



Article scientifique

Article

2024

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

## Modeling LAEs in the epoch of reionization with OBELISK : The connection between Lyman- $\alpha$ spectra and Lyman-continuum escape

---

Giovinazzo, Emma; Trebitsch, Maxime; Mauerhofer, Valentin; Dayal, Pratika; Oesch, Pascal

### How to cite

GIOVINAZZO, Emma et al. Modeling LAEs in the epoch of reionization with OBELISK : The connection between Lyman- $\alpha$  spectra and Lyman-continuum escape. In: Astronomy & astrophysics, 2024, vol. 688, p. A122. doi: 10.1051/0004-6361/202348765

This publication URL: <https://archive-ouverte.unige.ch/unige:187897>

Publication DOI: [10.1051/0004-6361/202348765](https://doi.org/10.1051/0004-6361/202348765)

© The author(s). This work is licensed under a Creative Commons Attribution (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0>

# Modeling LAEs in the epoch of reionization with OBELISK

## The connection between Lyman- $\alpha$ spectra and Lyman-continuum escape

Emma Giovinazzo<sup>1</sup>, Maxime Trebitsch<sup>2</sup>, Valentin Mauerhofer<sup>2</sup>, Pratika Dayal<sup>2</sup>, and Pascal A. Oesch<sup>1,3</sup>

<sup>1</sup> Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland  
e-mail: [emma.giovinazzo@unige.ch](mailto:emma.giovinazzo@unige.ch)

<sup>2</sup> Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

<sup>3</sup> Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N 2200, Denmark

Received 28 November 2023 / Accepted 29 April 2024

### ABSTRACT

**Context.** Lyman- $\alpha$  emitters (LAEs) are particularly useful objects in the study of the epoch of reionization. Lyman- $\alpha$  profiles can be used to estimate the number of ionizing photons that are able to escape galaxies, and therefore to understand which objects contributed to reionization. However, Lyman- $\alpha$  is a resonant line and its complex radiative transfer effects make the interpretation of the line challenging and require the use of appropriate radiative transfer methods for anything but the simplest gas distributions, such as uniform gaseous spheres, slabs, or cubes.

**Aims.** With this work, we aim to study the properties of simulated LAEs, and the robustness of these inferred properties during a change in the dust model. We also explore the Lyman continuum (LyC) escape fraction of these galaxies and compare our results with observationally calibrated methods to infer this quantity from the Lyman- $\alpha$  spectrum.

**Methods.** We used the radiative transfer code RASCAS to perform synthetic observations of 13 flux-selected galaxies from the OBELISK simulation at a redshift of  $z = 6$ , toward the end of the epoch of reionization. Each galaxy was observed in Lyman- $\alpha$ , as well as ionizing and nonionizing continuum from 48 different viewing angles.

**Results.** We show that the Lyman- $\alpha$  profiles emitted from a galaxy present large variations with a change in viewing angle and that the relation between peak separation and the Lyman- $\alpha$  escape fraction is not as strong as previously found, as we find lines of sight with both a low peak separation and a low escape fraction, due to their dust content. We also show that the properties of the Lyman- $\alpha$  line are reasonably robust during a change in the dust model. Lastly, we compare the LyC escape fractions that we derive from the simulation to three observationally calibrated methods of inferring this quantity. We determine that none of these relations reproduce the scatter that we find in our sample, and that high escape fraction lines of sight have both a low peak separation and a low dust extinction in the ultraviolet (UV).

**Key words.** line: profiles – radiative transfer – galaxies: high-redshift – dark ages, reionization, first stars

## 1. Introduction

The epoch of reionization (EoR) marks the last phase transition of the Universe, during which the initially cold and neutral intergalactic medium (IGM) became ionized. This period also marks the formation of the first luminous objects, which formed at the center of collapsed halos and produced sufficient amounts of ionizing radiation to drive the reionization process (e.g., Barkana & Loeb 2001; Dayal & Ferrara 2018).

Current constraints from quasar spectra place the end of the EoR around  $z \sim 5.5$ – $6$  (e.g., Wyithe et al. 2005; Fan et al. 2006; Eilers et al. 2018; Schroeder et al. 2012; Becker et al. 2021; Bosman et al. 2022), but it is still unclear which sources are responsible for the emission of the majority of the ionizing photons. It is commonly believed that the main driver of the reionization of hydrogen in the Universe is massive star formation in galaxies (e.g., Robertson et al. 2015; Bouwens et al. 2015; Finkelstein et al. 2019; Naidu et al. 2020; Yung et al. 2020a,b; Trebitsch et al. 2022), as young and massive stars emit copious amounts of Lyman continuum (LyC) radiation ( $\lambda < 912 \text{ \AA}$ ). This scenario implicates a very patchy process (e.g., D’Aloisio et al. 2015; Davies & Furlanetto 2016; Eilers et al. 2018), whereby galaxies are able to ionize gas in their surroundings, forming

growing bubbles of ionized gas around themselves until these bubbles overlap and the entirety of the IGM is ionized. Another potentially important source of ionizing photons are active galactic nuclei (AGNs; e.g., Madau & Haardt 2015; Giallongo et al. 2015). However, studies of the AGN luminosity function in the EoR have shown that AGNs are only minor contributors of reionization and fail to solely reionize the universe by  $z \sim 6$  (e.g., Onoue et al. 2017; Parsa et al. 2017; Yung et al. 2021; Matsuoka et al. 2023), although according to theoretical models, AGNs could dominate the ionizing photon production at the highest masses (Dayal et al. 2020).

To properly understand this process, it is necessary to determine which galaxies leak the most ionizing photons, which depends on large-scale properties of the IGM such as clumping, and on properties of the galaxies such as their star formation rates (SFRs) and the escape fraction of LyC photons  $f_{\text{esc}}(\text{LyC})$ . This last parameter quantifies the fraction of ionizing photons that are able to escape the galaxy, and therefore ionize the surrounding medium. Determining this quantity through observations is challenging, as it is only possible to observe the escaped radiation. Only recently have surveys been able to accumulate large numbers of directly detected escaping ionizing photons (e.g., Bergvall et al. 2006; Shapley et al. 2006; Heckman et al. 2011;

Leitet et al. 2013; Izotov et al. 2016a,b, 2018a,b, 2021; Fletcher et al. 2019; Davis et al. 2021; Flury et al. 2022a,b; Saxena et al. 2022b). Moreover, it is impossible to detect LyC photons at high redshifts,  $z \gtrsim 4$ , as they are completely absorbed by the residual HI in the IGM. It is therefore necessary to use indirect indicators that link  $f_{\text{esc}}(\text{LyC})$  to observable properties, which are often calibrated on low-redshift observations. Some previously studied indicators are based on metal lines and line ratios, such as the Mg II line (e.g., Chisholm et al. 2020; Xu et al. 2022), the C IV line (e.g., Saxena et al. 2022a; Schaerer et al. 2022), the O III  $\lambda 5007/\text{O III } \lambda 3727$  line ratio ( $\text{O}_{32}$ ; e.g., Jaskot & Oey 2013; Nakajima & Ouchi 2014; Izotov et al. 2018b; Paalvast et al. 2018; Tang et al. 2021), or also the ultraviolet (UV)  $\beta$  slope (e.g., Chisholm et al. 2022; Flury et al. 2022b).

Another notable indicator is the Lyman- $\alpha$  line (e.g., Verhamme et al. 2015, 2017; Izotov et al. 2021). Both Lyman- $\alpha$  and LyC are expected to escape galaxies from the same low neutral hydrogen column density paths (e.g., Dijkstra 2014; Verhamme et al. 2015; Gazagnes et al. 2020; Kakiichi & Gronke 2021; Begley et al. 2024), making this line a very interesting indicator. The most promising characteristic of this line for the estimation of  $f_{\text{esc}}(\text{LyC})$  is the peak separation ( $v_{\text{sep}}$ ), first suggested by Henry et al. (2015) and Yang et al. (2017) and then confirmed by Izotov et al. (2018b) and the Low-Redshift Lyman Continuum Survey (LzLCS; Flury et al. 2022b).

Galaxies that are particularly bright in the Lyman- $\alpha$  line, due to their star-forming nature (Dijkstra 2014), are called Lyman- $\alpha$  emitters (LAEs). The LAEs have been observed and studied at low redshift (Orlitová et al. 2018), although they are quite rare in the nearby universe (Hayes 2015). They are particularly interesting when used to study the EoR, both as probes of ionized bubbles in the partially neutral IGM and as contributors of reionization (Hu et al. 2016; Matthee et al. 2018; Songaila et al. 2018; Meyer et al. 2020). The resonant nature of Lyman- $\alpha$  does, however, add a layer of complexity to the study of this line in the EoR, as Lyman- $\alpha$  photons are unlikely to escape the neutral intergalactic gas. A relevant review of LAEs has been produced by Ouchi et al. (2020).

In this work, we analyze synthetic observations of a sample of LAEs extracted from the OBELISK cosmological simulation (Trebitsch et al. 2021) in order to study their Lyman- $\alpha$  properties in connection with the escape fraction of ionizing photons. In particular, we study how the modeling of the dust distribution inside the galaxies affects the Lyman- $\alpha$  transfer, and how this, in turn, impacts the usability of Lyman- $\alpha$  diagnostics in the study of the sources of reionization. In this work, we define  $f_{\text{esc}}(\text{LyC})$  as the absolute escape fraction along the observed line of sight.

This work is structured as follows. In Sect. 2, we present the OBELISK simulation and the radiative transfer code used in the Lyman- $\alpha$  and LyC post-processing. Then, in Sect. 2.6, we study the impact of the dust modeling on our results. In Sect. 3, we show the properties of the synthetic Lyman- $\alpha$  profiles. In Sect. 4, we compare the LyC escape fraction,  $f_{\text{esc}}(\text{LyC})$ , directly measured in the simulation to standard estimates used throughout the literature.

## 2. Methods

### 2.1. Obelisk simulation

In this paper, we analyze a sample of galaxies extracted from the OBELISK cosmological simulation (Trebitsch et al. 2021), which

re-simulates at a much higher resolution and down to  $z \simeq 3.5$  the evolution of a large, overdense region ( $V \simeq 23 h^{-3} \text{cMpc}^3$ ) within the HORIZON-AGN (Dubois et al. 2014) ( $100 h^{-1} \text{Mpc}^3$ ) cosmological volume. In this section, we highlight the features of the simulation most relevant to the study of LAEs and refer the reader to Trebitsch et al. (2021) for further details.

The simulation uses the same WMAP-7 (Komatsu et al. 2011)  $\Lambda$ CDM cosmology as HORIZON-AGN, with the cosmological parameters: the Hubble constant,  $H_0 = 70.4 \text{ km s}^{-1} \text{Mpc}^{-1}$ , the dark energy density parameter,  $\Omega_\Lambda = 0.728$ , the matter density parameter,  $\Omega_m = 0.272$ , the baryon density parameter,  $\Omega_b = 0.0455$ , amplitude normalization of the power spectrum,  $\sigma_8 = 0.81$ , and the spectral index,  $n_s = 0.967$ . The high-resolution region was defined in the initial conditions as the convex hull enclosing all the particles that will end up within a sphere of radius  $\approx 2.5 h^{-1} \text{cMpc}$  centered on the most massive halo in the HORIZON-AGN simulation at  $z = 2$ . In this region, dark matter (DM) particles have a mass resolution of  $1.2 \times 10^6 M_\odot$ , and gas cells are allowed to be refined up to 35 pc if the total mass in the cell exceeds eight times the mass resolution.

