

University of Groningen

## The Low-Redshift Lyman Continuum Survey. Unveiling the ISM properties of low-z Lyman-continuum emitters

Saldana-Lopez, Alberto; Schaerer, Daniel; Chisholm, John; Flury, Sophia R.; Jaskot, Anne E.; Worseck, Gábor; Makan, Kirill; Gazagnes, Simon; Mauerhofer, Valentin; Verhamme, Anne

*Published in:*  
Astronomy and astrophysics

*DOI:*  
[10.1051/0004-6361/202141864](https://doi.org/10.1051/0004-6361/202141864)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2022

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Saldana-Lopez, A., Schaerer, D., Chisholm, J., Flury, S. R., Jaskot, A. E., Worseck, G., Makan, K., Gazagnes, S., Mauerhofer, V., Verhamme, A., Amorín, R. O., Ferguson, H. C., Giavalisco, M., Grazian, A., Hayes, M. J., Heckman, T. M., Henry, A., Ji, Z., Marques-Chaves, R., ... Xu, X. (2022). The Low-Redshift Lyman Continuum Survey. Unveiling the ISM properties of low-z Lyman-continuum emitters. *Astronomy and astrophysics*, 663, [A59]. <https://doi.org/10.1051/0004-6361/202141864>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# The Low-Redshift Lyman Continuum Survey

## Unveiling the ISM properties of low- $z$ Lyman-continuum emitters

Alberto Saldana-Lopez<sup>1</sup>, Daniel Schaerer<sup>1,2</sup>, John Chisholm<sup>3</sup>, Sophia R. Flury<sup>4</sup>, Anne E. Jaskot<sup>5</sup>, Gábor Worseck<sup>6</sup>, Kirill Makan<sup>6</sup>, Simon Gazagnes<sup>7</sup>, Valentin Mauerhofer<sup>1</sup>, Anne Verhamme<sup>1</sup>, Ricardo O. Amorín<sup>8</sup>, Harry C. Ferguson<sup>9</sup>, Mauro Giavalisco<sup>4</sup>, Andrea Grazian<sup>10</sup>, Matthew J. Hayes<sup>11</sup>, Timothy M. Heckman<sup>12</sup>, Alaina Henry<sup>9,12</sup>, Zhiyuan Ji<sup>4</sup>, Rui Marques-Chaves<sup>1</sup>, Stephan R. McCandliss<sup>12</sup>, M. Sally Oey<sup>13</sup>, Göran Östlin<sup>11</sup>, Laura Pentericci<sup>14</sup>, Trinh X. Thuan<sup>15</sup>, Maxime Trebitsch<sup>7</sup>, Eros Vanzella<sup>16</sup>, and Xinfeng Xu<sup>12</sup>

(Affiliations can be found after the references)

Received 23 July 2021 / Accepted 16 December 2021

### ABSTRACT

**Aims.** Combining 66 ultraviolet (UV) spectra and ancillary data from the recent Low-Redshift Lyman Continuum Survey (LzLCS) and 23 LyC observations by earlier studies, we form a statistical sample of star-forming galaxies at  $z \sim 0.2\text{--}0.4$  with which we study the role of cold interstellar medium (ISM) gas in the leakage of ionizing radiation. We also aim to establish empirical relations between the H I neutral and low-ionization state (LIS) absorption lines with different galaxy properties.

**Methods.** We first constrain the massive star content (stellar ages and metallicities) and UV attenuation by fitting the stellar continuum with a combination of simple stellar population models. The models, together with accurate LyC flux measurements, allow us to determine the absolute LyC photon escape fraction for each galaxy ( $f_{\text{esc}}^{\text{abs}}$ ). We then measure the equivalent widths and residual fluxes of multiple H I and LIS lines, and the geometrical covering fraction of the UV emission, adopting the picket-fence model.

