m M_{\text{vir}}}{8 \pi \lambda_{\text{vir}}^2 m_H} \]  \hspace{1cm} (41)

Here \(r_g\) describes the gas radius, for which we assume \(r_g = 4.5 \lambda_{\text{vir}}\). \(X_c\) and \(Y\) are the cold gas and helium mass fractions, respectively. Gas accretion and SN feedback processes determine the relation between the initial gas mass and the halo mass, \(f_m\), which ranges typically between \(~10^{-3}\) for low-mass galaxies to \(~10^{-1}\) for more massive galaxies. \(\xi\) is a geometrical correction factor that depends on \(\tau_0\) and the dust optical depth at the Ly\(\alpha\) resonance \(\tau_d\). Its maximum values is 0.35 and we describe its derivation and dependencies in Appendix A. For the cosmological parameters assumed in this paper, we yield

\[ N_{\text{HI}} = 6.5 \times 10^{17} \text{cm}^{-2} (1 + z)^2 \frac{\xi f_m}{\lambda^2} \left( \frac{M_{\text{vir}}}{10^9 M_\odot} \right)^{1/3}. \]  \hspace{1cm} (42)

3.3 Dust attenuation
We employ two different dust models. The first one links the Ly\(\alpha\) escape fraction to the escape fraction of UV continuum photons, \(f_{\text{esc}}\). The second one is more complex. It assumes a clumpy medium where the attenuation of Ly\(\alpha\) by dust follows different relations in the regimes identified in Gronke (2017). Both models assume a slab-like geometry and we describe their details in the following.

3.3.1 Simple attenuation model
In this model, we assume that (i) dust and gas are perfectly mixed, (ii) the dust distribution is slab-like, and (iii) the dust attenuation of Ly\(\alpha\) photons is proportional to the dust attenuation of UV continuum photons. The escape fraction of Ly\(\alpha\) photons, \(f_{\text{esc}}^{\text{Ly}\alpha}\), is then directly related to the escape of UV continuum photons, \(f_{\text{esc}}\), derived in Section 2.2.2.

\[ f_{\text{esc}}^{\text{Ly}\alpha} = p f_{\text{esc}} \]  \hspace{1cm} (43)

We use \(p\) as a free parameter to obtain the observed Ly\(\alpha\) luminosity functions at \(z = 6.6 - 7.3\).

3.3.2 Refined attenuation model
This model assumes that dust and gas are perfectly mixed and distributed in clumps. The dust attenuation of Ly\(\alpha\) photons depends on
the total optical depth of the dust, $\tau_{d,\text{total}}$, the optical depth of a clump, $\tau_{d,\text{cl}}$, and the number of clumps, $f_c$, encountered along the sightline from the midplane to the surface of the slab. We derive its value by estimating the dust absorption cross section. Following Galliano (2022) and assuming the radius and density of graphite/carbonaceous grains (see Section 2.2.2), we assume $\kappa_{\text{abs}} = \frac{Q_{\text{abs}}}{\alpha_s} = 2 \times 10^5 \, \text{cm}^2/\text{g}$ with $Q_{\text{abs}} = 1$ being the absorption efficiency.

$$\tau_{d,\text{total}} = \frac{4}{3} f_c \tau_{d,\text{cl}} = \kappa \frac{M_d}{\tau_d} \frac{M_H}{\sigma_H} = \kappa \frac{M_d}{\tau_d} \frac{M_H}{\sigma_H} \tau_0 \quad (44)$$

The resulting estimates for $\tau_{d,\text{total}}$ and $\tau_{d,\text{cl}}$ allow us to compute the Lyα escape fractions in the different escape regimes as follows.

### 3.3.2.1 Free-streaming regime:
In an optically thin slab ($\tau_0 < 1$), the Lyα photons stream through $\sim f_c$ clumps. On their way, they are attenuated by the dust in clumps and hence, the total dust optical depth determines the Lyα escape fraction, $\tau_{d,\text{total}}$, as

$$f_{\text{Lyα,fs}}^\text{esc} = \exp(-\tau_{d,\text{total}}) = \exp\left(-\frac{4}{3} f_c \tau_{d,\text{cl}}\right) \quad (45)$$

We note that in this regime, the number of clumps along the sightline $f_c$ and clump optical depth $\tau_0$ are degenerate.

### 3.3.2.2 Random walk regime:
In the random walk regime, both the slab and individual clumps are optically thick ($\tau_{d,\text{cl}} \geq 1$). As a result, Lyα photons escape by mostly being scattered by the clumps, and their escape fraction is determined by the number of clumps encountered along their random walk, $N_c(f_c)$, and the absorption probability per clump interaction $\epsilon$. According to Hansen & Oh (2006), it is then given by

$$f_{\text{Lyα,rw}}^\text{esc} = f_{\text{Hon6}} = \frac{1}{\cosh(\sqrt{2N_c(f_c)} \epsilon)} \quad (46)$$

We assume $N_c(f_c) \approx \frac{3}{2} f_c^2 + 2 f_c$ as found in Gronke (2017). The scaling of $N_c$ with $f_c$ also agrees with the findings in Hansen & Oh (2006) and prefactors vary slightly due to different geometries of the scattering surface. However, since $\epsilon$ is sensitive to how deep the photons permeate the clump, it depends non-trivially on the clump optical depth and movement. For simplicity, we assume $\epsilon = 1 - \exp(-\tau_{d,\text{cl}}^2/10)$, which we have found to be in rough agreement with the fits and results shown in Gronke (2017).

### 3.3.2.3 Homogeneous regime:
In the homogeneous regime, the slab is optically thick ($\tau_0 \geq 1$), while the individual clumps are optically thin at the escape frequencies ($\tau_{d,\text{cl}}(f_{\text{esc}}) < 0$). During their initial random walk, the Lyα photons scatter with $N_c(f_{\text{esc}},\tau_{\text{crit}})$ clumps before they diffuse into the wings and escape by free-stream through $f_c$ clumps. The resulting Lyα escape fraction

$$f_{\text{Lyα,hom}}^\text{esc} = f_{\text{Hon6}}(f_c,\tau_{\text{crit}}) \exp(-\tau_{d,\text{total}}) \quad (47)$$

depends on $f_c, \tau_{\text{crit}}$, and $\tau_{d,\text{total}}$, with $f_c, \tau_{\text{crit}}$ being determined by $\tau_0$ and $\tau_{d,\text{cl}}$.

$$f_{\text{Lyα,crit}}^\text{esc} = \begin{cases} \frac{2}{\sqrt{\pi} \sigma_{\text{abs}}(f_c,\epsilon)} \sqrt{\tau_{\text{crit},0}} & \text{for } k \tau_0 \geq \sqrt{\pi}^2 \max \kappa \left(1 - \epsilon f_{\text{esc},\text{total}}^\text{crit}\right) \\ \frac{4}{3} f_c \tau_{d,\text{cl}} & \text{for } k \tau_0 < \sqrt{\pi}^2 \max \kappa \end{cases} \quad (48)$$

### 3.3.2.4 Porous regime:
In the porous regime, the individual clumps are optically thick ($\tau_{d,\text{cl}} \geq 1$), but only a fraction $1 - e^{-f_c}$ of the Lyα photons will encounter a clump along their sightlines. The other fraction of Lyα photons does not interact with any clumps and is thus not attenuated by dust as they escape the slab.

$$f_{\text{Lyα,por}}^\text{esc} = e^{-f_c} + \left[1 - e^{-f_c}\right] \exp\left(-\frac{4}{3} f_c \tau_{d,\text{cl}}\right) \quad (49)$$

### 3.3.3 Emerging Lyα line profile models
We briefly summarize the combinations of Lyα line and dust attenuation models that we will investigate in this paper.

**Gaussian:** The Lyα line profile emerging from a galaxy is given by the central Gaussian Lyα line profile (Section 3.1.1). To account for the attenuation by dust, we apply the Lyα escape fraction, $f_{\text{esc}}$, derived in our simple dust model (Section 3.3.1) to all frequencies $x$.

**Clumpy:** This model assumes a slab of dusty gas clumps, whereas gas and dust are perfectly mixed. It combines the Lyα line model described in Section 3.1.2 with the refined dust model depicted in Section 3.3.2. The gas in the galaxies is assumed to have a temperature of $T = 10^4 \, \text{K}$. In contrast to the Gaussian model, we dust-attenuate the Lyα line of each escape regime (homogeneous, random, porous) by its corresponding escape fraction $f_{\text{esc}}$. The emerging Lyα line profile is then the superposition of the line profiles of all relevant escape regimes.

**Porous:** This model is very similar to the Clumpy model. However, it considers the shell of clumps to be pierced with gas and dust-free channels through which a fraction $f_{\text{esc}}$ of the Lyα photons escape without scattering. It combines the Lyα line model described in Section 3.1.3 and assuming $f_{\text{channel}} = 0$ with the refined dust model depicted in Section 3.3.2. Again we assume the gas in the galaxy to be heated to a temperature of $T = 10^4 \, \text{K}$, and the Lyα line of each escape regime (homogeneous, random, porous) to be dust-attenuated by its corresponding escape fraction $f_{\text{esc}}$. The emerging Lyα line profile

---

5 We note that our expression is here a lower limit of $f_{\text{esc}}$, as we assume the Lyα radiation interacting with clumps to experience attenuation as if they streamed through the clump. It might be more appropriate to consider these Lyα photons to be absorbed as in the random walk regime, $f_{\text{esc}} = e^{-f_c} + \left[1 - e^{-f_c}\right] \cdot \frac{1}{\cosh(\sqrt{2N_c(f_c)} \epsilon)}$, however, in practise galaxies in the porous regime have not much, if any, dust.

6 We have chosen $T = 10^4 \, \text{K}$ for simplicity. If we were to assume the virial temperature ($T_{\text{vir}}$), the double-peak line profile would narrow as $T_{\text{vir}}$ increases.
profile is again a superposition of the Lyα photons escaping through the channels and the clumpy shell,
\[ L_\alpha^{gal}(x) = f_{esc} J_{channel}(x) + (1 - f_{esc}) J_{shell}(x) \]  
with \( f_{esc} \) denoting the fraction of radiation at frequency \( x \) given by
\[ f_{esc} = \int_{z_{min}}^{z_{max}} \sigma_0 \phi(x + x_p(r(z))) \frac{n_{HI}(r(z))}{(1 + z)H(z)} dz. \]

3.4 IGM attenuation

The Lyα radiation escaping from a galaxy is attenuated by the H I it encounters along the line of sight from the location of emission, \( r(z_{em}) \), to the location of absorption, \( r(z_{obs}) \). Expressing the frequency \( x \) of a photon in terms of its rest-frame velocity \( v \) relative to the Lyα line centre, the transmitted fraction of radiation at frequency \( x \) is given by
\[ T_{\alpha,x}(x) = \exp\left( -\tau_{\alpha}(x) \right) \]  
where \( \tau_{\alpha} = c \int_{z_{em}}^{z_{max}} \sigma_0 \phi(x + x_p(r(z))) \frac{n_{HI}(r(z))}{(1 + z)H(z)} dz \).