OBELISK was run with the RAMSES-RT (Rosdahl et al. 2013; Rosdahl & Teyssier 2015) radiation-hydrodynamics module of the adaptive mesh refinement (AMR) code RAMSES (Teyssier 2002), which follows the evolution of DM, gas, stars, black holes, and the (ionizing) radiation field. Gas in the simulation is described as a monoatomic gas with an adiabatic index of  $\gamma = 5/3$  and was evolved using an unsplit second-order MUSCL-Hancock scheme (van Leer 1979) with the HLLC Riemann solver (Toro et al. 1994). The masses of DM and star particles were projected onto the grid using a cloud-in-cell interpolation (with the DM particles projected on a coarser grid with  $\Delta x \simeq 540 \text{ pc}$ ) and combined with the gas density to evolve the gravitational field. The radiation field is split into three frequency intervals corresponding to HI, He I, and He II-ionizing photons. The ionizing radiative transfer used in OBELISK then follows the standard RAMSES-RT methodology, using a first-order Godunov method with the M1 closure (Levermore 1984; Dubroca & Feugeas 1999) to evolve the radiation field. As is described in Trebitsch et al. (2021), we used the variable speed of light approximation of Katz et al. (2017), which uses the reduced speed of light approximation in regions resolved better than  $\Delta x \simeq 2 \text{ kpc}$  (this is the case for all the intra-halo gas in this study). While OBELISK separately tracks the contribution of massive stars and AGNs to the radiation field, we treat the two components as one in this study. Finally, the gas thermodynamical evolution was followed using the nonequilibrium hydrogen and helium thermochemistry method from RAMSES-RT, described in Rosdahl et al. (2013), which includes a cooling contribution of metals down to a floor temperature of 50 K. Importantly for this work, we followed on the fly the ionization state of hydrogen and helium, which is necessary to accurately predict the Lyman- $\alpha$  emissivity from the gas.

In the simulation, star particles are formed with a mass of  $m_\star \simeq 10^4 M_\odot$  in the dense and turbulent ISM (above a number density of  $n_{\text{SF}} = 5 \text{ cm}^{-3}$  and a turbulent Mach number of  $M \geq 2$ ), assuming a local star formation efficiency per free-fall time,  $\epsilon_\star$ , defined using the thermo-turbulent model presented, for example, in Trebitsch et al. (2017) and Kimm et al. (2017). After they are formed, star particles emit ionizing radiation following the BPASS v2.2.1 binary stellar population model (Eldridge et al. 2017; Stanway & Eldridge 2018), as a function of the age and metallicity of the particle. After a delay of  $t_{\text{SN}} = 5 \text{ Myr}$ , a mass fraction,  $\eta_{\text{SN}} = 0.2$ , of the initial stellar population in each

**Table 1.** Characteristics of the 13 galaxies in the sample.

ID	$M_*$ [ $M_\odot$ ]	SFR [ $M_\odot \text{ yr}^{-1}$ ]	$M_{\text{UV}}^{\text{obs}}$	$M_{\text{UV}}^{\text{intr}}$	$f_{\text{esc}}(\text{LyC})$	$M_{\text{dust}}$ [ $M_\odot$ ]
2899	$1.4 \times 10^{10}$	31.0	-19.4	-21.5	0.001	$1.1 \times 10^7$
24688	$6.4 \times 10^9$	35.7	-19.4	-23.4	0.004	$8.7 \times 10^6$
55468	$4.9 \times 10^9$	17.9	-19.0	-21.9	0.002	$9.8 \times 10^6$
36112	$3.7 \times 10^9$	22.8	-19.4	-22.3	0.007	$1.4 \times 10^7$
41163	$3.2 \times 10^9$	8.4	-19.0	-21.9	0.007	$5.5 \times 10^6$
3639	$2.6 \times 10^9$	18.1	-19.1	-21.8	0.001	$5.9 \times 10^6$
42655	$1.7 \times 10^9$	10.0	-19.2	-21.5	0.023	$5.8 \times 10^6$
8776	$1.6 \times 10^9$	6.2	-19.0	-20.9	0.026	$8.3 \times 10^6$
39021	$1.5 \times 10^9$	9.5	-19.0	-21.1	0.033	$8.0 \times 10^6$
76168	$1.5 \times 10^9$	8.0	-19.2	-21.6	0.011	$3.1 \times 10^6$
10345	$2.5 \times 10^8$	1.5	-19.1	-20.0	0.235	$1.5 \times 10^6$
57172	$2.4 \times 10^8$	1.5	-19.3	-19.8	0.485	$1.2 \times 10^6$
67244	$7.0 \times 10^7$	0.7	-19.0	-19.4	0.195	$2.1 \times 10^5$

**Notes.** Here,  $M_*$  is their stellar mass, SFR is the star formation rate, calculated over 10 Myr,  $M_{\text{UV}}^{\text{obs}}$  is the observed absolute dust-attenuated UV magnitude,  $M_{\text{UV}}^{\text{intr}}$  is the intrinsic absolute UV magnitude,  $f_{\text{esc}}(\text{LyC})$  is the LyC escape fraction, and  $M_{\text{dust}}$  is the dust mass.

particle explodes as supernovae, injecting mass, energy, momentum, and metals into their environments. Supernova feedback was modeled following the mechanical feedback implementation of Kimm & Cen (2014), Kimm et al. (2015), with an additional boost due to the preprocessing of the ISM in (unresolved) H II regions inspired by the results of Geen et al. (2015).

The simulation also includes a detailed model for black hole formation, growth through accretion and mergers, a dual-mode AGN feedback model, and follows ionizing radiation resulting from accretion. In this work, we focus on galaxies less massive than  $\lesssim 10^{10} M_\odot$ , where the black hole growth is inefficient (see for example Habouzit et al. 2017). As a result, we ignore the AGN effect on the galaxies: for more details on the black hole model, we refer the reader to Trebitsch et al. (2021).

Importantly for this work, OBELISK models the evolution of dust separately from the metals, such that the dust-to-metal ratio is allowed to vary locally. Dust is released in the ISM in supernovae, and the dust content in a cell grows via the accretion of metals from the gas phase. Conversely, dust is destroyed in supernova shocks and through thermal sputtering in hot gas. While the dust is not coupled with the radiative transfer or with the cooling in the OBELISK simulation, it plays an important role in setting the Lyman- $\alpha$  properties of galaxies. We describe how we use this spatially varying dust content to estimate the attenuation of Lyman- $\alpha$  and continuum photons in Sect. 2.5.1.

Finally, galaxies (and their DM haloes) were identified using the ADAPTAHOP algorithm (Aubert et al. 2004), operating on all collisionless particles at the same time (DM and stars). Only galaxies with more than 100 star particles and 100 DM particles are considered in this work: as such, the lowest possible stellar mass is  $\approx 10^6 M_\odot$ . For all galaxies, the (dust-attenuated) UV magnitude was computed by casting 192 rays isotropically from each star particle, integrating the dust column density along these lines of sight, and computing the direction-averaged total UV luminosity. We used these UV magnitudes to define our galaxy sample, as is described in the next section.

## 2.2. Galaxy sample

We selected galaxies in order to create a synthetic flux-selected sample targeting faint Lyman-break galaxies corresponding to

typical LAEs at the end of the EoR. For this, we focused on the  $z = 6$  snapshot of the OBELISK simulation and selected all (central) galaxies with a 3D-averaged, dust-attenuated UV magnitude in the range of  $-19.5 \leq M_{\text{UV}} \leq -19$ . Because different galaxies have different levels of attenuation, this corresponds to a much broader range of intrinsic UV luminosities, with magnitudes between  $-19.4$  and  $-23.4$ . We summarize the properties of our sample in Table 1.

## 2.3. Ly $\alpha$ emission and transfer

The radiative transfer of Lyman- $\alpha$  photons through our galaxy sample was performed in post-processing with the 3D Monte Carlo radiative transfer code RASCAS (Michel-Dansac et al. 2020). We used the standard ‘‘peeling-off’’ technique (e.g., Yusef-Zadeh et al. 1984; Zheng & Miralda-Escude 2002; Whitney 2011) to produce synthetic observations of the Lyman- $\alpha$  line on the fly. At each scattering, both the probability that the photon would have gone in the direction of the observer and the probability that the photon would have escaped are calculated. This method allows for the production of synthetic spectra and images in any direction and is performed at the same time as the radiative scattering.

The first step in the Lyman- $\alpha$  transfer is the emission of Monte Carlo photons, each representing a fraction of the total Lyman- $\alpha$  flux. This was done by computing the Lyman- $\alpha$  emissivity directly from the cells in the simulation, accounting both for recombination and collisional excitation mechanisms. The emitting gas was then sampled using  $10^5$  photons, following the luminosity distribution of the cells. In a given cell, the number of photons emitted due to recombination per unit time was computed as

$$\dot{N}_{\gamma, \text{rec}} = n_e n_p \epsilon_{\text{Ly}\alpha}^B(T) \alpha_B(T) \times (\Delta x)^3, \quad (1)$$

where  $n_e$  and  $n_p$  are the electron and proton number densities, and  $(\Delta x)^3$  is the cell volume. The  $B$  stands for case B recombination, with  $\alpha_B$  being the recombination coefficient given by (Hui & Gnedin 1997):

$$\alpha_B = 2.753 \times 10^{-14} \text{ cm}^3 \text{ s}^{-1} \frac{\lambda_{\text{HI}}^{1.5}}{(1 + (\lambda_{\text{HI}}/2.740)^{0.407})^{2.242}}, \quad (2)$$

and  $\epsilon_{\text{Ly}\alpha}^B(T)$  being the fraction of recombination producing Lyman- $\alpha$  photons. We evaluated it using the fit from Cantalupo et al. (2008), which is based on the tabulated values from Pengelly (1964) for  $T > 10^3$  K and Martin (1988) for  $T < 10^3$  K:

$$\epsilon_{\text{Ly}\alpha}^B(T) = 0.686 - 0.106 \log(T_4) - 0.009 \times (T_4)^{-0.44}. \quad (3)$$

Instead, the number of Lyman- $\alpha$  photons emitted from the gas per unit time in each cell due to collisions is given by

$$\dot{N}_{\gamma,\text{col}} = n_e n_{\text{HI}} C_{\text{Ly}\alpha}(T) \times (\Delta x)^3, \quad (4)$$

where  $n_{\text{HI}}$  is the number density of neutral hydrogen atoms and  $C_{\text{Ly}\alpha}$  is the rate of collisional excitation from level  $1s$  to  $2p$ , given by (Goerdt et al. 2010):

$$C_{\text{Ly}\alpha}(T) = \frac{2.41 \times 10^{-6}}{T^{0.5}} \left( \frac{T}{10^4} \right)^{0.22} \times \exp\left(-\frac{h\nu_{\text{Ly}\alpha}}{kT}\right) \text{cm}^3 \text{s}^{-1}. \quad (5)$$

The Monte Carlo photons are then placed in a random position within their emission cell and emitted in a direction randomly drawn from an isotropic distribution. The initial wavelength is taken from a Gaussian distribution centered around the Lyman- $\alpha$  frequency in the rest-frame of emitting gas, with a width following the thermal velocity dispersion of the gas.