**Results.** The LyC escape fraction spans a wide range, with a median  $f_{\text{esc}}^{\text{abs}}$  (0.16, 0.84 quantiles) of 0.04 (0.02, 0.20), and 50 out of the 89 galaxies detected in the LyC ( $1\sigma$  upper limits of  $f_{\text{esc}}^{\text{abs}} \lesssim 0.01$  for non-detections, typically). The H I and LIS line equivalent widths scale with the UV luminosity and attenuation, and inversely with the residual flux of these lines. Additionally, Ly $\alpha$  equivalent widths scale with both the H I and LIS residual fluxes, but anti-correlate with the corresponding H I or LIS equivalent widths. The H I and LIS residual fluxes are correlated, indicating that the neutral gas is spatially traced by the low-ionization transitions. We find that the observed trends of the absorption lines and the UV attenuation are primarily driven by the geometric covering fraction of the gas. The observed nonuniform gas coverage also demonstrates that LyC photons escape through low-column-density channels in the ISM. The equivalent widths and residual fluxes of both the H I and LIS lines strongly correlate with  $f_{\text{esc}}^{\text{abs}}$ : strong LyC leakers (highest  $f_{\text{esc}}^{\text{abs}}$ ) show weak absorption lines, low UV attenuation, and large Ly $\alpha$  equivalent widths. We provide several empirical calibrations to estimate  $f_{\text{esc}}^{\text{abs}}$  from UV absorption lines. Finally, we show that simultaneous UV absorption line and dust attenuation measurements can, in general, predict the escape fraction of galaxies. We apply our method to available measurements of UV LIS lines of 15 star-forming galaxies at  $z \sim 4\text{--}6$  (plus 3 high- $z$  galaxy composites), finding that these high-redshift, UV-bright galaxies ( $M_{\text{UV}} \lesssim -21$ ) may have low escape fractions,  $f_{\text{esc}}^{\text{abs}} \lesssim 0.1$ .

**Conclusions.** UV absorption lines trace the cold ISM gas of galaxies, which governs the physics of the LyC escape. We show that, with some assumptions, the absolute LyC escape can be statistically predicted using UV absorption lines, and the method can be applied to study galaxies across a wide redshift range, including in the epoch of cosmic reionization.

**Key words.** ISM: structure – dust, extinction – galaxies: ISM – galaxies: starburst – galaxies: stellar content – ultraviolet: galaxies

## 1. Introduction

Since the discovery of the Gunn–Peterson effect (Gunn & Peterson 1965; Becker et al. 2001), understanding how the Universe was reionized around  $z \approx 6\text{--}9$  (Planck Collaboration XLVII 2016) requires not only identification of the sources that contributed to that change, but also knowledge of the physical mechanisms under the last major cosmic phase transition. There are two principal instigators of cosmic reionization. The first is massive stars that emit sufficient ionizing photons to keep hydrogen in an ionized state (Robertson et al. 2015), and the second is active galactic nuclei (AGN, Madau & Haardt 2015). Depending on the principal source for reionization (galaxies or AGN, or both; see Dayal et al. 2020), the thermal history of the late Universe changes accordingly (e.g., Kulkarni et al. 2017). For example, models primarily driven by

star-forming (SF) galaxies generally have delayed reionization compared to AGN-dominated reionization, where the Universe is reionized at a higher redshift (e.g., see Finkelstein et al. 2019).

To establish the contribution of galaxies and AGN in terms of ionization rate, one needs to quantify three main properties: (1) the ionizing photon production efficiency ( $\xi_{\text{ion}}$ ), (2) the ionizing photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ), and the number density of such galaxies or AGN at high  $z$ . The ionizing production efficiency ( $\xi_{\text{ion}}$ ) is the intrinsic number of ionizing photons emitted by the stars per unit UV luminosity, while  $f_{\text{esc}}^{\text{abs}}$  is the fraction of these photons escaping from them to the intergalactic medium (IGM).

Although always powerful in terms of ionizing emissivity ( $f_{\text{esc}}^{\text{abs}} \sim 1$ , Stevans et al. 2014; Lusso et al. 2015), AGNs appear to be very rare at high redshifts (Fontanot et al. 2014; Matsuoka et al. 2018) and with lower accretion rates and emissivities than in the late Universe (Dayal et al. 2020); although there is still the

possibility of having an extending low-luminosity AGN population that is numerous enough to provide a substantial contribution to the total ionizing budget (Giallongo et al. 2015; Grazian et al. 2020). Star-forming galaxies are much more numerous at high- $z$ , and the question now becomes whether more massive or compact low-mass SF galaxies contribute the most. On the one hand, the less numerous but massive SF galaxies with enormous gas and dust reservoirs at high- $z$  can intrinsically generate a large amount of Lyman continuum (LyC) photons (Naidu et al. 2020), but these photons can be absorbed and attenuated within the gaseous reservoirs of their host galaxies at the same time. On the other hand, compact low-mass SF galaxies are more numerous at high  $z$  (Rosdahl et al. 2018; Finkelstein et al. 2019). Additionally, they take advantage of their weak gravitational potential and feedback effects to clean the interstellar medium (ISM) from gas and dust, increasing the fraction of LyC photons escaping the galaxy (Trebitsch et al. 2017). The true contribution to reionization from different galaxy types is still under debate, but models require a threshold of  $f_{\text{esc}}^{\text{abs}} = 5\text{--}20\%$  for  $M_{\text{UV}} \lesssim -15$  galaxies to drive reionization around  $z \sim 7$  (Robertson et al. 2013, 2015).