Here \( \tau_{\alpha} \) describes the optical depth to Lyα, while \( n_{HI}(r) \) and \( v_p(r) = b x_p(r) \) the H I density and peculiar velocity (in the rest-frame of the emitted Lyα radiation) at a physical distance \( r \) from the emitter, respectively. \( \sigma_0 \) is the specific absorption cross section, described in the cgs system as
\[ \sigma_0 = \frac{\pi e^2 f}{m_e c^2} = \frac{3.54 \times 10^{-19}}{x_\alpha}. \]

and Lorentzian damping wings
\[ \phi_{\text{Lorentz}}(x) = \frac{A_{21}^2}{4 \pi^2 (x b)^2 + \frac{1}{2} A_{21}^2 x^2}. \]

Finally, we derive the observed, i.e. dust and IGM attenuated, Lyα luminosity and line profile along each major axis (resulting in 6 lines of sight) as
\[ L_{\alpha,x}(x) = L_\alpha^{gal}(x) T_{\alpha,x}(x) = L_{\alpha,\text{true}} \frac{f_{esc} \phi_{\text{esc}}^L J_{\alpha}}{f_{esc,\text{cen}} J_{\alpha}}, \]

for temperatures between \( T = 0.01 \)K and \( 10^8 \)K.

Our calculations of \( T_{\alpha} \) include the Hubble flow and peculiar velocities \( v_p \): outflows (inflows) of gas from a galaxy that correspond to positive (negative) \( v_p \) values will redshift (blueshift) the Lyα photons and lead to an increase (decrease) in \( T_{\alpha} \). For each galaxy in a simulation snapshot we derive \( T_{\alpha} \) along all directions along the major axes (i.e. along and against the x, y and z axes). By stepping through the simulation box that is divided into 512 cells on the side (and each cell having a size of 461ckpc), we derive the \( n_{HI}(r) \) and \( v_p(r) \) profiles from the ASTRAEUS ionisation and vsm0 density and velocity grids. For any galaxy, we start the profiles at the galaxy position \( r_{em} = 0 \) and end them once the highest frequency \( x_{\text{max}} = x_{\text{max,gal}}/b = 40 \) tracked in our Lyα line profiles has redshifted out of absorption at \( r = x_{\text{max}}/|H_0| \Omega_m^{1/2}(1 + z)^{1/2} \approx 13.6(1 + z)^{1/2} \)Mpc. We assume \( T_{\alpha} \) for the Lyα line redshifts out of resonance very quickly (the light travel time for distance \( r \) at \( z = 7 \) is less than 2 Myrs, shorter than the simulation time steps), a single simulation snapshot suffices for computing the \( T_{\alpha} \) values of the galaxies in that snapshot. We also assume periodic boundary conditions when computing \( T_{\alpha} \).

Finally, we derive the observed, i.e. dust and IGM attenuated, Lyα luminosity and line profile along each major axis (resulting in 6 lines of sight) as
\[ L_{\alpha,x}(x) = \int_{-\infty}^{\infty} L_{\alpha,x}(x) \, dx \]

where \( f_{esc} \) and \( J \) are the respective Lyα escape fraction and line profile for a one of the models as outlined in Section 3.3.3. The total observed Lyα luminosity \( L_\alpha \) and total fraction of Lyα radiation transmitted through the IGM are yielded when integrating the respective quantity over the frequency \( x \).

\[ L_{\alpha} = \int_{-\infty}^{\infty} L_{\alpha,x}(x) \, dx \]
\[ T_{\alpha} = \int_{-\infty}^{\infty} T_{\alpha,x}(x) \, dx \]

In the following, we use all lines of sight as independent probes when line-of-sight-sensitive Lyα quantities are analysed.

We derive the observed Lyα luminosities \( L_{\alpha} \) for all galaxies at \( z = 20, 15, 12, 10, 9, 8, 7, 3, 7 \) and 6.6 for any combination of emerging Lyα line model (Gaussian, Clumpy, Porous) and reionisation scenario (mHDEC, mhconst, mHINC). Free model parameters \( p \) for the Gaussian model, \( \tau_{0,\text{cl}} \) for the Clumpy and Porous models) have been chosen to visually best-fit the observed Lyα LFs at \( z \approx 6.7, 7.0 \) and 7.3 (see Tab. 2). For simplicity and better comparison we assume in all models the gas in galaxies to have the temperature of photo-ionised gas, \( T = 10^4 \) K. Moreover, we note that since the mhconst scenario represents an intermediate case and provides no further insights, we limit our discussion to the mHDEC and mHINC scenarios in the remainder of this paper.

| Parameter | Scenario | Gaussian | Clumpy | Porous |
|-----------|----------|----------|--------|--------|
| \( \tau_{0,\text{cl}} \) | mHDEC | \( 1.2 \times 10^6 \) | \( 2.4 \times 10^6 \) | |
| \( \tau_{0,\text{cl}} \) | mHINC | \( 5 \times 10^5 \) | \( 1.8 \times 10^6 \) | |
| \( p \) | mHDEC | 1.0 | | |
| \( p \) | mHINC | 1.4 | | |
| \( T \) | all | \( 10^4 \) K | \( 10^4 \) K | \( 10^4 \) K |
4 NUMBERS AND PROPERTIES OF LY\(\alpha\) EMITTING GALAXIES

In this Section, we aim to identify which physical process – the intrinsic Ly\(\alpha\) production \(\left(L_{\text{int}}^{\alpha,\gamma}\right)\), the absorption by dust within the galaxies \(f_{\text{esc},\alpha}^\gamma\), or the scattering by \(\text{H}\,\text{I}\) in the IGM \(T_{\alpha,x}\) – dominates the observed Ly\(\alpha\) emission. To this end, we analyse (i) how the IGM attenuation profile \(T_{\alpha,x}(x)\) depends on galaxy mass and the \(f_{\text{esc},\alpha}\) sensitive ionisation topology, (ii) how the Ly\(\alpha\) line profiles emerging from a galaxy depend on the density and velocity distributions of gas and dust within a galaxy and \(f_{\text{esc},\alpha}\), and how much it affects the fraction of Ly\(\alpha\) radiation that is transmitted through the IGM, and (iii) to which degree the \(f_{\text{esc},\alpha}\) dependency of \(I_{\alpha,\gamma}^\alpha\), \(f_{\text{esc},\alpha}\), and \(T_{\alpha,x}\) leave characteristic imprints in the Ly\(\alpha\) luminosity functions and the population emitting visible Ly\(\alpha\) emission.

4.1 The transmission through the IGM

We start by discussing the frequency-dependent IGM transmission \(T_{\alpha,x}\) shown in the top row of Fig. 3. These profiles depend solely on the underlying ionisation topology and density distribution of the IGM. From the different panels depicting the average \(T_{\alpha,x}\) in different halo mass bins of width \(\Delta \log M_\text{h} = 0.125\), we see that all \(T_{\alpha,x}\) profiles follow a common trend: \(T_{\alpha,x}\) decreases towards higher frequencies with an stronger decline around the Ly\(\alpha\) resonance \((x = 0)\). Photons bluewards the Ly\(\alpha\) resonance redshift into the Ly\(\alpha\) resonance as they propagate through the IGM and have the largest likelihood to be absorbed by the \(\text{H}\,\text{I}\) present. Photons redwards the Ly\(\alpha\) resonance are also redshifted, but their likelihood of being absorbed by \(\text{H}\,\text{I}\) decreases significantly as their energy drops.

In each panel in the top row of Fig. 3 we show \(T_{\alpha,x}\) for the two reionisation scenarios \(\text{mHinc}\) (yellow/orange/brown lines) and \(\text{mHinc}\) (blue lines) and redshifts \(z = 8.0, 7.3, 7.0, 6.6\) (bright to dark lines as redshift decreases). In general, i.e. for both reionisation scenarios and all halo masses, \(T_{\alpha,x}\) increases as the ionised regions grow around galaxies and the IGM is increasingly ionised (bright to dark lines): firstly, a larger ionised region shifts not only the point of strongest Ly\(\alpha\) absorption to higher frequencies \(x\) but also reduces the absorption in the damping wings of the Ly\(\alpha\) absorption profile. Secondly, lower \(\text{H}\,\text{I}\) fractions in ionised regions diminish the number of \(\text{H}\,\text{I}\) atoms absorbing Ly\(\alpha\) photons.

These two mechanisms shape \(T_{\alpha,x}\) redwards and bluewards the Ly\(\alpha\) resonance. As Ly\(\alpha\) photons travel through the IGM and redshift, photons emitted at frequencies \(x \geq 0\) see the Gaussian core of the Ly\(\alpha\) absorption profile \(\phi(x)\) and are absorbed by \(\text{H}\,\text{I}\) abundances as low as \(\chi_{\text{HI}} \gtrsim 10^{-4}\); thus they are sensitive to the residual \(\text{H}\,\text{I}\) fraction in ionised regions. Correspondingly, we see in Fig. 3 that \(T_{\alpha,x}\) increases for \(x \geq 0\) with decreasing redshift as the photoionisation rate around galaxies increases and lowers the residual \(\text{H}\,\text{I}\) fraction in ionised regions. However, photons emitted at frequencies \(x \leq 0\) are absorbed by the damping wings of the Ly\(\alpha\) absorption profile \(\phi(x)\). Since the Ly\(\alpha\) absorption cross section is lower in the damping wings, the abundance of \(\text{H}\,\text{I}\) needs to be significantly higher for Ly\(\alpha\) photons to be absorbed; thus, as the sizes of ionised regions decrease, photons emitted at these frequencies are increasingly absorbed by the neutral regions located beyond the ionised regions around the emission sites.

For this reason, we find \(T_{\alpha,x}\) for \(x \leq 0\) to increase as the sizes of the ionised regions around galaxies rise with increasing halo mass and decreasing redshift. The rising sizes of ionised regions also become manifest in the shift of the frequency at which \(T_{\alpha,x}\) has a value of 0.5 to higher frequencies.

Its dependence on the size of the ionised regions around galax-
Figure 3. Intrinsic (top) and observed (bottom) Ly$\alpha$ line profile and IGM transmission (centre) at $z = 8.0$, 7.3, 7.0, 6.6 for a homogeneous static gas shell (left), a clumpy gas shell (centre), and a clumpy gas shell with holes through which Ly$\alpha$ radiation escapes without scattering. Solid (dashed dotted) lines show results for the reionisation scenario where $f_{\text{esc}}$ decreases (increases) with halo mass $M_h$. 