Once the photons have been emitted, they travel until their next scattering event, which can be either on HI or on a dust grain. In practice, the optical depth to the next scattering event is drawn from  $\tau_{\text{event}} = -\ln(r)$ , where  $r$  is a random number between zero and one. The photons are then propagated until they reach this optical depth. In each cell, the total optical depth (gas + dust) is given by

$$\tau_{\text{tot}} = r(n_{\text{HI}}\sigma_{\text{HI}} + n_d\sigma_d) = \tau_{\text{HI}} + \tau_d. \quad (6)$$

The neutral hydrogen cross section is defined as

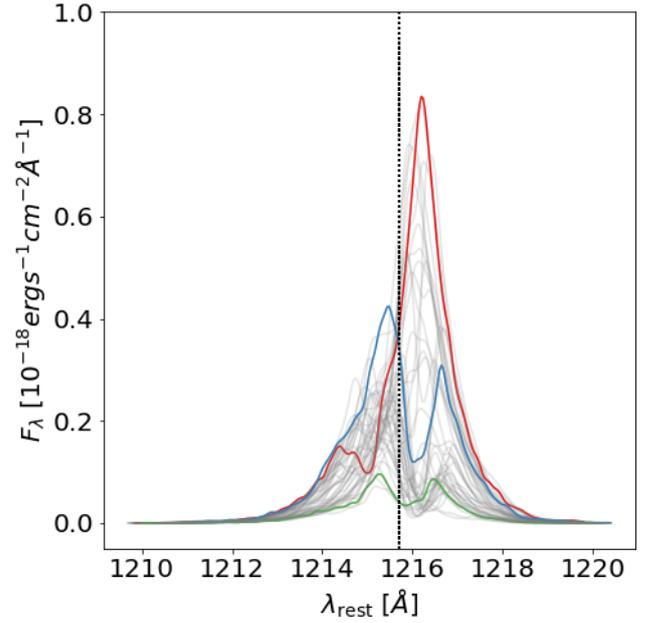
$$\sigma_{\text{HI}} = f_{12} \frac{\pi e^2}{m_e c \Delta \nu_D} \phi(x), \quad (7)$$

where  $f_{12}$  is the Lyman- $\alpha$  oscillator strength, and  $\phi(x)$  is the (Voigt) line profile. We discuss in Sect. 2.5 the details of the evaluation of the dust optical depth,  $\tau_d$ . We propagated photons until they were either destroyed by a dust grain or escaping the galaxy, defined here as crossing the virial radius,  $R_{\text{vir}}$ , of the halo.

#### 2.4. Mock observation

Each galaxy was observed along 48 different lines of sight, uniformly distributed around a sphere, selected using the Hierarchical Equal Area isoLatitude Pixelization (HEALPIX) decomposition (Gorski et al. 2005; Calabretta & Roukema 2007), which divides a sphere into equal area pixels. While none of these directions are strictly independent, previous works (e.g., Blaizot et al. 2023; but see also Verhamme et al. 2012; Behrens et al. 2014; Smith et al. 2018) have shown that the Lyman- $\alpha$  line emerging from galaxies is highly anisotropic and can mimic a large range of observed spectra. We used this characteristic to our advantage by treating each line of sight as a separate point.

For each sight line, RASCAS outputs a data cube with two spatial directions and one spectral direction, from which we can extract the spectra by collapsing the cube in the spatial directions. The extracted Lyman- $\alpha$  spectra are observed in a rest-frame wavelength range of 1210 Å–1220 Å in 250 pixels. This corresponds to a wavelength resolution of 0.04 Å or  $\sim 10 \text{ km s}^{-1}$ .



**Fig. 1.** All 48 spectra of galaxy ID 8776, where the three colored spectra are meant to highlight the diversity of spectral shapes that are possible from the same galaxy. The dotted vertical black line is the Lyman- $\alpha$  wavelength.

These profiles can then be shifted to the observer frame and analyzed. To identify the peaks of the profiles, we used the `findpeaks` python module (Taskesen 2020). We first reduced the Monte Carlo noise by applying a Gaussian smoothing with a sigma of two pixels to the profile. Smoothing on this scale should ensure that we do not remove any significant peaks. We then applied the `findpeaks` module on this new profile, to identify all the local maxima in the distribution. Next, we computed the noise level as the standard deviation of our unsmoothed data when subtracted to a polynomial fit to the profiles, which were smoothed over four pixels to perform the fit. We used the noise to discern between significant peaks and peaks resulting from noise, by only selecting from all the peaks only those with a signal-to-noise ratio above three. The peak separation was then computed as the separation between the identified peaks. If only one peak could be identified, the peak separation was set as 0, and in the case of triple peaks we determined the peak separation to not be identifiable without a visual inspection so we removed the spectrum from the peak separation dataset. Overall, we find 277 spectra where only one peak could be identified, 44% of the total sight lines, and only nine spectra where more than two peaks were identified, which is only 1.4% of the total dataset. For all other uses, the spectra still need to be smoothed to reduce the Monte Carlo noise, but that was done directly with the RASCAS function, which also loads the mock observations and shifts them into the observer frame. This function takes as the Gaussian sigma a wavelength value, which we chose as 0.5 Å. This level of smoothing is also necessary to produce spectra that can be comparable to observations, like those seen in Fig. 1.

Finally, we performed three additional RASCAS runs to measure the ionizing and nonionizing UV luminosities and estimate the continuum around Lyman- $\alpha$  for each galaxy along the same lines of sight. For the nonionizing UV, we propagated photons in the wavelength range 1480 Å–1520 Å and only considered

dust absorption using the same approach as for Lyman- $\alpha$  (see Sect. 2.5 for details). The ionizing photon run considers both gas and dust absorption and was only used to investigate the variation in  $f_{\text{esc}}(\text{LyC})$  with direction. To estimate the continuum at the Lyman- $\alpha$  wavelength, we sampled the emission from stars with  $10^5$  photons, using the tabulated spectral shape, between the wavelengths of 1150 Å–1200 Å. In all three cases, we obtained data cubes from which we extracted the spectra by collapsing the cube and smoothing all the spectra with the same Gaussian, and we then calculated the luminosity and flux, just as was done with Lyman- $\alpha$ .

## 2.5. Dust models

We now focus on the calculations required to assess the dust optical depth,  $\tau_d$ , in the Lyman- $\alpha$  transfer. Since OBELISK includes a model for dust evolution, we can take two different approaches to estimate  $\tau_d$  in this work. First, we directly make use of the OBELISK dust model. Then, we also compute  $\tau_d$  following the formulation of Laursen et al. (2009), which is a relatively standard approach and is the default model used in RASCAS.

### 2.5.1. Dust from OBELISK

The first model of dust attenuation that we considered relies on the dust distribution tracked by the OBELISK simulation: the dust mass density,  $\rho_d$ , and the dust-to-metal ratio,  $\mathcal{D}/Z = \rho_d/\rho_Z$ , is therefore allowed to vary within a galaxy.

However, the knowledge of  $\rho_d$  is insufficient to estimate the dust optical depth: indeed, the dust cross section depends on the grain size, which is not tracked directly in OBELISK<sup>1</sup>. Instead, we computed the dust optical depth over a path of length  $L$  directly as  $\tau_d = \alpha^{\text{ext}}L = \rho_d\kappa_dL$ , where  $\kappa_d$  is the absorption coefficient and

$$\alpha^{\text{ext}}(\lambda) = \int_{a_{\text{min}}}^{a_{\text{max}}} \pi a^2 \sum_{i=s,c} n(a) Q_{\text{ext}}^i(a, \lambda) da. \quad (8)$$

Here,  $Q_{\text{ext}}$  is the extinction efficiency factor as a function of grain size and wavelength taken from Weingartner & Draine (2001) and Laor & Draine (1993) for two grain types (silicates,  $s$ , and carbonaceous grains,  $c$ ) and  $n(a)$  is the number density of grains of size  $a$ . We assume that silicates have a mass fraction of 54% and 46% for the carbonaceous grains (Hirashita & Yan 2009). We then converted the mass density into a number density by assuming that the grain size is well described by the MRN distribution (Mathis et al. 1977),  $dn/da = Ca^{-3.5}$ , where  $C$  is a proportionality constant. We note that  $n_d$  is the total number density, such that  $n(a) = n_dCa^{-3.5}$ . For a dust grain specific density,  $\mu_g$ , we can write the mass density as

$$\rho_d = \int_{a_{\text{min}}}^{a_{\text{max}}} \frac{4}{3} \pi a^3 \mu_g C n_d a^{-3.5} da = \frac{8}{3} \pi \mu_g n_d C (\sqrt{a_{\text{max}}} - \sqrt{a_{\text{min}}})$$

for grain sizes between  $a_{\text{min}}$  and  $a_{\text{max}}$ , and so

$$n(a) = \frac{3\rho_d}{8\pi\mu_g (\sqrt{a_{\text{max}}} - \sqrt{a_{\text{min}}})} a^{-3.5}.$$

<sup>1</sup> While the dust evolution model implicitly assumes that all grain have an average size of  $a = 0.1 \mu\text{m}$ , this is not appropriate to derive an extinction curve.

When evaluating the dust optical depth, RASCAS directly uses the form  $\tau_d = \rho_d\kappa_dL$ . We therefore rewrote Eq. (8) as

$$\kappa_d = \frac{3}{8\mu_g (\sqrt{a_{\text{max}}} - \sqrt{a_{\text{min}}})} \int_{a_{\text{min}}}^{a_{\text{max}}} a^{-1.5} \sum_{i=s,c} Q_{\text{ext}}^i(a, \lambda) da. \quad (9)$$

We integrated Eq. (9) numerically to obtain  $\kappa_d(\lambda)$ , and obtained a fit (used for numerical efficiency) using the python package LMFIT (Newville et al. 2015), with a relative difference to the tabulated values lower than 5% around the Lyman- $\alpha$  wavelength.

### 2.5.2. Dust following the metals

The second model that we considered is the default dust model used in RASCAS that assumes a dust-to-metal ratio proportional to the ionization fraction of the cell. It has been used previously in, for example, Blaizot et al. (2023) to post-process simulations where the dust is not tracked self-consistently. Specifically, we used the Large Magellanic Cloud (LMC) model of Laursen et al. (2009), based on previous works by Pei (1992) and Gnedin et al. (2008).

The dust optical depth along a path of length  $L$  in a cell was computed as  $\tau_d = n_d\sigma_dL$ , where  $n_d$  is a dust ‘‘pseudo-density’’ and  $\sigma_d$  is the dust cross section per hydrogen atom: the dust (pseudo-)density,  $n_d$ , therefore gives the effective number of hydrogen atoms required to reach a given optical depth. The dust density is given by

$$n_d = (n_{\text{H I}} + f_{\text{ion}}n_{\text{H II}}) \frac{Z}{Z_0}, \quad (10)$$

where  $f_{\text{ion}} = 0.01$  is the fraction of dust grain surviving in the ionized gas,  $Z$  is the local metallicity of the gas, and  $Z_0$  is a reference metallicity, here taken to be that of the LMC. Our galaxies are therefore assumed to have an LMC-like dust-to-metal ratio, roughly proportional to the neutral hydrogen fraction of each cell. The cross section,  $\sigma_d$ , per hydrogen atom is given, close to the Lyman- $\alpha$  frequency, by (Laursen et al. 2009):

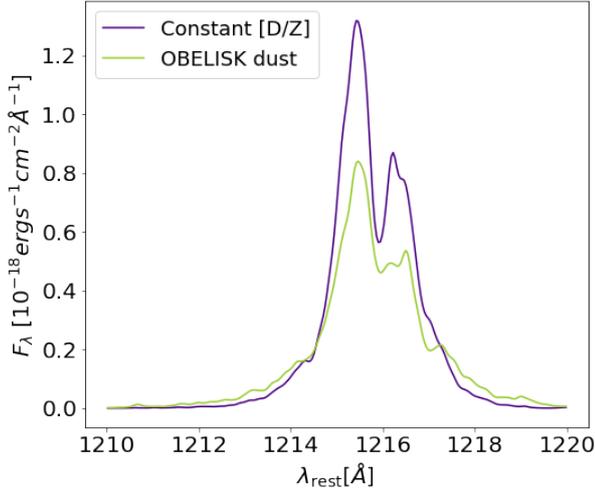
$$\sigma_d/10^{-21} \text{ cm}^2 = 0.723 + 4.46 \times 10^{-5} (T/10^4 \text{ K})^{1/2} x. \quad (11)$$

Here,  $x$  is the parametrization of the frequency,  $\nu$ , defined as  $x \equiv (\nu - \nu_0)/\Delta\nu_D$ , where  $\nu_0$  is the line center and  $\nu_D$  is the Doppler width of the line.