Over the last decade, several campaigns have searched for LyC emitters at low and intermediate redshifts. Using the Far Ultraviolet Spectroscopic Explorer (FUSE), Bergvall et al. (2006) and Leitet et al. (2013) reported the first LyC detections, although later re-observations revealed that some of them were affected by foreground scattered light. Using the *Hubble* Space Telescope (HST), about 30 new leakers were later discovered at  $z \sim 0.3$  (Leitherer et al. 2016; Borthakur et al. 2016; Izotov et al. 2016a,b, 2018a,b, 2021; Puschnig et al. 2017; Wang et al. 2019; Malkan & Malkan 2021), with  $f_{\text{esc}}^{\text{abs}}$  values spanning from a few to 70%. The majority consist of compact SF galaxies with relatively high star formation rates (SFR), strong optical emission lines, and ionizing photon production efficiencies comparable to their high- $z$  analogs (Schaerer et al. 2016). Nevertheless, it remains challenging to robustly detect faint LyC flux at the sensitivity limit of HST (see e.g., the re-analysis of Chisholm et al. 2017 versus Leitherer et al. 2016 and Puschnig et al. 2017).

At intermediate redshifts, investigating the LyC leakage is more difficult, as LyC photons are absorbed by Lyman limit systems in the IGM, which makes the detection extremely challenging beyond  $z \geq 4$ . The presence of low-redshift interlopers also renders potential LyC detections at high  $z$  difficult (Vanzella et al. 2010; Inoue et al. 2014; Naidu et al. 2018). Rivera-Thorsen et al. (2019) recently reported LyC emission from the Sunburst Arc, a lensed compact dwarf galaxy at redshift  $z \sim 2.2$ , with an escape fraction of  $f_{\text{esc}}^{\text{abs}} = 32\%$  averaged over 11 different gravitationally lensed HST images. Similar works are Bian et al. (2017) and the Cosmic Horsehoe in Vasei et al. (2016). The Keck Lyman Continuum Survey (KLCS, Steidel et al. 2018, see also Shapley et al. 2016) spectroscopically confirmed 13 individual LyC detections around  $z \sim 3.1$  in 9 independent fields (see the follow-up HST observations in Mostardi et al. 2015; Pahl et al. 2021). Using HST/F336W and narrow-band photometry, another 12 Lyman break galaxies (LBGs) and AGNs – enclosed in the SSA22 proto-cluster field (cf., Shapley et al. 2006; Micheva et al. 2017a,b) – were discovered to be leaking in the Lyman Continuum Escape Survey (LACES) by Fletcher et al. (2019). In Vanzella et al. (2015, 2016) and de Barros et al. (2016), the authors used the CANDELS  $U$ -band and HST/F336W imaging to select LyC leaker candidates in the GOODS-S field, resulting in the robust detection of the Ion2 galaxy at  $z = 3.2$ . With a similar methodology, Ion1 was discovered by Vanzella et al. (2012) and Ion3 found by

Vanzella et al. (2018), which currently holds the record as the most distant SF galaxy observed to leak LyC photons ( $z \sim 4$ ). The recent re-analysis of Ion1 by Ji et al. (2020) suggests a 5%–10% escape, although this was originally measured at 60% (Vanzella et al. 2012). At  $z \sim 3$ , three individual detections were performed by Grazian et al. (2016, 2017) in the CANDELS/GOODS-N, EGS, and COSMOS fields thanks to deep ground-based observations with the Large Binocular Telescope (LBT). Finally, the recently launched AstroSAT telescope discovered a remarkable  $z = 1.42$  SF galaxy that emits ionizing photons at a rest-frame of 650 Å (Saha et al. 2020). Some other tentative detections have been conducted by Siana et al. (2015) and Smith et al. (2020).

So far, low- $z$  galaxy studies have shown that individual SF galaxies can have  $f_{\text{esc}}^{\text{abs}}$  larger than the 10%–20% threshold for galaxies to reionize the Universe. However, at high  $z$ , the actual number of individual LyC detections is not only low but biased to the sightlines with the highest IGM transmissivity. As the number density and the properties of such high- $z$  leakers are not yet known (and we cannot extrapolate them from the few clear detections), stacking analyses have become the most popular method in reionization studies, and are used to obtain the average  $f_{\text{esc}}^{\text{abs}}$  of the SF galaxy population. Deep broad-band (Grazian et al. 2016, 2017; Rutkowski et al. 2017; Japelj et al. 2017; Naidu et al. 2018), narrow-band imaging (Guaita et al. 2016), and spectral stacking (Marchi et al. 2018; Meštrić et al. 2021) have yielded relatively low upper limits for the escape fraction of the SF population at  $z = 2\text{--}4$  (<5%–10% for  $M_{\text{UV}} \lesssim -19$  galaxies, cf., Naidu et al. 2018; Alavi et al. 2020).