$M_h \simeq 10^{9.0} M_\odot$  $M_h \simeq 10^{10.0} M_\odot$  $M_h \simeq 10^{11.0} M_\odot$

$\nu$ [km/s]  $\nu$ [km/s]  $\nu$ [km/s]
Lyα luminosities emerging from the galaxies and fit the observed Lyα LFs in each of our reionisation scenarios. In the \textit{mhinc} scenario the more massive galaxies – that dominate the observed Lyα LF – have higher $f_{\text{esc}}$ values than in the \textit{mhdec} scenario; to compensate the corresponding lower $L_{\alpha}^{\text{int}}$ values (and steeper slope of the intrinsic Lyα LF), we need a higher $f_{\text{esc}}^\text{Lyα}$/$L_{\alpha}^{\text{int}}$ ratio (1.4) than in the \textit{mhdec} scenario (1.0). Despite this compensation, the slopes of the observed Lyα LFs at $z \lesssim 8$ (c.f. left panel in Fig. 4) is still steeper for the \textit{mhinc} than for the \textit{mhdec} scenario.

4.2.2 The Clumpy model

In the \textit{Clumpy} model, the clumpiness of the gas in the shell and the attenuation by dust molecules in these clumps determine the shape of the Lyα line profile. We note that in the following discussion the number of clumps in the dusty gas shell, i.e. a higher clumpiness corresponds to fewer clumps and thus a higher ratio between the clump ($\tau_{0,\text{cl}}$) and total line optical depth ($\tau_0$). We find the following characteristic trends for the Lyα line profile: Firstly, the clumpier the gas in the shell is, the more Lyα radiation escapes around the Lyα resonance (profile showing a central peak), and the fewer Lyα photons escape through excursions or via the wings (double peak profile). Secondly, when assuming the same clump size for all galaxies – as we do in this paper – the gas clumpiness decreases as galaxies become more massive and contain more gas. Thus, from low-mass to more massive galaxies, we find the Lyα line profile to shift from a central peak dominated to a double-peak dominated profile (see the fourth row in Fig. 3 from left to right), reflecting the transition from the random to the homogeneous regime (see Section 3.1.2). This transition also goes in hand with an increased transmission through the IGM, which we can see when comparing the Lyα profiles emerging from galaxies (fourth row) with those after having traversed the IGM (fifth row in Fig. 3). The Lyα luminosity at $z = 0$ decreases by $\sim 0.5$ orders of magnitude for all halo masses (from $10^{41.1}$ erg s$^{-1}$ to $10^{41.1}$ erg s$^{-1}$ for $M_h \approx 10^{11} M_\odot$ and from $10^{39.7}$ erg s$^{-1}$ to $10^{39.2}$ erg s$^{-1}$ for $M_h \approx 10^9 M_\odot$ for e.g. \textit{mhdec} model), while the peak Lyα luminosity of the red wing decreases only about $\lesssim 0.3$ orders of magnitude at halo masses. While the blue wing is similarly or more attenuated than the central peak in the IGM, the total fraction of Lyα radiation transmitted through the IGM for a fully-double peaked profile exceeds that of profiles with a central peak component. Furthermore, as the galaxies’ gravitational potentials flatten with decreasing redshift, $\tau_0$ decreases and leads to (i) a narrower double-peak profile (following the dependence of the peak position on $\tau_0^{1/3}$) and (ii) a stronger central peak (the gas becomes clumpier as the ratio $\tau_{0,\text{cl}}/\tau_0$ increases).

A change in the clumpiness of the gas and dust shell (or clump optical depth $\tau_{0,\text{cl}}$ and $\tau_{d,\text{cl}}$) goes not only in hand with a change in the Lyα profile affecting $\tau_0$ but also an altered attenuation of the escaping Lyα radiation by dust. Thus, adjusting the clump optical depth allows us to enhance and reduce the Lyα luminosities and reproduce the observed Lyα LFs: As we increase the size of the clumps, i.e. increase $\tau_{0,\text{cl}}$ Lyα photons will scatter with fewer clumps, leading to (i) a higher fraction $f_{\text{esc}}^\text{Lyα}$ escaping, and (ii) a higher fraction escaping at the Lyα resonance, which again leads to stronger attenuation by H I in the IGM. However, we note that once the emerging Lyα profile is fully double-peak, the attenuation by H I in the IGM can not be further decreased (by changing the injected Lyα line profile). The observable Lyα emission can only be enhanced by decreasing $\tau_{0,\text{cl}}$ as long as the $f_{\text{OH06}}(f_{\text{esc}})$ factor in $f_{\text{esc}}^\text{Lyα, hom}$ remains significantly below unity. With the observed Lyα LF being dominated by the more massive galaxies ($M_h \gtrsim 10^{10} M_\odot$, as we will discuss in the next Section), we find the $\tau_{0,\text{cl}}$ value to reflect the factor by which the bright end of the intrinsic Lyα LF needs to be reduced to reproduce the observational Lyα LF data points (filled points in Fig. 4). As the intrinsic Lyα LFs is lower at the bright end in the \textit{mhinc} scenario, a lower $\tau_{0,\text{cl}}$ value ($5 \times 10^5$) is required than for the \textit{mhdec} scenario ($1.2 \times 10^6$). Nevertheless, the slopes of the resulting observed Lyα LFs at $z \lesssim 8$ keep the trends of the intrinsic Lyα LFs, with the bright ends of the Lyα LFs being steeper in the \textit{mhinc} than in the \textit{mhdec} scenario.
4.2.3 The Porous model

The Porous model represents a refinement of the Clumpy model. It adds gas-free channels through which Lyα and ionising photons escape freely. This explains why, to first order, we find the trends in the last two rows of Fig. 3 to be similar to those in the fourth and fifth rows: a lower clumpiness of gas and dust in the shell induces a stronger prevalence of the double-peak component in the Lyα line profile emerging from a galaxy, enhancing the IGM transmission $T_\alpha$ and absorption by dust within the galaxy, and causing the corresponding Lyα LFs to shift to lower values. On the other hand, it differs from the Clumpy model substantially, as $f_{\text{esc}}$ determines the minimum fraction of Lyα radiation that escapes at the Lyα resonance and contributes to the central peak in our modelling. Hence, as long as $\tau_{0,cl}$ remains above the $\tau_{0,cl}$ value that leads to the same fraction of Lyα escaping in the central peak than given by $f_{\text{esc}}$ (referred to as $\tau_{0,cl}$ in the following), the Porous model inherits the trend of the Clumpy model. As $\tau_{0,cl}$ drops below $\tau_{0,cl}$, a further decrease in $\tau_{0,cl}$ affects the Lyα line profile emerging from a galaxy hardly, and once the $f_{\text{1000}}(f_{\text{esc}},\tau_{0,cl})$ factor of $f_{\text{esc}}$ approaches unity, the observed Lyα LFs remain “fixed”. The resulting upper limit of $f_{\text{esc}}$ (determined by the total dust optical depth $\tau_{\text{total}}$) is essential, as together with $I_{\alpha}^{\text{intr}}$ it provides an upper limit to $f_{\text{esc}}$ values that fit the observed Lyα LFs. We find this upper limit to be about $f_{\text{esc}} \sim 0.5$ in our Astraeus model.

Due to their opposing dependencies of $f_{\text{esc}}$ with halo mass, the Lyα profiles in the Porous model show the most noticeable differences between the MHDEC and MHNHC scenarios among our three Lyα line profile models. While the double-peak component is more prominent in the most massive galaxies ($M_h \approx 10^{10} M_\odot$) in the MHDEC scenario, the central peak is slightly stronger in the MHNHC scenario. To fit the observed Lyα LFs, we find that we require for both reionisation scenarios a more clumpy gas and dust distribution than in the Clumpy model, i.e. a (higher) $\tau_{0,cl}$ value of $1.8 - 2.4 \times 10^5$. These increased $\tau_{0,cl}$ values enhance the corresponding $f_{\text{esc}}$ values and thus the dust attenuation in the homogeneous regime giving rise to the double-peak components and counteract the increased escape close to the Lyα resonance. This model-integrated correlation between $f_{\text{esc}}$ and $f_{\text{esc}}$ counteracts the trend of flattening (steepening) the slope of the intrinsic Lyα LFs due to $f_{\text{esc}}$ decreasing (increasing) with rising halo mass: If $f_{\text{esc}}$ is low (high), more (less) Lyα radiation is subject to dust attenuation. This model feature explains why the observed Lyα LFs of the MHNHC simulation are shallower than in the Clumpy model and hardly changes for the MHDEC simulation due to its low $f_{\text{esc}}$ values for more massive galaxies.

As the dust composition and absorption cross section for Lyα remain highly uncertain during the EoR, we note that a lower (higher) dust absorption cross section $\kappa_{\text{abs}}$ could still reproduce the observed Lyα LFs in our Clumpy and Porous models by raising (decreasing) the clump optical depth $\tau_{0,cl}$. However, this would go along with an enhanced (reduced) double-peak and reduced (enhanced) central-peak component in the average Lyα line profile emerging from galaxies.

Finally, we briefly comment on how our emerging and IGM-attenuated Lyα profiles compare to those obtained from radiative hydrodynamical simulations of clouds and small cosmological volumes ($\sim 10^3\text{Mpc}^3$). While the Clumpy and Porous reproduce the double- and triple-peak profiles and their dependence on $N_{\text{HI}}$ and $f_{\text{esc}}$ found in cloud simulations (Kakichi & Gronke 2021; Kimm et al. 2019, 2022) by construction, their Lyα line profiles differ from those obtained from the SPHINX simulation (Garel et al. 2021). In SPHINX the median angle-averaged Lyα line profile has been found to be less double-peak towards brighter galaxies, with the blue peak being seemingly increasingly suppressed. This is the opposite trend of our findings. The discrepancy lies in the differently assumed or simulated ISM structures: While our LAE models assume an idealised scenario of same-sized dusty gas clumps, the SPHINX simulation follows the formation of star-forming clouds within galaxies. With rising galaxy mass, we expect the simulated SPHINX galaxies to contain a higher number of star-forming clouds with various velocity and size distributions. A single or very few star-forming clouds – as found in low-mass galaxies – will give rise to a double-peak Lyα line profile. Adding the profiles of multiple/more star-forming clouds at different velocities will give rise to increasingly more complex Lyα line profiles as galaxies become more massive. Adjusting our current Lyα line profile models to the complex structure of the ISM will be the subject of future work.