## 2.6. Dust model comparison

In this section, we compare the two different dust models considered in this work to assess how they affect the predicted observable properties of our simulated galaxies.

In Fig. 2, we present a sample spectrum of a single line of sight of galaxy ID 2899 for both models. The green line shows the fiducial model from OBELISK and the purple line shows the Laursen et al. (2009) model in which the dust density is inferred from the local metallicity, based on LMC values, denoted as the constant  $\mathcal{D}/Z$ . We see that the Lyman- $\alpha$  flux observed from the run with the OBELISK dust distribution is lower than in the case of the other dust model. This might be due to a difference in the spatial distribution of dust around the galaxy, which in the case of our fiducial dust model might be more concentrated around Lyman- $\alpha$  emitting regions. However, the main features of the line are mostly unaffected by the choice of dust modeling. In particular, the position of each peak and the relative prominence of the blue peak are consistent between the two runs.



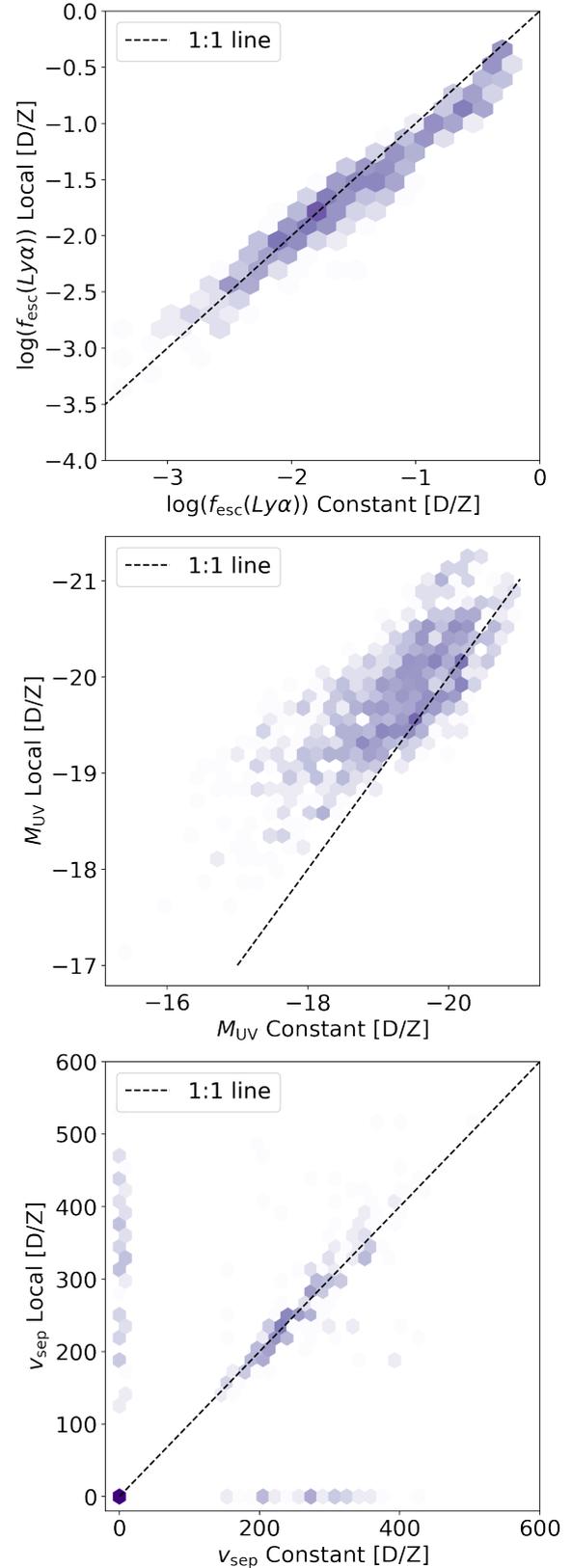
**Fig. 2.** Comparison of the Lyman- $\alpha$  profile of the same line of sight of galaxy ID 2899 for the two different dust models. The purple line is the observed spectrum for the constant  $\mathcal{D}/Z$  model and the green line is the observed spectrum for the dust in OBELISK.

We now take a more quantitative view of the difference between the two sets of runs in Fig. 3. We compare the distribution of the Lyman- $\alpha$  escape fraction  $f_{\text{esc}}(\text{Ly}\alpha)$  (top), which, in simulations, depends on the absorption of photons by dust in the ISM, the absolute observed UV magnitude  $M_{\text{UV}}$  (middle), and the Lyman- $\alpha$  peak separation,  $v_{\text{sep}}$  (bottom). The values measured for our fiducial model, the dust measured from OBELISK, are presented on the  $y$  axis denoted by a local  $\mathcal{D}/Z$ , while the other model is denoted by a constant  $\mathcal{D}/Z$ . In each panel, the color indicates the density of points (i.e., the number of sight lines, with darker colors meaning more points), and the dashed black line shows the 1:1 line.

The top panel shows that the Lyman- $\alpha$  escape fraction,  $f_{\text{esc}}(\text{Ly}\alpha)$ , is consistent between the two dust models, although there is some small deviation from the 1:1 line at  $f_{\text{esc}}(\text{Ly}\alpha) \gtrsim 10\%$ , where the constant  $\mathcal{D}/Z$  dust model is higher. This is qualitatively consistent with the results of Fig. 2.

In the second panel of Fig. 3, we see that the observed UV magnitudes are generally brighter in the run with the OBELISK dust distribution. Once again we can explain this by studying the difference in dust distribution between the two models. Indeed, assuming a constant  $\mathcal{D}/Z$  over the whole ISM will lead to overestimating it in lower-density regions, where the  $\mathcal{D}/Z$  in the simulation tends to be lower. Since UV photons escape more easily from these low-density regions, this will lead to higher attenuation, and so the model with constant  $\mathcal{D}/Z$  will attenuate the UV more strongly, which explains the faint observed UV magnitudes. In Appendix B we show two maps of  $\mathcal{D}/Z$  for our fiducial dust model for two of our galaxies to show that taking a constant  $\mathcal{D}/Z$  would overestimate the dust content, especially in the low-density region.

Finally, in the last panel, we find that the measurement of the peak separation is mostly unaffected by the dust model used. This is consistent with the expectations from simpler expanding shell models: for example, Verhamme et al. (2015) found that the position of the (red) peak is mostly independent of the dust optical depth of the shell. However, for many lines of sight, we see that spectra presenting a double peak for a dust model present a single peak for the other. In most cases, this is due to one of the peaks becoming weaker, and therefore not significant enough to be identified by our peak finding routine, as is seen for example



**Fig. 3.** Comparison of  $\log(f_{\text{esc}}(\text{Ly}\alpha))$ , UV magnitude, and Lyman- $\alpha$  peak separation (top, middle, and bottom panels, respectively) inferred from the synthetic spectra run with the two different dust models. In general, we see that the values are consistent with each other, as they do not substantially deviate from the one-to-one line, which is shown with the dashed black line in each panel.

in Fig. 2, where the red peak is much less marked when using the OBELISK dust model.

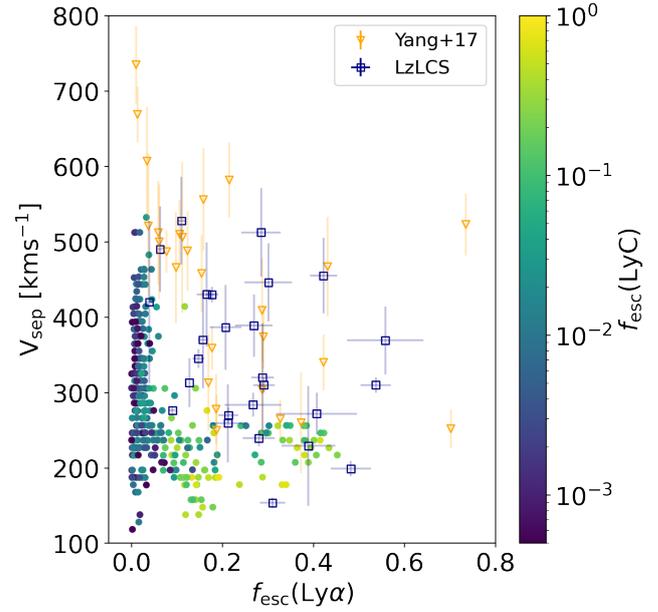
Overall, the two dust models seem fairly consistent with each other. The  $f_{\text{esc}}(\text{Ly}\alpha)$  is consistent, with small differences at high escape fractions, and the UV magnitude is also fairly consistent, especially for brighter galaxies, while faint galaxies are about 1.5 dex from the 1:1 line. Lastly, the peak separation is consistent, when both peaks are detected in both profiles. This shows that although the choice of dust model can change the luminosity (in Lyman- $\alpha$  or in UV) of the galaxy, the Lyman- $\alpha$  spectral shape is mostly unaffected. In the next section, we explore the usability of the shape of the Lyman- $\alpha$  line as a diagnostic tool to understand the escape of ionizing radiation in our simulated LAEs. Motivated by the results of Fig. 3, we choose to focus on a single dust model, namely the one from OBELISK, corresponding to a local  $\mathcal{D}/Z$ .

### 3. Lyman- $\alpha$ emission from simulated galaxies

In this section, we discuss the integrated Lyman- $\alpha$  properties of our simulated sample and consider the diversity of Lyman- $\alpha$  profiles resulting from the synthetic observations of our sample for the OBELISK dust model, our fiducial model.

We first present in Fig. 1 the emerging Lyman- $\alpha$  profile of galaxy ID 8776 along 48 different lines of sight, each shown as a thin gray line. We highlight three of the lines of sight displaying representative but extremely different line shapes: in red a prototypical red peak, in green a (fairly attenuated) double peak, and in blue a line dominated by its blue peak. In general, we expect red-dominated profiles to trace outflows, and blue-dominated profiles to trace inflows (Verhamme et al. 2006). This diversity is reminiscent of the results of, for example, Blaizot et al. (2023), and corresponds to different orientations of both the galaxy and gas flows around it. In Appendix A we show the same figure for the other 12 galaxies. In addition to the different shapes of the line, the total Lyman- $\alpha$  flux also varies with direction: the double peak highlighted in green in Fig. 1 appears to be substantially fainter than the spectrum with a red peak highlighted in red. This implies that the flux-selected sample of LAEs could miss some of the galaxies that happen to be observed from a direction where the flux is low, even if the galaxy emits a strong Lyman- $\alpha$  line in other directions. If we define LAEs as objects with a rest frame equivalent width of  $EW_0 \geq 20 \text{ \AA}$  (Ouchi et al. 2020), the galaxy in Fig. 1 would result in an LAE in only 22 out of 48 lines of sight, so 46% of the time. In the whole sample, this fraction varies from 31% to 98%, with a mean of 71%.