In order to detect and quantify LyC emission at high  $z$  ( $z \geq 4$ ), including in the Epoch of Reionization, the development of indirect techniques that estimate the LyC escape are required. Low- $z$  galaxies are the most suitable laboratories for testing indirect LyC indicators, because the LyC, the far-UV (FUV) stellar, and the interstellar absorption spectra, as well as the rest-frame optical nebular lines, can be observed at the same time and with better S/N than at high- $z$ , allowing spatially resolved studies.

Starting from the closest to the Lyman limit (912 Å), the depth of the Lyman series lines directly links the neutral gas column density to the geometrical distribution of the same gas along the line of sight. The residual flux of the high-order Lyman absorption lines has been related with the neutral gas porosity (see Reddy et al. 2016b, 2022). Properly accounting for dust attenuation, these lines have proved to be an effective LyC tracer (Gazagnes et al. 2018, 2020; Chisholm et al. 2018; Steidel et al. 2018), although simulations show different results (Mauerhofer et al. 2021). Very prominent when present, the  $\text{Ly}\alpha\lambda 1216$  line – accessible from the ground at  $z \geq 2$  – can also be used because  $\text{Ly}\alpha$  and LyC photons interact with the same neutral gas. Jaskot & Oey (2014) and Verhamme et al. (2015) proposed that the  $\text{Ly}\alpha$  peak separation can trace the presence of low-density gas: the physical mechanism driving the leakage is expected to be connected for both  $\text{Ly}\alpha$  and LyC (Rivera-Thorsen et al. 2015; Verhamme et al. 2015; Dijkstra et al. 2016). Indeed, the separation and shape of double- and triple-peaked  $\text{Ly}\alpha$  profiles have been associated with LyC leakage in observations (Rivera-Thorsen et al. 2017; Verhamme et al. 2017; Vanzella et al. 2020; Izotov et al. 2020, 2021). However, the Lyman series are still efficiently suppressed by IGM absorption at high- $z$  and are therefore almost impossible to observe during the Epoch of Reionization. The use of indirect LyC indicators redwards  $\text{Ly}\alpha$  is essential.

The most promising indirect tracers of LyC escape to date are the UV low-ionization state (LIS) absorption lines (Erb 2015; Chisholm et al. 2018) and the  $\text{Mg II}\lambda\lambda 2796, 2803$  emission

doublet (Henry et al. 2018; Chisholm et al. 2020). On the one hand, measurements of some UV metal absorption line depths such as Si II  $\lambda$ 1260 (Gazagnes et al. 2018) and C II  $\lambda$ 1334 (Grimes et al. 2009) have been shown to uniquely trace the neutral gas covering fraction (Reddy et al. 2016b, 2022). The observed depth of a given absorption transition depends on the optical depth and the covering fraction. If the optical depth is high enough that the lines are saturated, then the covering fraction can be inferred from the observed spectral line depths. On the other hand, the intrinsically bright Mg II nebular transitions have an ionization potential similar to that of warm H I. This makes the Mg II and LyC transitions optically thick at similar H I column densities (for typical Mg/H abundances), making Mg II  $\lambda$ 2796,2803 a good tracer of LyC escape as well.

In the rest-frame optical range, some previous studies suggest that both low- $z$  (Jaskot & Oey 2013; Izotov et al. 2018b) and high- $z$  leakers (Nakajima et al. 2016) have high [O III]  $\lambda$ 4959,5007 versus [O II]  $\lambda$ 3726,3729 ratios ( $O_{32}$ , see also Faisst 2016). Subsequent efforts showed that this quantity is not a good predictor of the LyC escape (Bassett et al. 2019; Nakajima et al. 2020), mainly because  $O_{32}$  alone cannot probe the outer extent of the ionization regions, the region that is most sensitive to the optical depth and thus chiefly affects the escape of LyC radiation. Different kinematic effects can also impact the LyC leakage, complicating the use of emission lines in the diagnostics (McKinney et al. 2019; Hogarth et al. 2020). Worth mentioning is also the [S II]  $\lambda$ 6717,6731 doublet (Wang et al. 2019; Ramambason et al. 2020), which shows a deficit in leakers and was recently investigated more thoroughly by Wang et al. (2021).