4.3 The dependence of Lyα properties on halo mass

In this Section, we provide a more detailed discussion of how the intrinsic Lyα luminosity ($L_{\alpha}^{\text{intr}}$), the Lyα escape fraction, the Lyα transmission through the IGM, the observed Lyα luminosity, and Lyα equivalent width depend on halo mass and evolve with redshift for the different reionisation scenarios. To this end, we show these quantities as a function of halo mass for both reionisation scenarios (MHDEC: yellow/orange/brown lines; MHNHC: blue lines) and redshifts $z = 8, 7.3, 7, 6.6$ in Fig. 5 and list the corresponding average H I fractions in Table 3. Solid and dot-dashed lines in Fig. 5 depict the median value for galaxies in the given halo mass bin, and shaded regions indicate the range spanned by $68\%$ of the values. For line-of-sight-dependent Lyα properties ($T_\alpha$, $L_{\alpha}$, $EW_\alpha$), we include all 6 lines of sight.

Intrinsic Lyα luminosity $L_{\alpha}^{\text{intr}}$: As the most recent star formation dominates the production of ionising photons within galaxies, we find $L_{\alpha}^{\text{intr}}$ to follow the SFR-$M_h$ relation (for a detailed discussion, see Hutter et al. 2021a). While the range of SFR values is broad for low-mass halos ($M_h \lesssim 10^{9.5} M_\odot$) where SN feedback drives stochastic star formation, the SFR-$M_h$ relation becomes tighter towards more massive galaxies as SN feedback ejects an increasingly lower fraction of gas from the galaxy. Being mainly produced by recombinig hydrogen atoms within a galaxy, the Lyα radiation produced within the galaxy correlates with the escape fraction of ionising photons as $1 - f_{\text{esc}}$. As we can see from the first row in Fig. 5, this dependency on $f_{\text{esc}}$ leads to higher (lower) Lyα luminosities for more massive galaxies, lower (higher) Lyα luminosities for low-mass galaxies, and thus a shallower (steeper) LFs in the MHDEC (MHNHC) scenario.

Lyα escape fraction $f_{\text{esc}}$: As the dust content in galaxies increases with their mass, we find $f_{\text{esc}}$ to decrease with rising halo mass at all redshifts and for all Lyα line models. However, the different assumed distributions of dust and their resulting attenuation of Lyα radiation lead to differences in the details of this global trend: Firstly, the Gaussian model shows a steeper decline in $f_{\text{esc}}$ for galaxies with

| $z$ | $\langle f_{\text{H I}}^{\text{MHINC}} \rangle$ | $\langle f_{\text{H I}}^{\text{MHDEC}} \rangle$ |
|-----|-------------------------------|-------------------------------|
| 8.0 | 0.84                          | 0.71                          |
| 7.3 | 0.69                          | 0.59                          |
| 7.0 | 0.52                          | 0.49                          |
| 6.6 | 0.23                          | 0.34                          |

Table 3. The evolution of the global H I fractions of the IGM for our reionisation scenarios.
Figure 5. Median of indicated galactic properties (lines) and their $\sim 1.3\sigma$ uncertainties (shaded regions) as a function of halo mass $M_h$ at $z = 8.0, 7.3, 7.0, 6.6$ for a homogeneous static gas shell. Solid (dashed dotted) lines show results for the reionisation scenario where $f_{esc}$ decreases (increases) with halo mass $M_h$. 

LAEs in the EoR
Withincreasinghalomass,galaxiesaresurroundedbylargerionisedregionsaroundthem.Mh≳10^{10.5}M⊙thantheClumpyandPorousmodels.Secondly,fescLyαisalwayshigherinthe\texttt{mhnc}thaninthe\texttt{mhdec}scenario.ThisisnecessarytoreproducetheobservedLyαLFsbycompensatingthelowintrinsicLyαluminositieswithamoreclumpygasdustdistributioninthe\texttt{mhnc}scenario.IncaseoftheClumpymodel,itisalsoshownthatadecreaseintheclumpopticaldepthbyafactor~2canincreasefescLyαbyreducingthefractionofLyαphotonsescapinginhomogeneousenvironment(i.e.aincreasedecreaseinfecentraland\tt{t}0,centralleadstoreducednumberofclumpsencountersNclandtheclumpalbedo\tt{e}).Thirdly,forthe\texttt{mhdec}(\texttt{mhnc})scenario,thefescLyαvaluesshowhigher(lower)valuesinthe\texttt{Porous}modelthanintheClumpymodelforMh≤10^{10}M⊙.Thereasonforthisdifferenceisasfollows.Inbothscenariosthehigher\tt{t}0,centralvaluesinthe\texttt{Porous}modelincreasetheattenuationofLyαescapinginhomogeneousenvironment.Butonlyafraction1−fescoftheLyαphotonsissubjecttoattenuation.ThissimulatedattenuatedescapeofLyαradiationimprintsthemass-dependencyoffescintheJLyαvalues.However, forgalaxieswithMh≥10^{10}M⊙,thisimprint(JLyα enhancementin\texttt{Porous}model)isonlyvisibleinthe\texttt{mhnc}scenariowherethesufficentlarge(>0.1);inthe\texttt{mhdec}scenariofescvaluesaretoo-small.

\textbf{LyαIGM transmission Tα}: As outlined in Section 4.1, the surrounding ionised region (in particular its size and residual H I fraction) and the Lyα line profile emerging from a galaxy determine how much of a galaxy’s escaping Lyα radiation is transmitted through the IGM.

For more massive galaxies with M_h ≥ 10^{10}M⊙, Tα is mainly shaped by the Lyα profile. This is because the ionised regions surrounding them are sufficiently large - due to their enhanced ionising emissivity and their clustered neighbourhood - for the Lyα radiation to redshift out of absorption. Hence, at these high halo masses, any trends in Tα reflect the ratio between the Lyα radiation escaping around the Lyα resonance and escaping through the wings: the more Lyα escapes in the central peak, the lower is the Tα value. Indeed, as can be seen in Fig. 5, the Gaussian model concentrating the emerging Lyα radiation around the Lyα resonance shows the lowest median Tα values at M_h ≥ 10^{10}M⊙ among all Lyα line profile models. In the Clumpy model, where the fraction of Lyα escaping through the wings increases with rising halo mass, we find the median Tα value to increase accordingly. This effect is more evident for the \texttt{mhnc} scenario as it transitions from a Lyα line profile with a dominating central peak at M_h ≈ 10^{10}M⊙ to one with a prevailing double peak component at M_h ≈ 10^{11}M⊙. The Porous model also confirms that Tα is highly sensitive to the Lyα line profile. In the \texttt{mhnc} scenario, the double peak component is weaker and increases less with halo mass, leading to slightly lower Tα values than in the Clumpy model for M_h ≈ 10^{10}−11M⊙ and Tα hardly changing with halo mass. In the \texttt{mhdec} scenario, we see the same effect but to a lower degree.

However, for less massive galaxies (M_h ≤ 10^{10}M⊙), Tα is more sensitive to the properties of their surrounding ionised regions. Since the ionised regions around less massive galaxies can differ significantly depending on their environment and phase in their stochastic star formation cycle (see Hutter et al. (2021b) and Legrand et al. (2023) for environment dependence), their Tα values span across an extensive range from as low as effectively zero to as high as ≈ 70%. Nevertheless, the median Tα value shows a definite trend. It increases with rising halo mass for all models and at all stages of reionisation. With increasing halo mass, galaxies are surrounded by larger ionised regions as they form more stars emitting ionising photons and are more likely to be located in clustered regions that are reionised earlier. The larger the surrounding ionised regions are, the higher the transmission of Lyα radiation through the IGM. We can see this relationship when comparing the median Tα values of the \texttt{mhnc} and \texttt{mhdec} simulations. In the \texttt{mhnc} scenario low-mass galaxies are surrounded by larger ionised regions at z ≳ 7 than in the \texttt{mhnc}, causing their corresponding Tα values to be raised (c.f. orange/brown solid lines vs dark blue/blue lines in the third row of Fig. 5). At z ≤ 7, however, reionisation progresses faster and the photoionisation rate in clustered ionised regions yields a lower residual H I fraction in the \texttt{mhnc} simulation, both leading to a higher median Tα value for the \texttt{mhnc} than \texttt{mhdec} scenario at z ≳ 6.6. Finally, we briefly discuss how the Lyα line profile emerging from a galaxy affects Tα for less massive galaxies. From Fig. 5 we see that the Tα values differ between our three different Lyα line profile models: While at all stages of reionisation the Tα values for M_h ≥ 10^{10}M⊙ are very similar in the \texttt{Porous} and \texttt{Clumpy} model, the \texttt{Porous} model shows lower Tα values for M_h ≥ 10^{10}M⊙ at z ≤ 7 than the \texttt{Clumpy} model in the \texttt{mhnc} scenario. This drop goes in hand with the increased central peak component in these more massive galaxies (c.f. Fig. 3 and the previous Section). The median Tα values of the Gaussian model always lie below those of the \texttt{Clumpy} and \texttt{Porous} models; a larger fraction of Lyα radiation escapes closer to the Lyα resonance and is thus subject to stronger attenuation by the IGM.

\textbf{Variance of the IGM transmission along different lines of sight:} To investigate how strongly the transmission of Lyα radiation through the IGM depends on the direction, we show the standard deviation of Tα values over the 6 lines of sight aligning with the major axes in relation to the corresponding mean value, σ_{Tα}/⟨Tα⟩ = √{(⟨Tα⟩^2−⟨Tα⟩^2)/⟨Tα⟩^2}, in the fourth row of Fig. 5. At all redshifts and for all models, σ_{Tα}/⟨Tα⟩ decreases with rising halo mass and decreasing redshift for the following reason. As galaxies grow in mass, they produce more ionising photons that can ionise larger regions around them and are also more likely to be located in more strongly clustered ionised regions, both enhancing and homogenising Lyα transmission through the IGM along different line of sights. However, we note that parts of the decrease of σ_{Tα}/⟨Tα⟩ with decreasing redshift is also due to ⟨Tα⟩ rising. Since it is hard to disentangle these two effects, we will focus on relative differences between the different reionisation scenarios and Lyα line profile models. Firstly, the more the emerging Lyα line profile is concentrated around the Lyα resonance, the more sensitive is Tα to the varying H I abundance around a galaxy, and the larger is the variance across different lines of sight (c.f. the higher σ_{Tα}/⟨Tα⟩ values in the Gaussian compared to the other two models, and in the Porous compared to the Clumpy model for M_h ≥ 10^{10.5}M⊙ when central peak component dominates). Secondly, we focus on Lyα line profiles more sensitive to the environmental H I abundance of a galaxy (Gaussian model). When accounting for the ⟨Tα⟩ values to be lower in the \texttt{mhnc} than in the \texttt{mhdec} scenario at z ≳ 7 (see median Tα values in the third row of Fig. 5), we can deduce that the variance of Tα across different lines of sight is higher in the \texttt{mhnc} than in the \texttt{mhdec} scenario. Indeed in the \texttt{mhnc} scenario, the shape of ionised regions is closer to spheres and less filamentary, which results in more “homogeneous” Tα values.