We note that Fig. 1 also shows that in many directions the spectra present a somewhat unexpectedly strong blue peak, which is the case for all of the galaxies in our sample. There are several explanations for this: first of all, we do not model the attenuation from the IGM in this work, which would suppress our strong blue peaks. However, even at lower redshift, where IGM attenuation is less strong, observations tend to show a stronger red peak (e.g., Verhamme et al. 2017; Leclercq et al. 2017; Izotov et al. 2018b; Orlitová et al. 2018). A second possibility is that our strong blue peaks come from our choice of limiting the Lyman- $\alpha$  transfer only up to the virial radius of each galaxy: Blaizot et al. (2023) have shown that transfer further than the virial radius tends to reduce the strength of the blue peaks. By excluding the gas between 1 and  $3R_{\text{vir}}$ , we are (artificially) making this effect weaker. Indeed, in relatively low-mass systems like the LAEs we are modeling here, galactic outflows typically extend beyond the virial radius (Mitchell et al. 2017,

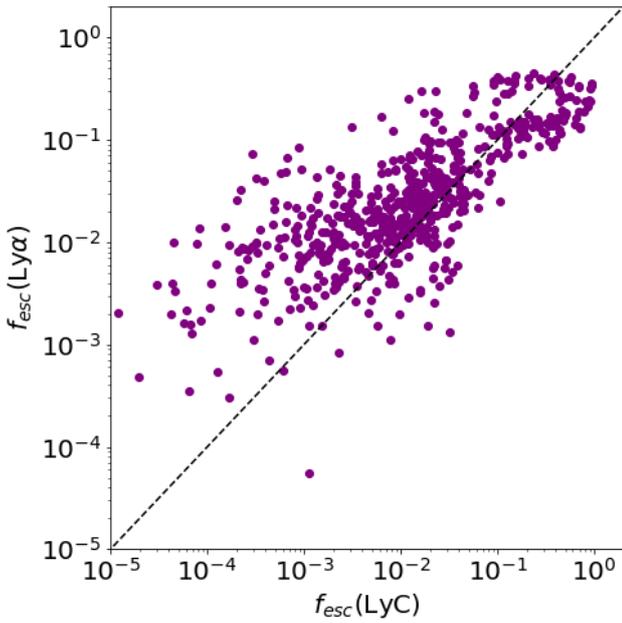


**Fig. 4.** Lyman- $\alpha$  peak separation against Lyman- $\alpha$  escape fraction, color-coded by the LyC escape fraction. We also show observational data from Yang et al. (2017) in the orange triangles and Flury et al. (2022b) in the blue squares. Here, we are showing that although spectra with large peak separations (above  $300 \text{ km s}^{-1}$ ) always have low LyC and Lyman- $\alpha$  escape fractions, the contrary is not true, as we have many lines of sight with both low peak separation and escape fractions.

2020), and so by restricting the Lyman- $\alpha$  transfer to the halo we might miss part of the outflow.

From each synthetic observation, we measure the peak separation, as detailed in Sect. 2.4, and using the intrinsic Lyman- $\alpha$  luminosity of each galaxy (derived directly from the simulation), we measure the Lyman- $\alpha$  escape fraction for each galaxy along each line of sight. In Fig. 4, we show the separation between the two peaks in the Lyman- $\alpha$  profiles against the Lyman- $\alpha$  escape fraction and we compare it to observational data from Yang et al. (2017) and Flury et al. (2022b). We choose here to color-code our points by the LyC escape fraction along the same line of sight,  $f_{\text{esc}}(\text{LyC})$ , where lighter points correspond to a higher  $f_{\text{esc}}(\text{LyC})$ . Here, we note that the (apparent) discreteness of the Lyman- $\alpha$  peak separation values is completely artificial and comes from our choice of spectral resolution for the synthetic observations. All of the spectra were observed with the same resolution, in the same wavelength range, and were also smoothed in the same way, which leads to the discreteness in the values of the peak separation.

Overall, the figure displays a trend of peak separation decreasing with  $f_{\text{esc}}(\text{Ly}\alpha)$ , in good agreement with expectations from both earlier models (e.g., Verhamme et al. 2015) and observations of low- $z$  LAEs (Yang et al. 2017; Flury et al. 2022b): the sight lines with very high  $f_{\text{esc}}(\text{Ly}\alpha)$  tend to have low peak separations, and sight lines with peak separations above  $300 \text{ km s}^{-1}$  tend to have low  $f_{\text{esc}}(\text{Ly}\alpha)$ . However, we find that the low separation, low  $f_{\text{esc}}(\text{Ly}\alpha)$  corner of the figure is also populated: galaxies with  $f_{\text{esc}}(\text{Ly}\alpha) \lesssim 10\%$  have a wide range of possible peak separations, in contrast both with observations and with shell models such as those of Verhamme et al. (2015), which predict a low HI column density in lines of sight with low peak separation, and in contrast with observational data. We believe that



**Fig. 5.**  $f_{\text{esc}}(\text{LyC})$  versus  $f_{\text{esc}}(\text{Ly}\alpha)$ . The black dashed line corresponds to the 1:1 relation.

this might be due to the clumpy nature of the galaxies modeled in this work. The Lyman- $\alpha$  radiation can originate from different clumps in the galaxy, which might have different properties and different intrinsic velocities, which could contribute to the creation of multiple close peaks. For a detailed discussion on the nature of the difference between idealized models such as the picket fence model and results from radiative transfer in hydrodynamical simulations, we refer the reader to [Mauerhofer et al. \(2021\)](#).

A similar behavior is also seen for  $f_{\text{esc}}(\text{LyC})$ . High  $f_{\text{esc}}(\text{LyC})$  sight lines have low peak separation but low  $f_{\text{esc}}(\text{LyC})$  sight lines correspond to a large range of peak separations. This implies a relation between  $f_{\text{esc}}(\text{LyC})$  and  $f_{\text{esc}}(\text{Ly}\alpha)$ , which is confirmed by Fig. 5, which shows that an increase in  $f_{\text{esc}}(\text{Ly}\alpha)$  corresponds to an increase in  $f_{\text{esc}}(\text{LyC})$ , although with some scatter, especially at low  $f_{\text{esc}}(\text{LyC})$ . This relation is also expected from theory (e.g., [Dijkstra et al. 2016](#); [Kimm et al. 2022](#); [Maji et al. 2022](#)) and observations (e.g., [Flury et al. 2022b](#); [Begley et al. 2024](#)). A possible explanation for this scatter between  $f_{\text{esc}}(\text{LyC})$  and the  $f_{\text{esc}}(\text{Ly}\alpha)$ , and therefore also the Lyman- $\alpha$  peak separation, could be the much more localized emission of the LyC radiation with respect to the Lyman- $\alpha$  radiation, such as that observed in the Sunburst Arc by [Kim et al. \(2023\)](#). Indeed, they show that non-LyC-leaking regions of the Sunburst Arc still have substantial  $f_{\text{esc}}(\text{Ly}\alpha)$  of about 13%, and it is therefore not unexpected to have very low  $f_{\text{esc}}(\text{LyC})$  while still being able to observe Lyman- $\alpha$ .

#### 4. Estimate of $f_{\text{esc}}(\text{LyC})$

Several quantitative measurements based on the Lyman- $\alpha$  line shape have been proposed in the literature to infer the escape of LyC radiation. In this section, we will describe and analyze some of these methods: the Lyman- $\alpha$  peak separation (e.g., [Izotov et al. 2018b](#)), the Lyman- $\alpha$  central flux fraction proposed by [Naidu et al. \(2022\)](#), and the fraction of flux in the Lyman- $\alpha$  trough compared to the continuum flux proposed by

[Gazagnes et al. \(2020\)](#). Having explored the Lyman- $\alpha$  line shape in the previous section, we will now assess how these proposed diagnostics perform for our simulated galaxy sample, and compare that to the “true”  $f_{\text{esc}}(\text{LyC})$  measured directly in the simulation.

##### 4.1. Peak separation, $v_{\text{sep}}$

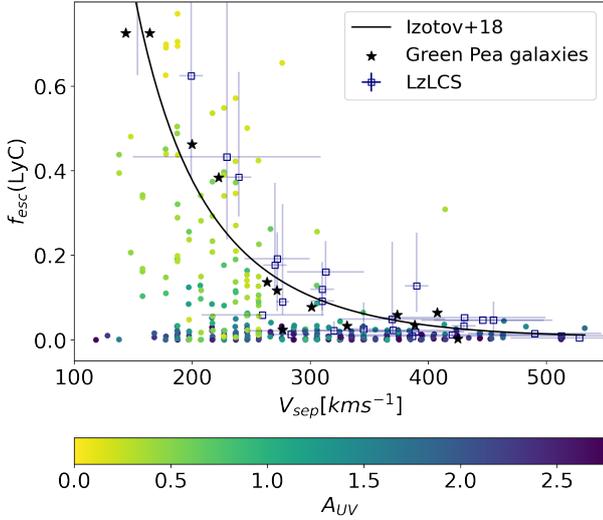
From theoretical studies of spherical shell models, we expect a relation between the peak separation and the column density neutral gas in the line of sight (e.g., [Verhamme et al. 2006, 2015](#); [Orsi et al. 2012](#)), which is itself related to  $f_{\text{esc}}(\text{LyC})$ , where a higher column density corresponds to a lower escape fraction ([Verhamme et al. 2017](#)). We therefore expect a relation between the Lyman- $\alpha$  peak separation and  $f_{\text{esc}}(\text{LyC})$ . This relation was first shown in simulations by [Verhamme et al. \(2017\)](#) and then calibrated by [Izotov et al. \(2018b\)](#), using a set of low-redshift LyC-leaking galaxies known as “Green Peas” ([Cardamone et al. 2009](#)). This particular set of galaxies was used as they are confirmed LyC leakers, at a redshift high enough that the Milky Way’s neutral gas would not absorb the LyC radiation, but low enough to avoid LyC absorption by the neutral IGM and to be detected by the Cosmic Origin Spectrograph on the *Hubble* Space Telescope (HST). The Green Pea galaxies were selected for their high  $\text{O}_{32}$ . Their SFRs are comparable to the most massive galaxies in our sample but are brighter than our sample, with UV magnitudes around  $-19.5$  to  $-20.5$ . The stellar masses of the Green Peas are also comparable with those of the galaxies in our sample. The relation found by [Izotov et al. \(2018b\)](#) is

$$f_{\text{esc}}(\text{LyC}) = \frac{3.23 \times 10^4}{V_{\text{sep}}^2} - \frac{1.05 \times 10^2}{V_{\text{sep}}} + 0.095. \quad (12)$$

This is empirically derived and breaks down at  $v_{\text{sep}} < 140 \text{ km s}^{-1}$ , where it gives unphysical values of  $f_{\text{esc}}(\text{LyC}) > 1$ . It can nonetheless be used to infer the escape fraction of ionizing radiation, or a lower limit in the case of low peak separations, and has already been applied at high redshift (e.g., [Meyer et al. 2020](#)). The anticorrelation between peak separation and the LyC escape fraction found by [Izotov et al. \(2018b\)](#) was also found in other observations of low-redshift LyC leakers ([Verhamme et al. 2017](#); [Flury et al. 2022b](#); [Izotov et al. 2022](#)) and several radiation hydrodynamic simulations of cloud-scale LyC escape ([Kimm et al. 2019](#); [Kakiichi & Gronke 2021](#)).

In Fig. 6, we show where our simulated data lie in the  $f_{\text{esc}}(\text{LyC})$  to  $v_{\text{sep}}$  plane, color-coded by the dust attenuation in the UV for each line of sight, and compare it with the best-fit relation from [Izotov et al. \(2018b\)](#). We also show the Green Pea galaxies this relation was calibrated on with the black stars and the LzLCS results in the empty blue squares.

Overall, there is a reasonable agreement between the simulation and the observations, although with a large scatter: OBELISK reproduces the trend that a larger peak separation indicates a lower  $f_{\text{esc}}(\text{LyC})$ . Nevertheless, we also note a large number of sight lines with low peak separation and low  $f_{\text{esc}}(\text{LyC})$ , suggesting that the observed relation is not fully representative of our simulated sample. This is in agreement with the results from the SPHINX<sup>20</sup> simulation ([Choustikov et al. 2024](#)). We find 97 sight lines with  $0 < v_{\text{sep}} < 250 \text{ km s}^{-1}$  and  $f_{\text{esc}}(\text{LyC}) < 0.1$ , or 15% of the total sight lines, which is a range not yet probed by observations, although new HST observations were recently designed to probe this parameter space ([Leclercq et al. 2022](#), GO-17153, PI Leclercq). We find that these results are consistent with the work of [Kimm et al. \(2019\)](#), who also find a population of sight lines



**Fig. 6.** Comparison of our data, color-coded by the dust attenuation in the UV of each line of sight, with the relation found by [Izotov et al. \(2018b\)](#), shown with the black line. The Green Pea galaxies that the relation was calibrated on are shown with black stars. The more recent measurements of peak separations of low-redshift galaxies by [Flury et al. \(2022b\)](#) are shown with empty blue squares.