Finally, other indicators have been investigated in the literature, such as the SFR surface density (Heckman et al. 2001; Clarke & Oey 2002; Izotov et al. 2018b, 2021; Naidu et al. 2020; Cen 2020), and UV/optical lines such as He I  $\lambda$ 3889 (Izotov et al. 2017; Guseva et al. 2020) or O I  $\lambda$ 6300 (Ramambason et al. 2020), but these require further calibrations to be used as robust predictors of LyC escape. A detailed analysis of the various indirect LyC diagnostics is presented in Flury et al. (2022b).

All these indicators are sensitive to the ISM properties, and we can therefore expect correlations between different ISM-related quantities and LyC escape. Revealing the ISM properties of leakers versus nonleakers, one could illustrate the limitations and strengths of the previous tracers. In summary, two scenarios describing the ISM structure compete in the escape of LyC radiation (Zackrisson et al. 2013). The first, called density-bounded regions (Jaskot & Oey 2013; Nakajima & Ouchi 2014), envision a massive stellar population surrounded by a homogeneous and low-density neutral gas media, so that photons can leak from the galaxy at a rate that is inversely proportional to the hydrogen column density. Alternatively, in the “picket-fence” model (Heckman et al. 2001, 2011; Vasei et al. 2016; Reddy et al. 2016b, 2022; Gazagnes et al. 2018; Steidel et al. 2018), stars are surrounded by an optically thick ISM and the ionizing radiation can only escape through low-column-density channels: the ISM is not homogeneous but patchy. So far, local observations of leakers (Gazagnes et al. 2018, 2020) and also simulations (Trebitsch et al. 2017; Mauerhofer et al. 2021) favor a patchy ISM, and only the galaxies with the highest  $f_{\text{esc}}^{\text{abs}}$  have a mostly ionized ISM consistent with a density-bounded scenario (Jaskot et al. 2019; Ramambason et al. 2020; Kakiichi & Gronke 2021). On top of the geometric distribution, other aspects like photoionized outflows and feedback effects (Chisholm et al. 2017; Hogarth et al. 2020; Carr et al. 2021) also change the ISM struc-

ture, affecting the profile of emission/absorption lines and the escape of LyC radiation.

In this paper, we statistically explore the impact of the neutral and low-ionized ISM absorption on LyC escape, taking a novel approach and using a large set of low- $z$  LyC galaxy observations. We use spectra from the HST Low- $z$  Lyman Continuum Survey (LzLCS, see Sect. 2), a set of 66 SF galaxies of which 35 have a secure LyC detection. The methodology for modelling the stellar continuum of the galaxies in the sample, the adopted absolute LyC photon escape fraction, and the absorption line measurements are presented in Sect. 3. The absorption line results are shown in the same section, and the possible inclusion of systematic errors is studied through absorption line simulations. In Sect. 4, we discuss the fitting results, secondary products, and limitations of the modelling. The main empirical relations between H I and LIS lines and different galaxy properties are then given in Sects. 5 and 6. A physical interpretation of the results and literature comparisons are presented in Sect. 7. We finally conclude in Sect. 8, also highlighting the relevance of these results for future high- $z$  searches.

Throughout this paper, a standard  $\Lambda$ CDM cosmology is used, with a matter density parameter  $\Omega_M = 0.3$ , a vacuum energy density parameter  $\Omega_\Lambda = 0.7$ , and Hubble constant of  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . All magnitudes are in the AB system, and we adopt a solar metallicity value of  $12 + \log(\text{O}/\text{H})_\odot = 8.69$  ( $Z_\odot = 1$ ).

## 2. Data

In this study, we use UV spectra in the observed frame wavelength range 800–1950 Å from the Low-Redshift Lyman Continuum Survey (LzLCS, Flury et al. 2022a), the most comprehensive spectroscopic campaign so far to trace the LyC emission of galaxies in the nearby Universe ( $z \sim 0.3$ ). In addition, we include and reanalyze previous LyC emitters from the studies of Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) as a comparison sample. For all the sources, we also use ancillary data derived in a homogeneous fashion by the LzLCS team (see Sect. 2.3).

### 2.1. The Low- $z$ Lyman Continuum Survey

The Low- $z$  Lyman Continuum Survey (LzLCS, see Flury et al. 2022a) is composed of 66 SF galaxies at  $z \sim 0.3$  whose UV spectra were observed by the Cosmic Origin Spectrograph (COS) on board the HST. At  $z \sim 0.3$ , the COS/G140L grating not only probes the LyC but also traces an entire window of wavelengths redder than Ly $\alpha$  up to 1500 Å rest-frame.