\textbf{Observed Lyα luminosity Lα}: For any model and reionisation scenario, the trend of Lα with rising halo mass depends on the respective trends of J{\textit{α}}_{\textit{intr}}, J{\textit{α}}_{\textit{esc}}, and Tα. Being surrounded by smaller ionised regions, the low Tα values of less massive galaxies (M_h ≤ 10^{10}M⊙) strongly suppress and shape their emerging Lyα radiation. In contrast, the Tα values of more massive galaxies (M_h ≥ 10^{10}M⊙) show only
weak trends with halo mass and similar values throughout reionisation. For this reason, the trends of their \( L_\alpha \) values with halo mass are predominantly shaped by the corresponding trends of \( L^{\text{Ly}\alpha}_{\text{esc}} \) and \( f^{\text{Ly}\alpha}_{\text{esc}} \). Though, for model parameters that reproduce the observed \( \text{Ly}\alpha \) LFs, a relative increase (decrease) of \( L^{\text{int}}_{\text{Ly}\alpha} \) towards higher halo masses, such as in the \( \text{mHnc} \) (\( \text{mHdec} \)) scenario, is compensated by an \( f^{\text{Ly}\alpha}_{\text{esc}} \) that decreases (more) strongly with halo mass. Nevertheless, the resulting relation between \( L_\alpha \) and halo mass does not significantly change. It shows that only more massive galaxies where SN and radiative feedback do not considerably suppress star formation exhibit observable \( \text{Ly}\alpha \) emission of \( L_\alpha \gtrsim 10^{41} \text{erg s}^{-1} \).

**Observed \( \text{Ly}\alpha \) equivalent width \( \text{EW}_{\alpha} \):** We compute the \( \text{Ly}\alpha \) equivalent width \( \text{EW}_{\alpha} \) from \( L_\alpha \) and the observed UV continuum luminosity at 1500Å (\( L_c \)). The trend of the median \( \text{EW}_{\alpha} \) with halo mass follows that of \( L_\alpha \), with median \( \text{EW}_{\alpha} \) values ranging from \( \sim 5 \text{--} 30 \)Å for galaxies in \( M_h \sim 10^{10} \text{M}_\odot \) halos to \( \sim 25 \text{--} 100 \)Å for galaxies in \( M_h \sim 10^{11.3} \text{M}_\odot \) halos. More massive galaxies with a strongly attenuated UV continuum – the fraction of these galaxies increases towards higher halo masses due to the higher abundance of dust – and high \( L_\alpha \) values show \( \text{EW}_{\alpha} \) values up to \( \sim 300 \)Å in the Clumpy and Porous models. However, these high \( \text{EW}_{\alpha} \) values are not present in the Gaussian model for the following reason: in this model, the escape of \( \text{Ly}\alpha \) and UV continuum radiation differs just by a constant factor, while the dust attenuation of \( \text{Ly}\alpha \) and UV continuum photons within a galaxy are not only linked via the dust mass in the Clumpy and Porous models.

In summary, we find that only more massive galaxies (\( M_h \gtrsim 10^{10} \text{M}_\odot \)) where star formation is not substantially suppressed by SN and radiative feedback from reionisation show significant \( \text{Ly}\alpha \) emission of \( L_\alpha \gtrsim 10^{41} \text{erg s}^{-1} \). This limitation of observable \( \text{Ly}\alpha \) emission to more massive galaxies allows the \( f^{\text{esc}} \)-dependency of the intrinsic \( \text{Ly}\alpha \) luminosity to be compensated by a weaker or stronger attenuation of \( \text{Ly}\alpha \) by dust within a galaxy. If less massive galaxies were visible in \( \text{Ly}\alpha \), they would break this degeneracy as they would not contain enough dust to attenuate the \( \text{Ly}\alpha \) radiation in all scenarios sufficiently.

### 5 The Spatial Distribution of \( \text{Ly}\alpha \) Emitting Galaxies

In this Section, we analyse where galaxies with observable \( \text{Ly}\alpha \) emission are located in the large-scale structure and how their environment and \( \text{Ly}\alpha \) luminosity distributions differ in our reionisation scenarios (\( \text{mHdec} \) and \( \text{mHnc} \)). For this purpose, we discuss the environment of \( \text{Ly}\alpha \) emitting galaxies in terms of their large-scale spatial distribution (Fig. 6), their surrounding over-density (\( 1+\delta \)) and \( \text{H}\ I \) fraction (\( \chi_{\text{HI}} \)) (Fig. 7), and their 3D autocorrelation functions (Fig. 8). As we yield very similar results for our three \( \text{Ly}\alpha \) lines profile models, we use the Porous model as a representative case.

#### 5.1 The environment

Before detailing the location of \( \text{Ly}\alpha \) emitting galaxies in the large-scale matter distribution, we briefly discuss the ionisation structure of the IGU using Fig. 6 and 7. Fig. 6 shows the ionisation fields at \( z = 8, 7 \) and 6.7 for the \( \text{mHdec} \) (top) and \( \text{mHnc} \) scenarios (bottom). As can be seen in this Figure, if \( f^{\text{esc}} \) decreases with halo mass (\( \text{mHdec} \) scenario), reionisation is not only more extended but also ionised regions are on average smaller, follow more the large-scale density distribution and thus have less bubble-like shapes than if \( f^{\text{esc}} \) increases with halo mass (\( \text{mHnc} \) scenario). The grey contours in Fig. 7, showing the two-dimensional probability density distribution of the \( \text{H}\ I \) fraction (\( \chi_{\text{HI}} \)) and over-density of the IGU at \( z = 8, 7, 3, 7 \) and 6.7 (derived from all cells of the 512\(^3 \) ionisation and density grids output by \text{ASTRAEUS} \), complement the picture. These contours indicate that not only an increasing fraction of the volume becomes ionised as reionisation progresses (from right to left) but also the \( \chi_{\text{HI}} \) values in ionised regions decrease (e.g. from \( \chi_{\text{HI}} \sim 10^{-4} \) (4--10\(^{-4} \)) in average dense regions with \( \log_{10}(1+\delta) + 1 \) at \( z = 8 \) to \( \chi_{\text{HI}} \sim 10^{-4.7} \) (10\(^{-4.3} \)) at \( z = 6.7 \) for the \( \text{mHdec} \) (\( \text{mHnc} \)) scenario). The latter is because as galaxies grow in mass with decreasing redshift, their emission of ionising photons increases, leading to a rise of the photoionisation rates within ionised regions and thus lower \( \chi_{\text{HI}} \) values. Moreover, at the same time, as the photoionisation rate within ionised regions becomes increasingly homogeneous, the enhanced number of recombinations in denser regions (for the detailed modelling description see \text{Hutter} 2018) leads to a positive correlation between the \( \text{H}\ I \) fraction and density in ionised regions. However, the exact value of the photoionisation rate within ionised regions and its spatial distribution depends strongly on the ionising emissivities escaping from the galaxies into the IGU. If less clustered low-mass galaxies drive reionisation – as in the \( \text{mHdec} \) (top row in Fig. 7) –, the resulting photoionisation rate is more homogeneous and lower than if the more strongly clustered massive galaxies are the main drivers of reionisation (c.f. \( \text{mHnc} \) scenario in the bottom row of Fig. 7). The difference in the photoionisation rate’s magnitude explains the shift of the \( \chi_{\text{HI}} \) values by an order of magnitude to lower values in under-dense to moderately over-dense regions (\( \log_{10}(1+\delta) \leq 1.2 \)) when going from the \( \text{mHdec} \) to the \( \text{mHnc} \) scenario. In contrast, the more homogeneous distribution of the photoionisation rate’s values enhances this drop in \( \chi_{\text{HI}} \) in over-dense regions where the most massive galaxies are located.

As we can see from the red stars in Fig. 6 and coloured contours in Fig. 7, galaxies emitting \( \text{Ly}\alpha \) luminosities of \( L_\alpha \geq 10^{42} \text{erg s}^{-1} \) always lie in ionised regions for our \( \text{Ly}\alpha \) line profile models. Although these galaxies trace the ionisation topology, their populations (and thus locations) hardly differ for our two opposing reionisation scenarios. This absence of a significant difference is due to their massive nature (see also e.g. \text{Kusakabe et al.} 2018): hence, all \( \text{Ly}\alpha \) emitting galaxies lie in over-dense regions, with the ones brighter in \( \text{Ly}\alpha \) located in denser regions (c.f. green to blue to red contours). The latter trend is mainly because more massive galaxies, which exhibit higher star formation rates and produce more ionising and \( \text{Ly}\alpha \) radiation, are located in denser regions.

#### 5.2 The clustering

In this Section, we address the question whether the \( \text{Ly}\alpha \) luminosity-dependent distribution of LAEs could differ for reionisation scenarios with opposing trends of \( f^{\text{esc}} \) with halo mass. For this purpose, we analyse the 3D autocorrelation function for LAE samples with different minimum \( \text{Ly}\alpha \) luminosities (Fig. 8). We define a galaxy to be an LAE if it has an observed \( \text{Ly}\alpha \) luminosity of \( L_\alpha \geq 10^{42} \text{erg s}^{-1} \).

Before we discuss the differences between our opposing \( f^{\text{esc}} \) descriptions, we give a brief overview of the global trends and their origins. Firstly, as predicted by hierarchical structure formation, all autocorrelation functions in Fig. 8 decrease from small to large scales, implying stronger clustering of galaxies on small scales than on large scales. Secondly, the dropping amplitude of the LAE autocorrelation functions with decreasing redshift (from ochre to blue lines) reflects
Figure 6. Neutral hydrogen fraction fields at $z = 8.0$ (left), $z = 7.0$ (centre), and $z = 6.6$ (right) for the mhdec (top) and mhinc Porous models (bottom). We show a 1.6$h^{-1}$cMpc-thick (5 cells) slice through the centre of the simulation box. The blue color scale depicts the volume-averaged value of the neutral fraction in each cell. Red stars show the location of LAEs, with their sizes and colour scale encoding the observed Ly$\alpha$ luminosity along the $z$-direction.

Figure 7. 2D probability distribution in $\chi_{HI}$ and overdensity for all simulation cells (grey) and galaxies with $L_{\alpha} \geq 10^{42}$ erg s$^{-1}$ (green), $L_{\alpha} \geq 10^{42.5}$ erg s$^{-1}$ (blue), and $L_{\alpha} \geq 10^{43}$ erg s$^{-1}$ (red) in the Porous model. The top (bottom) row shows results for the reionisation scenario where $f_{esc}$ decreases (increases) with halo mass $M_h$. 