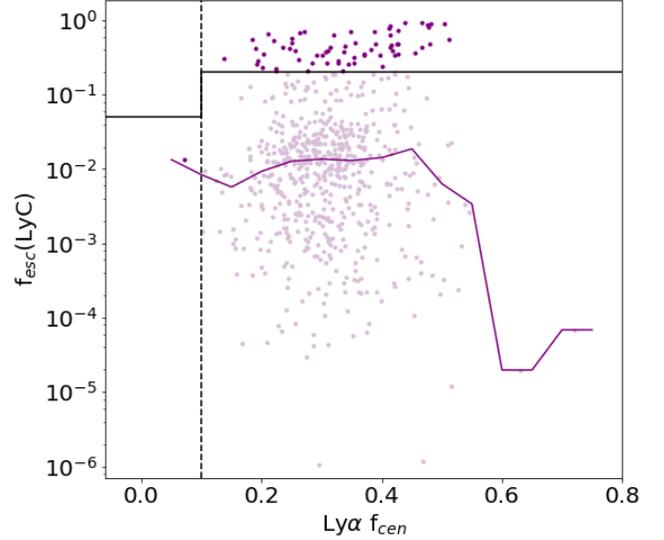
around a simulated molecular cloud that have low peak separation but low  $f_{\text{esc}}(\text{LyC})$ .

Interestingly, points with low dust attenuation (in yellow) tend to align better with the [Izotov et al. \(2018b\)](#) relation, while sight lines with high dust attenuation tend to depart from it. Using data from LzLCS, [Saldana-Lopez et al. \(2022\)](#) show that it is necessary to account for dust attenuation in the indirect diagnostics of LyC escape: our numerical results seem to confirm this. This, however, does not mean that dust absorbs LyC radiation: indeed, in general we do not see more than a 1% difference in the  $f_{\text{esc}}(\text{LyC})$  between runs with or without dust. Instead, our interpretation is that as the amount of dust in a given line of sight scales with the amount of hydrogen, sight lines with higher dust attenuation are also those with higher HI column density. Therefore, lines of sight with higher dust attenuation should correspond to a lower LyC escape fraction. Here, we would like to note that we find lines of sight with high dust attenuation but low peak separation, which may seem unexpected, but these are the same lines of sight as that in the bottom left corner of Fig. 4, which has a low peak separation and low  $f_{\text{esc}}(\text{Ly}\alpha)$ . We therefore believe that the explanation that was given in Sect. 3 about the clumpy nature of our galaxies applies here as well.

Another possible explanation for this discrepancy between the simulated data and the Green Peas is that Green Pea galaxies might be intrinsically different from the galaxies in our sample. Green Pea galaxies are slightly brighter, with  $-19.5 < M_{\text{UV}} < -20.5$ , and were selected for their strong oxygen line ratio,  $\text{O}_{32}$ . Similarly, the LzLCS galaxies are mostly brighter than ours, although their  $v_{\text{sep}}$  sample has magnitudes ranging between  $-18.5 < M_{\text{UV}} < -21$ . Both observational samples also comprise galaxies usually found in low-density environments at low redshift, while OBELISK simulates a high-density environment at high redshift, so it could also be an environmental effect.

#### 4.2. Central flux fraction

The amount of flux leftover at the line center is expected to be a good indicator of both the Lyman- $\alpha$  and the LyC escape frac-



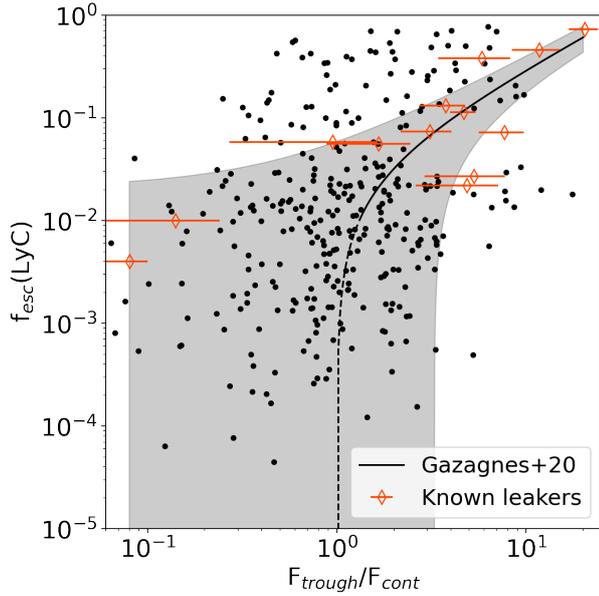
**Fig. 7.** Comparison of our data with the central flux fraction method derived by [Naidu et al. \(2022\)](#). The dashed black line here is the divide between leakers and non-leakers. The points in purple are the points from our sample that fit these criteria and the points in pastel are those that do not. We also show the running median of our sample here with the solid purple line.

tion, as it should correspond to Lyman- $\alpha$  photons that are able to escape at the systemic redshift and have therefore been scattered only a few times. These photons would have escaped through clear sight lines in a partially transparent ISM, where we also expect high  $f_{\text{esc}}(\text{LyC})$ . If there are many clear sight lines, we can expect a third, central peak, making the central flux fraction method derived by [Naidu et al. \(2022\)](#) particularly useful when dealing with Lyman- $\alpha$  profiles that have more than two peaks. [Naidu et al. \(2022\)](#) define the central flux fraction,  $f_{\text{cen}}$ , as the fraction of Lyman- $\alpha$  flux within  $\pm 100 \text{ km s}^{-1}$  of the systemic velocity:

$$\text{Central flux fraction } (f_{\text{cen}}) = \frac{\text{Ly}\alpha \text{ flux at } \pm 100 \text{ km s}^{-1}}{\text{Ly}\alpha \text{ flux at } \pm 1000 \text{ km s}^{-1}}, \quad (13)$$

and show that in their sample of 25 literature LyC leakers at low redshift ( $z \approx 0.3-4$ ) a central flux fraction above 10% corresponds to a high  $f_{\text{esc}}(\text{LyC})$  ( $f_{\text{esc}} > 20\%$ ), although some high  $f_{\text{esc}}(\text{LyC})$  galaxies (two out of nine) are missed with this method. The low  $f_{\text{esc}}(\text{LyC})$  galaxies ( $f_{\text{esc}}(\text{LyC}) < 5\%$ ) are instead selected as having both  $f_{\text{cen}} < 10\%$  and  $v_{\text{sep}} > 375 \text{ km s}^{-1}$ . It must also be noted that the choice of  $\pm 100 \text{ km s}^{-1}$  as the central flux is resolution-dependent and might therefore affect our estimates, as our resolution is higher.

In Fig. 7, we show  $f_{\text{esc}}(\text{LyC})$  as a function of the central flux fraction for our sample, with the divide from [Naidu et al. \(2022\)](#) as the dashed black line: the point on the left of the vertical dashed line would be expected to have  $f_{\text{esc}}(\text{LyC}) < 10\%$  solely based on its  $f_{\text{cen}}$ , while points to the right of the line would have  $f_{\text{esc}}(\text{LyC}) > 20\%$ . In the figure, the dark purple points are those in our sample that fit the criteria from [Naidu et al. \(2022\)](#), and the pastel points are those that do not. We also show the running median with the solid purple line, which clearly lies almost an order of magnitude below the criterion of [Naidu et al. \(2022\)](#), indicating that for any given  $f_{\text{cen}}$  most of the sight lines in our sample would be classified as non-leakers. This implies that for the typical sight line of our simulated LAEs,  $f_{\text{cen}}$  is not



**Fig. 8.** Comparison of our data with the relation found by Gazagnes et al. (2020) between  $f_{\text{esc}}(\text{LyC})$  and the fraction of the flux at the trough of Lyman- $\alpha$  and the continuum flux. The shaded gray area represents the error on the relation. The empty red diamonds are the known leakers used by Gazagnes et al. (2020) to derive this relation. Many of our points fall on their relation but mostly for low escape fractions, whereas high escape fractions are often underestimated.

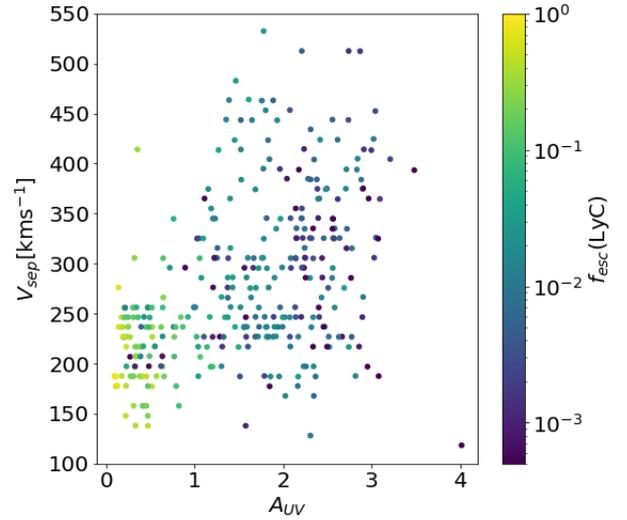
a good predictor of  $f_{\text{esc}}(\text{LyC})$ . Here, we would like to note that the upturn in  $f_{\text{esc}}(\text{LyC})$  at high Lyman- $\alpha$   $f_{\text{cen}}$  is purely due to low number statistics, as the median in the last two bins is only calculated on one data point.

#### 4.3. $F_{\text{trough}}/F_{\text{cont}}$

As was already described in the previous section, we expect leftover Lyman- $\alpha$  flux at the line center to be an indicator of  $f_{\text{esc}}(\text{LyC})$ . The amount of leftover flux can be estimated as the fraction of flux at the line center over the continuum flux, since the presence of flux at the trough of the Lyman- $\alpha$  profile above the continuum should indicate the escape through transparent channels. This ratio could therefore be a reliable indicator of properties of the ISM that favor LyC leakage. Gazagnes et al. (2020) derived a predictor for  $f_{\text{esc}}(\text{LyC})$  based on the ratio between the flux at the trough of the Lyman- $\alpha$  spectrum and the continuum flux, calibrated on a sample of 22 star-forming galaxies with spectroscopy in the rest-frame UV. The relation is as follows:

$$f_{\text{esc}}(\text{LyC}) = (0.032 \pm 0.006) \times \frac{F_{\text{trough}}}{F_{\text{cont}}} - 0.032 \pm 0.053. \quad (14)$$

In Fig. 8, we show our  $f_{\text{esc}}(\text{LyC})$  as a function of the  $\frac{F_{\text{trough}}}{F_{\text{cont}}}$  derived from the synthetic spectra, with the relation found by Gazagnes et al. (2020) shown with the black line, their uncertainty in gray, and the leakers used to derive this relation as empty red diamonds. At  $f_{\text{esc}}(\text{LyC}) < 1\%$  we show the relation in a dashed line to indicate that the relation was extrapolated. Although many of our lines of sight fall within the uncertainty of the relation, this is mostly the case for low LyC escape fractions. From Fig. 8 we can see that these high escape fraction lines of sight are actually often underestimated by this relation: we find



**Fig. 9.** Peak separation against UV dust attenuation, color-coded by  $f_{\text{esc}}(\text{LyC})$ . We find high escape fractions for low peak separations, as was predicted by Izotov et al. (2018b), but only in addition to low UV attenuation.

sight lines with  $f_{\text{esc}}(\text{LyC}) \geq 10\%$  even at  $F_{\text{trough}}/F_{\text{cont}} < 1$  (58 out of a total 452 sight lines have  $F_{\text{trough}}/F_{\text{cont}} < 1$ , so 13%), and 45% of points with high ( $>4$ )  $F_{\text{trough}}/F_{\text{cont}}$  have low  $f_{\text{esc}}(\text{LyC})$ . This large scatter between the  $f_{\text{esc}}(\text{LyC})$  and  $F_{\text{trough}}/F_{\text{cont}}$  could be due to the presence of dust in the line of sight, which affects Lyman- $\alpha$  and LyC emission differently, and by the fact that Lyman- $\alpha$  could scatter back into the line center, which could add flux at  $F_{\text{trough}}$  in low  $f_{\text{esc}}$  sight lines.