The parent sample of LyC-emitter candidates was drawn from Sloan Digital Sky Survey sources (SDSS Data Release 15, Blanton et al. 2017) with GALEX counterparts in both the FUV (0.1530  $\mu\text{m}$ ) and near-UV (NUV; 0.2310  $\mu\text{m}$ ) photometric bands (Morrissey et al. 2007). Applying the methods described in Jaskot et al. (2019), the equivalent widths and fluxes of different optical emission lines were measured in the SDSS spectra, and the SF galaxies in the redshift range  $0.22 \leq z < 0.45$  were selected by applying usual BPT (Baldwin et al. 1981) diagram diagnostics (HST/COS traces the LyC region at  $z > 0.22$ ). Star-formation rate surface densities ( $\Sigma_{\text{SFR}}$ ) were measured from the H $\alpha$  and H $\beta$  fluxes in the SDSS spectra and the SDSS  $u$ -band half-light radius, and UV  $\beta$ -indices for this subsample were computed using GALEX photometry. These measurements were later improved with the COS data in hand.



The final sample was built by selecting the galaxies that fulfill one or more of the following criteria (see Izotov et al. 2016a): high  $[\text{O III}]5007/[\text{O II}]3727$  flux ratios ( $\text{O}_{32} > 3$ ), high SFR densities ( $\Sigma_{\text{SFR}} > 0.1 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ ), and blue UV continuum slopes ( $\beta < -2$ ), resulting in 66 galaxies. Spectra of each galaxy were taken under the GO 15626 HST large program (P.I. Jaskot), with around two orbits per object on average. COS acquired each object via NUV imaging and centered its  $2.5''$  diameter spectroscopic aperture on the peak of the NUV emission. The G140L grating was then used to cover the 800–1950 Å observed wavelength range, with a nominal full width at half maximum (FWHM) resolution of  $R \equiv \lambda/\Delta\lambda \sim 1000$  at 1100 Å, increasing up to  $R \sim 1500$  at 1900 Å.

Science spectra were reduced following the methods described in Flury et al. (2022a) using the standard CALCOS pipeline (v3.3.9). The tackling of the systematic errors and the LyC signal-to-noise ratios (S/Ns) was improved thanks to the use of the FAINTCOS custom software (Makan et al. 2021), which properly estimates the dark current and scattered geocoronal Ly $\alpha$ , co-adding spectra from multiple visits to improve the S/N. Finally, each spectrum was binned to 0.5621 Å to yield a sampling of  $\geq 2$  pixels per resolution element. The spectra were corrected for Milky Way extinction using the Galactic  $E_{B-V}$  estimates from the Green (2018) dust maps and the Fitzpatrick (1999) attenuation law.

The LyC flux ( $f^{\text{LyC}}$ ) was measured as the mean over a 20 Å window as close as possible to but shortward of 900 Å in the rest-frame (typically 850–900 Å), while avoiding wavelengths redward of the observed frame of 1180 Å to reduce the impact of geo-coronal Ly $\alpha$  emission. Over the 66 galaxies in the LzLCS, 35 were significantly detected (LyC flux  $S/N \geq 2$ ) and 31 were not detected at LyC wavelengths. In the detected galaxies, LyC luminosities span from  $10^{38}$  to  $10^{40} \text{ erg s}^{-1}$ . In the undetected sample, the average upper uncertainty on the LyC flux density was reported as the  $1\sigma$  upper limit. See Flury et al. (2022a) for details.

## 2.2. Low- $z$ LyC emitters in the literature

In addition to the new LzLCS sources, we add 20 SF galaxies with LyC observations from the sample of Izotov et al. (2016a,b, 2018a,b, 2021). Some of these sources are among the strongest leakers in the literature at  $z \sim 0.3$  (i.e., with the highest ionizing escape fractions).

Originally, the Izotov et al. (2016a,b, 2018a,b) galaxies were selected from SDSS in a similar way to some of the LzLCS in terms of compactness, FUV brightness, and ionizing features in the optical spectra ( $\text{EW}(\text{H}\beta) > 100 \text{ \AA}$  and  $\text{O}_{32} \geq 5$ ). The selected galaxies exhibit narrow ( $\leq 130 \text{ km s}^{-1}$ ) double-peaked Ly $\alpha$  emission and high (up to 60%) Ly $\alpha$  escape fractions in most cases, low metallicities ( $1/10$ – $1/3 Z_{\odot}$ ), low stellar masses ( $0.15$ – $6 \times 10^9 M_{\odot}$ ), and high SFRs ( $14$ – $80 M_{\odot} \text{ yr}^{-1}$ ) and SFR densities ( $2$ – $35 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ ). The resulting LyC escape in those galaxies extends from a few up to  $\sim 70\%$ . Here, we used these Izotov et al. SF galaxies as a reference sample for the strongest leakers in the new LzLCS.