MNRAS 000, 1–25 (2022)
the growth and increasing ionisation of ionised regions. Thirdly, since the $L_{\alpha}$ value of a galaxy is strongly correlated to its halo mass in our galaxy evolution model, selecting galaxies with increasingly brighter Ly$\alpha$ luminosities (left to right in Fig. 8) corresponds to selecting more massive galaxies. The latter explains the increasing amplitude and stronger clustering. Comparing the correlation functions of the $L_{\alpha}$ selected galaxies with those of LBGs (galaxies with $M_{UV} \geq -17$) shows that the Ly$\alpha$ selected galaxies are more massive than our LBGs (solid grey lines). It also shows that the decrease in the clustering of LAEs is partially due to galaxies of a given mass becoming a less biased tracer of the underlying density field as the density of the Universe drops with decreasing redshift.

Comparing the autocorrelation functions of our two opposing $f_{\text{esc}}$ descriptions, we find that the $\text{mhnc}$ scenario (dotted lines) has higher autocorrelation amplitudes than the $\text{mhdec}$ scenario (solid lines) throughout reionisation and for all minimum Ly$\alpha$ luminosities studied. This difference decreases towards larger scales. The reason for these higher amplitudes is twofold: On the one hand, the $\text{mhnc}$ scenario has a lower global average ionisation fraction at $z \geq 7$ than the $\text{mhdec}$ scenario (see Fig. 2). Its ionised regions are located around more biased tracers of matter, i.e. more massive galaxies, leading to a stronger clustering. While the scenarios’ difference in $\langle \chi_{HI} \rangle$ reaches its maximum with $\sim 0.13$ around $z \approx 8$, the difference in the autocorrelation amplitudes rises even towards higher redshifts. This is because, with increasing redshift, galaxies of the same mass become more biased tracers of the underlying matter distribution. Thus, the same difference in $\langle \chi_{HI} \rangle$ at higher $\langle \chi_{HI} \rangle$ values leads to a larger difference in the clustering of LAEs, since the Ly$\alpha$ luminosity of a galaxy correlates strongly with its halo mass. We note that selecting LAEs with a higher minimum Ly$\alpha$ luminosity also corresponds to selecting more biased tracers and yields higher correlation amplitudes (c.f. different panels in Fig. 8). On the other hand, during the early stages of reionisation, ionised regions grow preferentially around the most biased tracers of the underlying matter field (most massive galaxies) in the $\text{mhnc}$ scenario. Thus, we would expect that, at the same $\langle \chi_{HI} \rangle$ value, LAEs in this scenario are more clustered than LAEs in the $\text{mhdec}$ scenario where the $f_{\text{esc}}$ decreasing with rising halo mass counteracts the biased growth of ionised regions. Indeed, at $z \leq 7$, the correlation amplitude in the $\text{mhnc}$ scenario is higher or similar than in the $\text{mhdec}$ scenario, although the Universe is similarly or more ionised in the former, respectively. This difference becomes more apparent as we consider higher minimum Ly$\alpha$ luminosities of $L_{\alpha} > 10^{42.5}$ erg s$^{-1}$. It is driven by the higher photoionisation rates in the ionised regions around massive galaxies.

We conclude that, since LAEs coincide with the most massive galaxies located in dense and ionised regions, their clustering is primarily a tracer of the global ionisation state of the IGM. While the exact ionisation topology at fixed $\langle \chi_{HI} \rangle$ values has only a secondary effect on the clustering of LAEs during the second half of reionisation, the spatial distribution of LAEs provides a relatively robust tool to map the detailed ionisation history at early times.

6 THE RELATION OF LAES TO LBGS

In this Section, we address the question of what defines whether an LBG shows Ly$\alpha$ emission and why the fraction of LBGs with observable Ly$\alpha$ emission changes as the observed UV continuum luminosity (at 1500Å) or the minimum Ly$\alpha$ equivalent width, $EW_{\alpha}$, rise. For this purpose, we show both the fraction of LBGs with a Ly$\alpha$ equivalent width of at least $EW_{\alpha} \geq 25$Å (top row) and $EW_{\alpha} \geq 50$Å (central row) and the median $EW_{\alpha}$ value (bottom row) as a function of the UV continuum luminosity in Fig. 9.

For our three different Ly$\alpha$ line profile models, we find the median $EW_{\alpha}$ to exhibit similar values of $\sim 4 - 40$Å at all redshifts shown. Furthermore, the $EW_{\alpha}$ values range to lower values as galaxies become UV fainter. As galaxies become less massive, this spread in $EW_{\alpha}$ values reflects the increasingly broader range of star formation...
rate values to lower values, which traces back to the larger variety of mass assembly histories that increasingly include progenitors with particularly SN feedback-suppressed star formation. The $M_{UV}$-dependency of the fraction of LBGs with Ly$\alpha$ emission ($f_{\text{LAE}}$) also reflects this shift towards lower EW$_{\alpha}$ values (c.f. top and central row of Fig. 9): firstly, $f_{\text{LAE}}$ decreases towards lower UV luminosities, and secondly, this decrease is stronger for lower than higher EW$_{\alpha}$ cuts. These trends imply that UV bright galaxies are more likely to show higher EW$_{\alpha}$ values for all our Ly$\alpha$ line profile models and reionisation scenarios. For example, while only <20% of galaxies with $M_{UV} \approx -18$ exceed EW$_{\alpha} > 25$Å, >40% of galaxies with $M_{UV} \approx -20$ exceed EW$_{\alpha} > 25$Å and >5% even EW$_{\alpha} > 50$Å.

Moreover, both the slight rise of EW$_{\alpha}$ and $f_{\text{LAE}}$ values with decreasing redshift and their variation among our different reionisation scenarios can be attributed to the increasing fraction of Ly$\alpha$ radiation that is transmitted through the IGM as the Universe becomes more ionised (see $T_{\text{esc}}$ in Fig. 5). For example, at a given redshift $z > 7$ ($z < 7$), the EW$_{\alpha}$ and $f_{\text{LAE}}$ values are on average higher (lower) in the $mh_{\text{dec}}$ than in the $mh_{\text{inc}}$ scenario, which is due to a more (less) ionised IGM. Similarly, the lower EW$_{\alpha}$ values reached in the Gaussian model for UV fainter galaxies are due to the stronger absorption of Ly$\alpha$ radiation by H\textsc{i} in the IGM. Finally, we note that since in the Clumpy and Porous models the attenuation of the UV continuum and Ly$\alpha$ by dust do not necessarily correlate with each other (as e.g. parts of Ly$\alpha$ can escape via random walk), a few galaxies that are attenuated strongly in the UV but less in Ly$\alpha$ show high EW$_{\alpha}$ values of $\sim 1000$Å. Thus, the main driver of these high EW$_{\alpha}$ values is the dust attenuation of the UV continuum assumed in our models.

Comparing our fraction of LBGs showing Ly$\alpha$ emission with those obtained in observations (e.g. Schenker et al. 2012, 2014; Caruana et al. 2014; Pentericci et al. 2014, 2018; Mason et al. 2019), we find that (i) the observed trend of $f_{\text{LAE}}$ decreasing towards higher UV luminosity agrees roughly with our results for EW$_{\alpha} > 50$Å but not for EW$_{\alpha} > 25$Å, and (ii) our $f_{\text{LAE}}$ values are higher than those inferred from observations (again more so for EW$_{\alpha} > 25$Å than EW$_{\alpha} > 50$Å). These discrepancies hint either at our model predicting too high Ly$\alpha$ or too low UV luminosities (particularly for more massive galaxies) despite reproducing the observed Ly$\alpha$ and UV LFs, or observations missing bright LAEs. Interestingly, we find that the fraction of LBGs with high EW$_{\alpha}$ values...
of $f_{\text{LAE}}(\text{EW}_\alpha > 100\text{Å}) \approx 1-12\%$ and $f_{\text{LAE}}(\text{EW}_\alpha > 240\text{Å}) \lesssim 1\%$ in the Clumpy and Porous models are in rough agreement with the results from deep MUSE observations at $z = 3-6$ that consider only LAEs with detected UV continuum (Kerutt et al. 2022). A higher abundance of high EW$_\alpha$ values has been found in various high-redshift LAE observations (e.g. Shibuya et al. 2018; Malhotra & Rhoads 2002; Shimakaku et al. 2006). Nevertheless, our $f_{\text{LAE}}$ values agree roughly with the results from radiative hydrodynamical simulations post-processed with Ly$\alpha$ radiative transfer, such as sprinx (c.f. Fig. B1 in Garel et al. 2021).

7 CONCLUSIONS

We apply our new framework for LAEs to different reionisation scenarios, and analyse how the escape fraction of H I ionising photons, $f_{\text{esc}}$, and its dependence on halo mass affect the luminosity-dependent number and spatial distributions of LAEs. Besides $f_{\text{esc}}$ affecting the IGM ionisation topology and the strength of the Ly$\alpha$ line produced in the ISM, its sensitivity to the density and velocity structure of ISM gas and dust has been found to correlate with the Ly$\alpha$ line profile emerging from a galaxy and the fraction of Ly$\alpha$ radiation escaping into the IGM. Notably, the emerging Ly$\alpha$ line profile reflects the attenuation by dust in the ISM and can also change the fraction of Ly$\alpha$ radiation that traverses the IGM unattenuated by H I. For this reason, we build an analytical model for Ly$\alpha$ line profiles that emerge from a Ly$\alpha$ source surrounded by a shell of dusty gas clumps interspersed with low-density channels. Our model reproduces the numerical radiative transfer results of a shell with dust gas clumps of different sizes as presented in Gronke (2017). By coupling this model to astraus, a semi-numerical model coupling galaxy evolution and reionisation self-consistently, we derive the Ly$\alpha$ line profiles emerging from the simulated galaxy population and explore the resulting large-scale distribution of LAEs for different dependencies of $f_{\text{esc}}$ on halo mass (decreasing, constant, increasing) and Ly$\alpha$ line profiles (Gaussian profile, shell of dusty clumps interspersed with low-density channels or not). For this parameter space, we analyse the resultant ionisation topologies, the dependencies of Ly$\alpha$ line profiles and Ly$\alpha$ properties on halo mass, and the location of galaxies with observable Ly$\alpha$ emission in the large-scale structure. Our main results are the following:

(i) For a shell consisting of clumps of the same size, the Ly$\alpha$ line profile emerging from a galaxy develops from a central peak at the Ly$\alpha$ resonance dominated to a double peak dominated profile as it becomes more massive. Adding low-density channels results in either a weakening of this trend, particularly as $f_{\text{esc}}$ increases with rising halo mass.