#### 4.4. Adding dust attenuation

None of the estimators presented in this section are sufficient to predict the range of escape fractions we see in our data. However, in Fig. 6 we see a clear trend where lines of sight with higher  $A_{\text{UV}}$  tend to have lower  $f_{\text{esc}}(\text{LyC})$  than lines of sight with lower  $A_{\text{UV}}$ . A similar trend has already been shown both in simulations (e.g., Ma et al. 2020; Yeh et al. 2023) and in observations of low-redshift LyC leakers (Chisholm et al. 2022; Saldana-Lopez et al. 2022). We investigate this behavior further in our sample by looking at the peak separation as a function of the dust attenuation, color-coded by the LyC escape fraction, shown in Fig. 9. Here, we can clearly see that the lines of sight with both low peak separation and low dust attenuation are those that have a high LyC escape fraction. In our sample, 68% of sight lines with  $v_{\text{sep}} < 300$  and  $A_{\text{UV}} < 1$  have  $f_{\text{esc}}(\text{LyC}) > 10\%$ , while outside of this selection only 1.6% of sight lines have  $f_{\text{esc}}(\text{LyC}) > 10\%$ . Moreover, both quantities can be found with observations, meaning that it should be possible to get better estimates for LyC escape fractions from observations of Lyman- $\alpha$  profiles if the dust attenuation is also taken into account. We want to stress again that our values for  $f_{\text{esc}}(\text{LyC})$  are not strongly dependent on the presence of dust in the run, and therefore the relation we see is due to the amount of HI in a sight line, which is a quantity traced by dust attenuation.

## 5. Conclusions

In this work, we have studied a sample of 13 galaxies from the OBELISK simulation at redshift  $z = 6$  by performing synthetic

observations of the Lyman- $\alpha$  line. Our sample was chosen to represent a flux-selected sample targeting faint Lyman-break galaxies corresponding to typical LAEs at the end of the EoR.

We have shown that a large variety of Lyman- $\alpha$  profiles can emerge from a single gas distribution, highlighting the challenges that come with the study of this line when it is used to trace intrinsic galaxy properties. We have also shown that although from theory we expected a low Lyman- $\alpha$  peak separation to correlate with a high LyC escape fraction (Verhamme et al. 2017; Izotov et al. 2018b), this does not hold for our sample, where we see a significant number of lines of sight with a low peak separation but also a low Lyman- $\alpha$  and LyC escape fraction.

We have also determined the impact that the choice of dust model has on our results by comparing the same properties found with two different dust models. Although there is a deviation from the 1:1 line in the case of the Lyman- $\alpha$  escape fraction and the UV magnitude, the properties seem robust during a change in the dust model. We have investigated the reason for this deviation and we hypothesize that it could be due to the OBELISK dust being more concentrated around star-forming regions.

Lastly, we compared methods of determining the LyC escape fraction calibrated with observations with the LyC escape fraction that we find from the simulation. We first considered the peak separation method, derived by Izotov et al. (2018b) based on Green Pea galaxies. This method relates the peak separation of the Lyman- $\alpha$  line to the LyC escape fraction. Many of our points do not follow the relation, although here we are able to see a trend between the points that follow the relation and those that do not, which is the amount of dust attenuation in the UV in a given line of sight. We have also compared our values with the central flux fraction method derived by Naidu et al. (2022) and shown that many of our lines of sight have high  $f_{\text{cen}}$  but low  $f_{\text{esc}}(\text{LyC})$ , unlike their prediction. We show that the running median of our  $f_{\text{esc}}(\text{LyC})$  sample is about an order of magnitude lower than the 20% value quoted by Naidu et al. (2022) as the escape fraction expected for galaxies with  $f_{\text{cen}} > 10\%$ . We lastly compared our values with the  $\frac{F_{\text{through}}}{F_{\text{cont}}}$  method, designed by Gazagnes et al. (2020). In this case, most of our lines of sight fall within the uncertainty but mostly at a low escape fraction. Most high-escape-fraction lines of sight, which are the most interesting in a study of the EoR, seem to be underestimated by this method. Following the trend we saw with the dust attenuation, we have looked at the relation between UV dust attenuation, peak separation, and LyC escape fraction. We have seen that lines of sight with low peak separation and UV attenuation are those with the highest LyC escape fractions. Overall, we see that adding a dust attenuation parameter seems to better estimate the LyC escape fractions, and in general we conclude that using multiple parameters is necessary to properly estimate  $f_{\text{esc}}(\text{LyC})$ .

*Acknowledgements.* We thank the referee for the useful comments on an earlier version of the manuscript. We thank Sophia Flury for providing the LzLCS data. MT, VM and PD acknowledge support from the NWO grant 016.VIDI.189.162 (“ODIN”). PD warmly thanks the European Commission’s and University of Groningen’s CO-FUND Rosalind Franklin program for support. This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00072, as well as from the Swiss National Science Foundation (SNSF) through project grant 200020\_207349. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140.

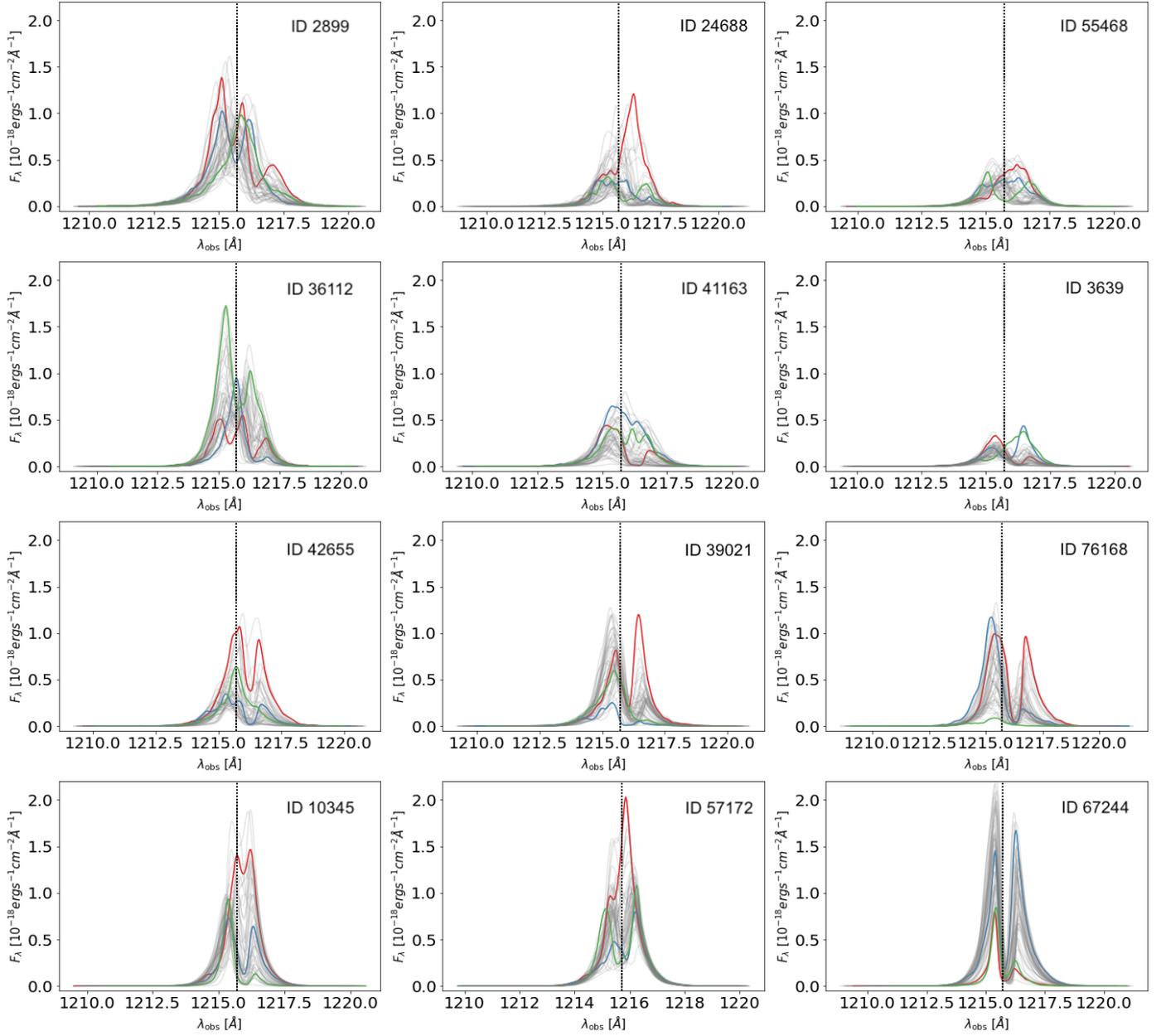
## References

Aubert, D., Pichon, C., & Colombi, S. 2004, *MNRAS*, 352, 376  
Barkana, R., & Loeb, A. 2001, *Phys. Rep.*, 349, 125