Similarly, the Izotov et al. (2021) sample consists of 9 dust-poor and low-mass SF galaxies ( $\leq 10^8 M_{\odot}$ ), with very high specific SFRs ( $\sim 150$ – $630 \text{ Gyr}^{-1}$ ), showing a wide range in Ly $\alpha$  peak velocity separations ( $\sim 200$ – $500 \text{ km s}^{-1}$ ). Their published LyC escape fractions reach  $\sim 35\%$ .

In order to increase the dynamic range for some of the galaxy properties in our working sample, we finally also included three more SF galaxies observed by Wang et al. (2019), which

were selected in terms of their [SII] deficiency (SDSS spectra), compactness ( $\leq 0''.5$  in the SDSS  $u$ -band), and relatively high GALEX FUV fluxes. These galaxies are more massive ( $\sim 3 \times 10^{10} M_{\odot}$ ), more metal-rich ( $0.9$ – $1 Z_{\odot}$ ), dustier, and less extreme in terms of ionization properties than the Izotov et al. (2016a,b, 2018a,b) sample, but also have exceptionally high SFR densities ( $35$ – $700 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ ). Their estimated escape fractions are significantly lower (up to 3%).

The LzLCS team re-processed and re-reduced these 23 COS/G140L spectra with the same methodology as that described above in order to maintain the homogeneity and consistency of the results (see Flury et al. 2022a). This results in 15 additional detections and 8 nondetections out of a total of 23 additional galaxies. The same total sample is described and analyzed in depth in Flury et al. (2022a,b).

## 2.3. Other LzLCS science products

The LzLCS team measured galaxy properties of every source in the LzLCS and in the selected literature samples, either from the SDSS spectra and the GALEX photometry, or directly from the COS spectra. In particular, for this work we use:

- Optical dust attenuation parameters:  $E_{B-V}$ -Balmer.
- Gas-phase metallicities:  $12 + \log(\text{O}/\text{H})$ .
- Ly $\alpha$  equivalent widths:  $W_{\text{Ly}\alpha}$ .

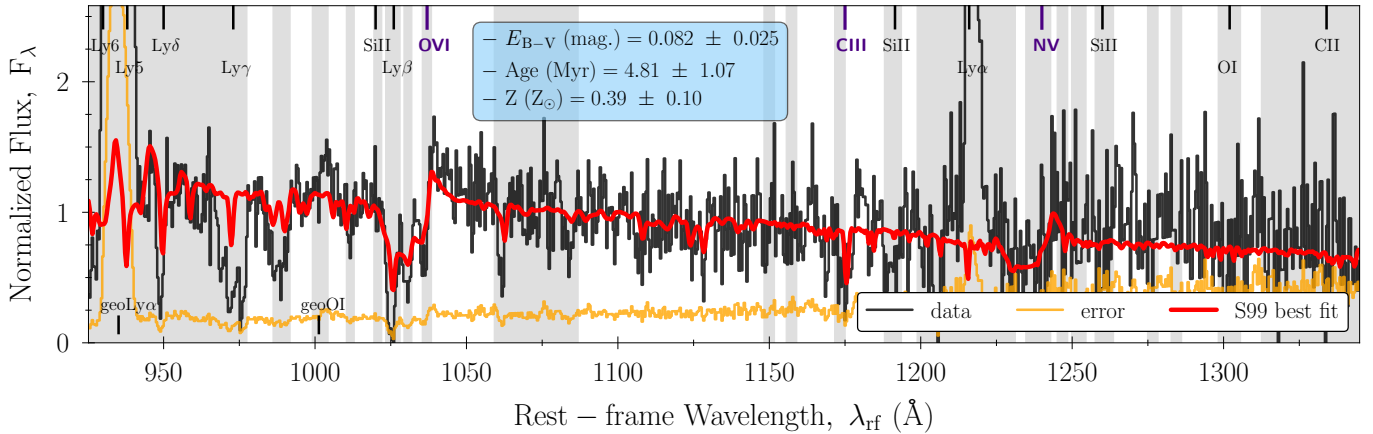
Optical attenuations were computed using an iterative Balmer decrement methodology from the measured SDSS H $\beta$  to H $\epsilon$  fluxes (following Izotov et al. 1994, see also Flury & Moran 2020). The gas-phase metallicities (relative oxygen abundances) were obtained using a direct-method approach using the extinction-corrected hydrogen and oxygen optical lines. Finally, Ly $\alpha$  equivalent widths were measured by integrating the COS spectral flux close to the Ly $\alpha$  line (continuum linearly fitted within a 100 Å window). Errors on every quantity come from  $10^4$  Monte-Carlo realizations of the corresponding spectra. See Flury et al. (2022a) for more details. The LzLCS team is also measuring UV morphological parameters from the COS acquisition images (Ji et al., in prep.).