(ii) In all reionisation scenarios and Ly$\alpha$ line profile models, LAEs (galaxies with $L_{\text{Ly}\alpha} \gtrsim 10^{42}\text{erg s}^{-1}$) are more massive galaxies with $M_h \gtrsim 10^{11}\text{M}_\odot$. These galaxies exhibit continuous star formation and are biased tracers of the underlying mass density distribution. Both allow efficient transmission of the Ly$\alpha$ line through the IGM by facilitating the build-up of ionised regions around them. In contrast, less massive galaxies are surrounded by smaller ionised regions, which results in their Ly$\alpha$ radiation being significantly attenuated by H I in the IGM.

(iii) As LAEs are more massive galaxies and the most biased tracers of the underlying mass density distribution, they are located in the densest and most highly ionised regions. This finding holds for any inside-out reionisation scenario where dense regions containing massive galaxies are ionised before under-dense voids and for Ly$\alpha$ line profiles exhibiting emission around/close to the Ly$\alpha$ resonance (see also Hutter et al. 2014, 2017). In such scenarios, the spatial distribution of LAEs is primarily sensitive to the global ionisation fraction and only in second-order to the ionisation topology or the trend of $f_{\text{esc}}$ with halo mass.

(iv) As the observable Ly$\alpha$ LFs are composed of the Ly$\alpha$ emission from more massive galaxies, a decrease in their intrinsic Ly$\alpha$ luminosities (Ly$\alpha$ produced in the ISM) due to higher $f_{\text{esc}}$ values can be compensated by reducing the attenuation by dust in the IGM (echoing the degeneracy found in Hutter et al. 2014). However, if $f_{\text{esc}}$ exceeds $\sim 0.5$ for the most massive galaxies ($M_h \gtrsim 10^{11}\text{M}_\odot$), their intrinsic Ly$\alpha$ luminosity is too low to reproduce the observed Ly$\alpha$ LFs (see also Hutter et al. 2014).

All combinations of our reionisation scenarios and Ly$\alpha$ line profile models result in Ly$\alpha$ and UV luminosities in reasonable agreement with observational constraints. However, although two of the three Ly$\alpha$ line profile models investigated use parameterisations of numerical Ly$\alpha$ radiative transfer simulation results, they represent idealised scenarios where the gas in each galaxy is distributed in clumps of the same mass. In reality, the density and velocity distributions of gas and dust in the ISM are more complex: Firstly, the dusty gas clumps will have different masses, with a distribution close to that of a scale-free one at the massive end. Such a mass distribution would result in more massive galaxies having larger clumps than less massive galaxies, which again would lead to a homogenisation of their Ly$\alpha$ line profiles where more massive (less massive) galaxies have an enhanced (weakened) central peak component and a weakened (enhanced) double-peak component. This change in the Ly$\alpha$ line profiles would result in the Ly$\alpha$ radiation being less (more) attenuated by dust in the ISM and traversing the IGM more (less) efficiently. Secondly, the medium between the clumps as well as the low-density channels might not be fully ionised (and very unlikely to be gas-free), causing the Ly$\alpha$ radiation escaping close to its resonance (central peak used in this work) to contribute to a narrower double-peak profile. Additionally, the gas may exhibit a turbulent velocity structure that could broaden the double-peak component. Both partially neutral low-density channels and an inhomogeneous velocity structure are likely to enhance the transmission of Ly$\alpha$ through the IGM. Thirdly, the attenuation of Ly$\alpha$ radiation by dust in the ISM depends on the distribution of dust in clumps. While our model assumes that gas and dust are perfectly mixed, a scenario where dust condensates in the centre surrounded by a shell of hydrogen gas would lower the absorption probability per clump and enhance the escape fraction of Ly$\alpha$ photons from a galaxy. Finally, simulations and observations of local analogues of high-redshift galaxies (i.e. regarding their extreme metallicity and ionisation continuum properties) indicate that stellar feedback, especially that of supernovae, heat the gas and drive gas outflows (e.g. Gronke & Oh 2020; Kakiichi & Gronke 2021; Carr et al. 2021; Fielding & Bryan 2022; Xu et al. 2023; Hu et al. 2023). Indeed expanding homogeneous shell models have been used to fit observed Ly$\alpha$ profiles from high-redshift analogues (e.g. Gronke 2017; Orlitová et al. 2018), however, the inferred outflow velocities are on average lower than those inferred from ultraviolet absorption lines of low-ionisation-state elements (Orlitová et al. 2018; Xu et al. 2023), hinting at more complex outflow geometries and kinematics of the neutral gas (see e.g. Carr et al. 2021; Blaizot et al. 2023). In general, outflowing neutral gas causes the Ly$\alpha$ photons to redshift, enabling easier escape from the galaxy and transmission through the IGM. While in principle outflows could enhance the observed Ly$\alpha$ emission, particularly from low-mass ($M_h \lesssim 10^{10}\text{M}_\odot$) galaxies, and make the large-scale LAE distribution more sensitive to the ionisation topology, their velocities or neutral gas fraction might be not...
DATA AVAILABILITY

The source code of the semi-numerical galaxy evolution and reionisation model within the astreaus framework is available on GitHub (https://github.com/annehutter/astreaus). The underlying N-body DM simulation, the astreaus simulations and derived data in this research will be shared on reasonable request to the corresponding author.

REFERENCES

Adams T. F., 1975, ApJ, 201, 350
Anderson L., Governato F., Karcher M., Quinn T., Wadsley J., 2017, MNRAS, 468, 4077
Asano R. S., Takeuchi T. T., Hirashita H., Inoue A. K., 2013, Earth, Planets and Space, 65, 213
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busca M. T., Klypin A. A., Primack J. R., 2013b, ApJ, 763, 18
Blaziot J., et al., 2023, MNRAS, 523, 3749
Bosman S. E. I., et al., 2022, MNRAS, 514, 55
Carr C., Scarlata C., Henry A., Panagia N., 2021, ApJ, 906, 104
Caruana J., Bunker A. J., Wilkins S. M., Stannan W. R., Lacy M., Jarvis M. J., Lorenzoni S., Hickey S., 2012, MNRAS, 427, 3055
Caruana J., Bunker A. J., Wilkins S. M., Stannan W. R., Lorenzoni S., Jarvis M. J., Ebert H., 2014, MNRAS, 443, 2831
Castellano M., et al., 2016, ApJ, 818, L3
Castellano M., et al., 2018, ApJ, 863, L3
Chardin J., Aubert D., Ocvirk P., 2012, A&A, 548, A9
Chen Z., Xu Y., Wang Y., Chen X., 2019, ApJ, 885, 23
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Dayal P., Maselli A., Ferrara A., 2011, MNRAS, 410, 830
Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, MNRAS, 445, 2545
Dayal P., et al., 2022, MNRAS, 512, 989
Dijkstra M., Gronke M., Venkatesan A., 2016, ApJ, 828, 71
Endsley R., Stark D. P., 2022, MNRAS, 511, 6042
Endsley R., et al., 2022, MNRAS, 517, 5642
Fan X., et al., 2006, AJ, 132, 117
Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, ApJ, 725, 633
Ferrara A., Loeb A., 2013, MNRAS, 431, 2826
Fielding D. B., Bryan G. L., 2022, ApJ, 924, 82
Forero-Romero J. E., Yepes G., Gottlöber S., Knollmann S. R., Cuesta A. J., Prada F., 2011, MNRAS, 415, 3666
Fuller S., et al., 2020, ApJ, 896, 156
Furlanetto S. R., Oh S. P., 2016, MNRAS, 457, 1813
Galliano F., 2022, Habilitation Thesis, p. 1
Garaldi E., Kamann R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, MNRAS, 512, 4009
Garel T., Blaizot J., Rosdahl J., Michel-Dansac L., Haehnelt M. G., Katz H., Kimm T., Verhamme A., 2021, MNRAS, 504, 1902
Gazagnes S., Chisholm J., Schreier D., Verhamme A., Izotov Y., 2020, A&A, 639, A85
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., Kauvar A. A., 2014, ApJ, 793, 30
Gronke M., 2017, A&A, 608, A139
Gronke M., Oh S. P., 2020, MNRAS, 494, L27
Hansen M., Oh S. P., 2006, MNRAS, 367, 979
He C.-C., Ricotti M., Geen S., 2020, MNRAS, 492, 4858
Heckman T. M., et al., 2011, ApJ, 730, 5
Howard C. S., Pudritz R. E., Harris W. E., Klessen R. S., 2018, MNRAS, 475, 3121
Hu E. M., Cowie L. L., Songaila A., Barger A. J., Rosenwasser B., Wold I. G. B., 2016, ApJ, 825, L7
Hu W., et al., 2021, Nature Astronomy, 5, 485
Hu W., et al., 2023, arXiv e-prints, p. arXiv:2307.04911
Hunter J. D., 2007, Computing In Science & Engineering, 9, 90

ACKNOWLEDGEMENTS

We thank Max Gronke and Peter Laursen for useful discussions and the anonymous referee for their comments. AH, GY, LL, PD and SG acknowledge support from the European Research Council’s starting grant ERC StG-717001 (“DELPHI”). AH, MT, PD also acknowledge support from the NWO grant 016.VID1.189.162 (“ODIN”) and the European Commission’s and University of Groningen’s CO-FUND Rosalind Franklin program. AH acknowledges support by the VILLUM FONDEN under grant 37459. PD thanks the Institute for Advanced Study (IAS) Princeton, where a part of this work was carried out, for their generous hospitality and support during the Bershadsky Fund. GY acknowledges financial support from MI- CIU/FEDER under project grant PGC2018-094975-C21. We thank Peter Behroozi for creating and providing the rockstar merger trees of the vsmdpl and esmdpl simulations. The authors wish to thank V. Springel for allowing us to use the L-Gadget2 code to run the different Multidark simulation boxes, including the vsmdpl and esmdpl used in this work. The vsmdpl and esmdpl simulations have been performed at LRZ Munich within the project pr87yi. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SUPERMUC-NG at Leibniz Supercomputing Centre (www.lrz.de). The CosmoSim database (www.cosmosim.org) provides access to the simulations and the Rockstar data. The database is a service by the Leibniz Institute for Astrophysics Potsdam (AIP). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant no. 140. This research made use of matplotlib, a Python library for publication quality graphics (Hunter 2007); and the Python library numpy (Oliphant 2006).
Appendix A: Geometrical Correction Factor $\xi$