- Becker, G. D., D’Aloisio, A., Christenson, H. M., et al. 2021, *MNRAS*, 508, 1853  
Begley, R., Cullen, F., McLure, R. J., et al. 2024, *MNRAS*, 527, 4040  
Behrens, C., Dijkstra, M., & Niemeyer, J. C. 2014, *A&A*, 563, A77  
Bergvall, N., Zackrisson, E., Andersson, B. G., et al. 2006, *A&A*, 448, 513  
Blaizot, J., Garel, T., Verhamme, A., et al. 2023, *MNRAS*, 523, 3749  
Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, *MNRAS*, 514, 55  
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, 811, 140  
Calabretta, M. R., & Roukema, B. F. 2007, *MNRAS*, 381, 865  
Cantalupo, S., Porciani, C., & Lilly, S. J. 2008, *ApJ*, 672, 48  
Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, *MNRAS*, 399, 1191  
Chisholm, J., Prochaska, J. X., Schaerer, D., Gazagnes, S., & Henry, A. 2020, *MNRAS*, 498, 2554  
Chisholm, J., Saldana-Lopez, A., Flury, S., et al. 2022, *MNRAS*, 517, 5104  
Choustikov, N., Katz, H., Saxena, A., et al. 2024, *MNRAS*, 532, 2463  
D’Aloisio, A., McQuinn, M., & Trac, H. 2015, *ApJ*, 813, L38  
Davies, F. B., & Furlanetto, S. R. 2016, *MNRAS*, 460, 1328  
Davis, D., Gebhardt, K., Cooper, E. M., et al. 2021, *ApJ*, 920, 122  
Dayal, P., & Ferrara, A. 2018, *Phys. Rep.*, 780, 1  
Dayal, P., Volonteri, M., Choudhury, T. R., et al. 2020, *MNRAS*, 495, 3065  
Dijkstra, M. 2014, *PASA*, 31  
Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, *ApJ*, 828, 71  
Dubois, Y., Pichon, C., Welker, C., et al. 2014, *MNRAS*, 444, 1453  
Dubroca, B., & Feugeas, J. 1999, *Académie des Sciences Paris Comptes Rendus Série Sciences Mathématiques*, 329, 915  
Eilers, A.-C., Davies, F. B., & Hennawi, J. F. 2018, *ApJ*, 864, 53  
Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, *PASA*, 34  
Fan, X., Carilli, C. L., & Keating, B. 2006, *ARA&A*, 44, 415  
Finkelstein, S. L., D’Aloisio, A., Paardekooper, J.-P., et al. 2019, *ApJ*, 879, 36  
Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, *ApJ*, 878, 87  
Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022a, *ApJ*, 260, 1  
Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022b, *ApJ*, 930, 126  
Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A&A*, 639, A85  
Geen, S., Rosdahl, J., Blaizot, J., Devriendt, J., & Slyz, A. 2015, *MNRAS*, 448, 3248  
Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *A&A*, 578, A83  
Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, *ApJ*, 672, 765  
Goerdt, T., Dekel, A., Sternberg, A., et al. 2010, *MNRAS*, 407, 613  
Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759  
Habouzit, M., Volonteri, M., & Dubois, Y. 2017, *MNRAS*, 468, 3935  
Hayes, M. 2015, *PASA*, 32  
Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, *ApJ*, 730, 5  
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, *ApJ*, 809, 19  
Hirashita, H., & Yan, H. 2009, *MNRAS*, 394, 1061  
Hu, E. M., Cowie, L. L., Songaila, A., et al. 2016, *ApJ*, 825, L7  
Hui, L., & Gnedin, N. Y. 1997, *MNRAS*, 292, 27  
Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016a, *Nature*, 529, 178  
Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, *MNRAS*, 461, 3683  
Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, 474, 4514  
Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, 478, 4851  
Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2021, *MNRAS*, 503, 1734  
Izotov, Y. I., Chisholm, J., Worseck, G., et al. 2022, *MNRAS*, 515, 2864  
Jaskot, A. E., & Oey, M. S. 2013, *ApJ*, 766, 91  
Kakiichi, K., & Gronke, M. 2021, *ApJ*, 908, 30  
Katz, H., Kimm, T., Sijacki, D., & Haehnelt, M. G. 2017, *MNRAS*, 468, 4831  
Kim, K. J., Bayliss, M. B., Rigby, J. R., et al. 2023, *ApJ*, 955, L17  
Kimm, T., & Cen, R. 2014, *ApJ*, 788, 121  
Kimm, T., Cen, R., Devriendt, J., Dubois, Y., & Slyz, A. 2015, *MNRAS*, 451, 2900  
Kimm, T., Katz, H., Haehnelt, M., et al. 2017, *MNRAS*, 466, 4826  
Kimm, T., Blaizot, J., Garel, T., et al. 2019, *MNRAS*, 486, 2215  
Kimm, T., Bieri, R., Geen, S., et al. 2022, *ApJS*, 259, 21  
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, 192, 18  
Laor, A., & Draine, B. T. 1993, *ApJ*, 402, 441  
Laursen, P., Sommer-Larsen, J., & Andersen, A. C. 2009, *ApJ*, 704, 1640  
Leclercq, F., Bacon, R., Wisotzki, L., et al. 2017, *A&A*, 608, A8  
Leclercq, F., Atek, H., Borthakur, S., et al. 2022, *HST Proposal, Cycle 30, ID. #17153*  
Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, *A&A*, 553, A106  
Levermore, C. D. 1984, *J. Quant. Spectrosc. Radiat. Transf.*, 31, 149  
Ma, X., Quataert, E., Wetzel, A., et al. 2020, *MNRAS*, 498, 2001  
Madau, P., & Haardt, F. 2015, *ApJ*, 813, L8  
Maji, M., Verhamme, A., Rosdahl, J., et al. 2022, *A&A*, 663, A66  
Martin, P. G. 1988, *ApJS*, 66, 125  
Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, *ApJ*, 217, 425

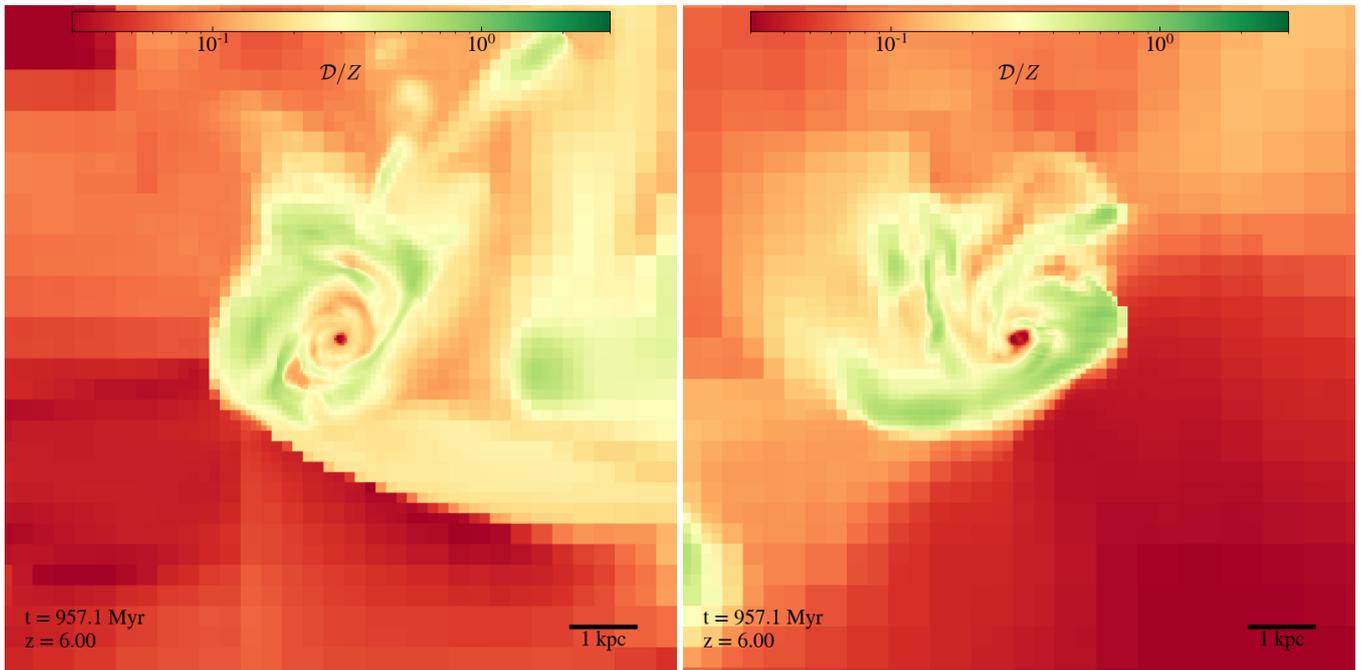
- Matsuoka, Y., Onoue, M., Iwasawa, K., et al. 2023, *ApJ*, 949, L42
- Matthee, J., Sobral, D., Gronke, M., et al. 2018, *A&A*, 619, A136
- Mauerhofer, V., Verhamme, A., Blaizot, J., et al. 2021, *A&A*, 646, A80
- Meyer, R. A., Laporte, N., Ellis, R. S., Verhamme, A., & Garel, T. 2020, *MNRAS*, 500, 558
- Michel-Dansac, L., Blaizot, J., Garel, T., et al. 2020, *A&A*, 635, A154
- Mitchell, P. D., Blaizot, J., Devriendt, J., et al. 2017, *MNRAS*, 474, 4279
- Mitchell, P. D., Schaye, J., & Bower, R. G. 2020, *MNRAS*, 497, 4495
- Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, *ApJ*, 892, 109
- Naidu, R. P., Matthee, J., Oesch, P. A., et al. 2022, *MNRAS*, 510, 4582
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, 442, 900
- Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. 2015, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python
- Onoue, M., Kashikawa, N., Willott, C. J., et al. 2017, *ApJ*, 847, L15
- Orlitová, I., Verhamme, A., Henry, A., et al. 2018, *A&A*, 616, A60
- Orsi, A., Lacey, C. G., & Baugh, C. M. 2012, *MNRAS*, 425, 87
- Ouchi, M., Ono, Y., & Shibuya, T. 2020, *ARA&A*, 58, 617
- Paalvast, M., Verhamme, A., Straka, L. A., et al. 2018, *A&A*, 618, A40
- Parsa, S., Dunlop, J. S., & McLure, R. J. 2017, *MNRAS*, 474, 2904
- Pei, Y. C. 1992, *ApJ*, 395, 130
- Pengelly, R. M. 1964, *MNRAS*, 127, 145
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, 802, L19
- Rosdahl, J., & Teyssier, R. 2015, *MNRAS*, 449, 4380
- Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., & Teyssier, R. 2013, *MNRAS*, 436, 2188
- Saldana-Lopez, A., Schaerer, D., Chisholm, J., et al. 2022, *A&A*, 663, A59
- Saxena, A., Cryer, E., Ellis, R. S., et al. 2022a, *MNRAS*, 517, 1098
- Saxena, A., Pentericci, L., Ellis, R. S., et al. 2022b, *MNRAS*, 511, 120
- Schaerer, D., Izotov, Y. I., Worseck, G., et al. 2022, *A&A*, 658, L11
- Schroeder, J., Mesinger, A., & Haiman, Z. 2012, *MNRAS*, 428, 3058
- Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, *ApJ*, 651, 688
- Smith, A., Ma, X., Bromm, V., et al. 2018, *MNRAS*, 484, 39
- Songaila, A., Hu, E. M., Barger, A. J., et al. 2018, *ApJ*, 859, 91
- Stanway, E. R., & Eldridge, J. J. 2018, *MNRAS*, 479, 75
- Tang, M., Stark, D. P., Chevallard, J., et al. 2021, *MNRAS*, 503, 4105
- Taskesen, E. 2020, Findpeaks is for the detection of peaks and valleys in a 1D vector and 2D array (image)
- Teyssier, R. 2002, *A&A*, 385, 337
- Toro, E. F., Spruce, M., & Speares, W. 1994, *Shock Waves*, 4, 25
- Trebtsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, *MNRAS*, 470, 224
- Trebtsch, M., Dubois, Y., Volonteri, M., et al. 2021, *A&A*, 653, A154
- Trebtsch, M., Dayal, P., Chisholm, J., et al. 2022, arXiv e-prints [arXiv:2212.06177]
- van Leer, B. 1979, *J. Comput. Phys.*, 32, 101
- Verhamme, A., Schaerer, D., & Maselli, A. 2006, *A&A*, 460, 397
- Verhamme, A., Dubois, Y., Blaizot, J., et al. 2012, *A&A*, 546, A111
- Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, *A&A*, 578, A7
- Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, 597, A13
- Weingartner, J. C., & Draine, B. T. 2001, *ApJ*, 548, 296
- Whitney, B. A. 2011, *Bull. Astron. Soc. India*, 39, 101
- Wyithe, J. S. B., Loeb, A., & Carilli, C. 2005, *ApJ*, 628, 575
- Xu, X., Henry, A., Heckman, T., et al. 2022, *ApJ*, 933, 202
- Yang, H., Malhotra, S., Gronke, M., et al. 2017, *ApJ*, 844, 171
- Yeh, J. Y.-C., Smith, A., Kannan, R., et al. 2023, *MNRAS*, 520, 2757
- Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., et al. 2020a, *MNRAS*, 496, 4574
- Yung, L. Y. A., Somerville, R. S., Popping, G., & Finkelstein, S. L. 2020b, *MNRAS*, 494, 1002
- Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., et al. 2021, *MNRAS*, 508, 2706
- Yusef-Zadeh, F., Morris, M., & White, R. L. 1984, *ApJ*, 278, 186
- Zheng, Z., & Miralda-Escude, J. 2002, *ApJ*, 578, 33

**Appendix A: Lyman- $\alpha$  Profiles**



**Fig. A.1.** Similar to Figure 1 but for the rest of the galaxy sample

**Appendix B: Dust-to-metal ratio in OBELISK**



**Fig. B.1.** Dust-to-metal ratio distribution in OBELISK for galaxies ID 2899 (left) and ID 55468 (right). Our constant  $\mathcal{D}/Z$  model would overestimate the amount of dust in all the red and orange pixels in the figure, as the color bar is centered with yellow as the  $\mathcal{D}/Z$  for the LMC