When deriving the attenuation of the UV continuum by dust, we have assumed the light sources and dust to be homogeneously distributed within a slab. However, the numerical $Ly\alpha$ radiative transfer simulations in Gronke (2017) assume a screen of dust and gas between the light sources and the observer. To make the escape fractions of the radiation consistent, we introduce a geometrical correction factor $\xi$.
relations for these two geometries are given by
\[ f_{\text{screen}}(\tau_0) = \frac{1}{\cosh \left( \frac{\xi_0}{\tau_0 (a/r_0)^{1/3}} \right)} \] (A1)
\[ f_{\text{lab}}(\tau_0) = \frac{1 - \exp(-P)}{P} \] (A2)
\[ P = \epsilon_0 \left( \frac{a/r_0}{1/3} \right)^3 \] (A3)
\[ \tau_a = \tau_a,\text{total}(1 - A) \] (A4)
where \( A \) is the albedo. By equating the expressions for the \( \text{Ly}_\alpha \) escape fractions,
\[ f_{\text{screen}}(\epsilon_{0,\text{eff}}) = f_{\text{lab}}(\tau_0), \] (A5)
we derive a correction factor \( \xi \) that reduces the \( \text{H}_\text{i} \) column density in the screen geometry to the slab geometry.
\[ \xi = \frac{\epsilon_{0,\text{eff}}}{\tau_0} = \min \left( \xi_{\text{max}}, \frac{\epsilon_0}{3/2} \left( \frac{\arccosh \left( \frac{1}{1 - P} \right)}{P} \right)^{3/2} \right) \] (A6)
\[ P = \epsilon_0 a^{1/4} \left( \frac{1}{\tau_0} (M_{\text{d}}/M_{\text{HI}})^{\kappa_{\alpha\beta}/4} \right)^3 \tau_0 \] (A7)

Here we assume \( A = 0.5 \) and \( \xi_0 = 2.48 \), with \( \xi_0 \) being adjusted to reproduce \( f_{\text{screen}} \) shown in Fig. 1 in Forero-Romero et al. (2011). \( \xi_{\text{max}} = 0.35 \) represents an upper limit for a dust free homogeneous distribution of gas and sources. We derived its value as follows: Firstly we sum the Neufeld solutions (Eqn. 30) for an equidistant set of \( \tau_0 \) values between \( [0, \tau_0] \). Secondly, from the resulting \( \text{Ly}_\alpha \) line profile we estimate the effective \( \tau_0^{\text{eff}} \) value by measuring the peak positions \( x_{p,\text{eff}} \). Relating these peak positions to those obtained for the single Neufeld solution for \( \tau_0 \) \( (x_p) \), we obtain the ratio \( x_{p,\text{eff}}/x = \left( \tau_0^{\text{eff}}/\tau_0 \right)^{1/3} = \left( N_{\text{HI}}/N_{\text{HI}} \right)^{1/3} \), and the correction factor \( \xi_{\text{max}} = (x_{p,\text{eff}}/x_p)^{1/3} \). We have also checked that applying the correction factor \( \xi \) to \( N_{\text{HI}} \) reproduces the correct shift in the \( \text{Ly}_\alpha \) peak positions \( x_p \) shown in Fig. A5 in Forero-Romero et al. (2011).

**APPENDIX B: DELAYED NON-BURSTY SUPERNOVA FEEDBACK SCHEME**

We briefly describe our new formalism for the number of SN exploding if the stellar mass formed in a time step is assumed to form at a continuous rate across that time step. For a given star formation history \( \text{SFR}(t) \), the differential number of SN after a time \( t \) is given by
\[ \frac{dN_{SN}}{dt}(t) = \int_0^t \text{SFR}(t') \nu(t - t') dt'. \] (B1)
\( \nu(t) \) is the differential number of SN per stellar mass formed at \( t' = 0 \) and exploding at \( t' = t \), and hence yields as
\[ \nu(t) = M_{SN}^{-\gamma} \frac{dM_{SN}}{dt} \Theta(t - t_{\text{high}}) \Theta(t_{\text{SN,low}} - t). \] (B2)
with \( \gamma \) being the slope of the assumed IMF, \( t_{\text{high}} \) being the time after which the most massive stars sampled by the IMF, \( M_{\text{high}} \), explode as SN, and \( t_{\text{SN,low}} \) the time that it takes a star with the lowest stellar mass to explode as SN \( (M_{\text{SN,low}} = 8\, M_\odot) \). Starks of mass \( M_{SN} \) explode after a time \( t \) and the corresponding relation is described by
\[ M_{SN}/M_\odot = (t/\text{Myr} - 3/1.2 \times 10^5)^{-1/1.85} = a^{-c} (t - 3)^{-c}. \] (B3)

For constant star formation with
\[ \text{SFR}(t) = \begin{cases} 0 & t < t_i \\ s_0 & t_i \leq t \leq t_f \\ 0 & t_f < t \end{cases} \] (B4)
we yield after inserting Eqn. B2 and B3 into Eqn. B1
\[ \frac{dN_{SN}}{dt}(t) = \int_{t_i}^{t_f} \text{SFR}(t) dt \] (B5)
\[ = \int_{t_i}^{t_f} \frac{dM_{SN}}{dt}(t) dt \] (B6)
\[ = \int_{t_i}^{t_f} \frac{dM_{SN}}{dt}(t) \frac{dM_{SN}}{dt}(t - t') dt' \] (B7)
\[ = \int_{t_i}^{t_f} \frac{dM_{SN}}{dt}(t) \frac{dM_{SN}}{dt}(t - t') dt' \] (B8)
\[ \text{and} \quad t_{\text{max}} = \max(t_i, t_{\text{SN,low}}). \] (B9)

Eqn. B6 describes the differential number of SN exploding between the onset of star formation \( t_i \) and time \( t \) assuming constant star formation from \( t_i \) to \( t_f \). However, to obtain the total number of SN exploding in a given time step, i.e. between \( t_{\text{j-1}} \) and \( t_f \), we need to integrate over all contributions from \( t_{\text{j-1}} \leq t \leq t_f \) (i.e. integrating Eqn. B6 over time \( t \)),
\[ N_{SN}(t_{\text{j}}, t_f, t_{\text{j-1}}, t_f) = \int_{t_{\text{j-1}}}^{t_f} dN_{SN} dt \] (B10)
\[ = \int_{t_{\text{j-1}}}^{t_f} \frac{dM_{SN}}{dt} \frac{dM_{SN}}{dt} \left[ f_{\text{min}}(t) - f_{\text{max}}(t) \right]. \] (B11)

We solve the different summands in the integral separately, yielding
\[ F_{\text{max}} = \int_{t_{\text{j-1}}}^{t_f} dt \frac{dM_{SN}}{dt} \] (B12)
\[ = \int_{t_{\text{j-1}}}^{t_f} dt (t - t_{\text{max}} - 3)c(y - 1) \] (B13)
\[ \times \Theta(t_f + t_{\text{SN,low}} - t) \Theta(t - t_{\text{high}}) \] (B14)
\[ = \left[ \int_{t_{\text{j-1}} - t_{\text{high}} + t_{\text{SN,low}}}^{t_{\text{j-1}} - t_{\text{high}} + t_{\text{SN,low}}} \frac{dM_{SN}}{dt} (t_{\text{SN,low}} - t) \right] \frac{dM_{SN}}{dt} \] (B15)
\[ + \left[ \int_{t_{\text{j-1}} - t_{\text{high}} + t_{\text{SN,low}}}^{t_{\text{j-1}} - t_{\text{high}} + t_{\text{SN,low}}} \frac{dM_{SN}}{dt} (t_f - 3)c(y - 1) \right] \frac{dM_{SN}}{dt} \] (B16)
\[ \text{and} \quad t_{\text{max}} = \max(t_{\text{j-1}}, t_{\text{SN,low}}). \] (B17)
and

\[ F_{\min} = \int_{t_{j-1}}^{t_j} dt \, f_{\min}(t) \]  
\( \text{(B13)} \)

\[ = \int_{t_{j-1}}^{t_j} dt \, (t - t_{\min} - 3)^c(y-1) \times \Theta(t_f + t_{SN,low} - t) \Theta(t - t_i + t_{\star,high}) \]
\[ = \int_{\max(t_{j-1},t_i + t_{SN,low})}^{\min(t_j,t_i + t_{SN,low})} dt \, (t - t_{\min} - 3)^c(y-1) \]
\[ + \int_{\min(t_j,t_i + t_{SN,low})}^{\max(t_{j-1},t_i + t_{SN,low})} dt \, (t_{SN,low} - 3)^c(y-1) \]
\[ = \left[ \frac{(t_{\star,high} - 3)^c(y-1)}{c(y-1) + 1} \right] \frac{\min(t_f,t_i + t_{SN,low})}{\max(t_{j-1},t_i + t_{SN,low})} + \left[ \frac{(t - t_f - 3)^c(y-1) + 1}{c(y-1) + 1} \right] \frac{\min(t_f,t_i + t_{SN,low})}{\max(t_{j-1},t_i + t_{SN,low})}. \]

Inserting Eqn. B13 and B14 into Eqn. B12, we obtain

\[ N_{SN}(t_i, t_f, t_{j-1}, t_j) = s_0 \frac{a^c(y-1)}{1 - \gamma} \left[ F_{\min} - F_{\max} \right] \]  
\( \text{(B14)} \)

For a given star formation law \( \text{SFR}(t) \), the total stellar mass formed across all time is

\[ M_{\star}^{\text{tot}} = \int_0^{\infty} dt \, \text{SFR}(t) \left[ \frac{M_{\star,low}^{2-\gamma}}{M_{\star,high}^{2-\gamma}} \right] \]  
\( \text{(B15)} \)

\[ = \int_0^{\infty} dt \, \text{SFR}(t) \left[ \frac{M_{\star,low}^{2-\gamma} - M_{\star,high}^{2-\gamma}}{2 - \gamma} \right]. \]

Hence, for a constant star formation between \( t_i \) and \( t_f \), we finally obtain

\[ M_{\star}^{\text{tot}}(t_i, t_f) = \frac{s_0}{t_f - t_i} \frac{M_{\star,low}^{2-\gamma} - M_{\star,high}^{2-\gamma}}{2 - \gamma} \]  
\( \text{(B16)} \)

Finally, from Eqn. B14 and B16, we derive the number of SN exploding between times \( t_{j-1} \) and \( t_j \) from stars formed between \( t_i \) and \( t_f \) per stellar mass as

\[ \frac{N_{SN}(t_i, t_f, t_{j-1}, t_j)}{M_{\star}^{\text{tot}}(t_i, t_f)} = \frac{2 - \gamma}{1 - \gamma} \frac{a^c(y-1)}{M_{\star,low}^{2-\gamma} - M_{\star,high}^{2-\gamma}} \frac{F_{\min} - F_{\max}}{t_f - t_i} \]  
\( \text{(B17)} \)