## 3. Methods

Measuring the properties of the ISM features in the UV requires accurate determination of the stellar continuum. In this section, we describe our methodology for modelling the UV stellar continuum of extragalactic spectra (Sect. 3.1, following Chisholm et al. 2019) and absolute LyC escape fractions in detail. The expressions used to determine the equivalent widths, residual fluxes, and corresponding escape fractions from the lines are also listed (Sect. 3.2).

### 3.1. Stellar continuum modelling

The stellar continuum modelling is achieved by fitting every COS spectrum with a linear combination of multiple bursts of single-age and single-metallicity stellar population models. We use the fully theoretical STARBURST99 single-star models without stellar rotation (S99, Leitherer et al. 2011a, 2014) using the Geneva evolution models (Meynet et al. 1994), and computed with the WM-BASIC method (Pauldrach et al. 2001; Leitherer et al. 2010). The S99 models use a Kroupa initial mass function (Kroupa 2001) with a high-(low-)mass exponent of 2.3 (1.3), and a high-mass cutoff at  $100 M_{\odot}$ . The spectral resolution of the

J091703 ( $z = 0.30035$ ) – stellar continuum


**Fig. 1.** Stellar continuum modelling for the galaxy *J091703*. Stellar continuum fit (red), observed spectrum (black), and error spectrum (yellow); spectral regions masked during the fit are shown in gray. The spectra, shown in  $F_\lambda$  units, have been normalized over 1070–1100 Å. Different ISM absorption lines and stellar features are indicated with black and dark-blue vertical lines in the top part of the figure, respectively. Geo-coronal emissions are labeled at the bottom. The best-fit dust attenuation color-excess ( $E_{B-V}$ , in mag, R16 law), light-weighted stellar age (Myr) and metallicity ( $Z_\odot$ ) inferred for this source are given in the inset. Uncertainties come from a consistent Monte-Carlo error propagation along the fitting routine.

models is  $R \equiv \lambda/\Delta\lambda \sim 2500$ , and remains constant at  $\leq 1500$  Å rest-frame.

A nebular continuum contribution was added to every single population by processing the stellar continuum models through the CLOUDY v17.0 code (Ferland et al. 2017), assuming similar gas-phase and stellar metallicities, an ionization parameter of  $\log(U) = -2.5$ , and a volume hydrogen density of  $n_H = 100 \text{ cm}^{-3}$ . The nebular continuum is negligible for the wavelengths that are typically fitted in this paper, but it can impact longer wavelength data ( $\geq 1200$  Å rest-frame).

We chose four different metallicities (0.05, 0.2, 0.4 and  $1 Z_\odot$ ) and seven ages for each metallicity (1, 2, 3, 4, 5, 8, and 10 Myr) as a representative set of 28 models for our low- $z$  UV spectra. In detail, ages are chosen to densely sample the stellar ages where the stellar continuum features appreciable change (Chisholm et al. 2019).

The observed spectra are first placed into the rest-frame by dividing by  $(1+z)$ . Both the observed spectra and the models are then normalized by the median flux in the 1070–1100 Å rest-frame wavelength interval (free of stellar and ISM features), and all the fits are performed in the same wavelength range: 975–1345 Å (rest-frame). Finally, the models are convolved by a Gaussian kernel<sup>1</sup> to the COS/G140L nominal spectral resolution ( $R \sim 1000$ ).

We first mask out the nonstellar features, the spectral regions that are affected by host-galaxy and Milky Way ISM absorptions, and geo-coronal emissions. We also mask bad pixels with zero flux, as well as pixels with  $S/N < 1$ . We then fit the UV stellar continuum ( $F^*(\lambda)$ ) with a linear combination of multiple S99 models ( $F_{i,j}^{S99}$ ) plus the nebular continuum. Adopting a uniform foreground dust-attenuation model for the galaxies (see Sect. 3.2), this yields:

$$F^*(\lambda) = 10^{-0.4k_\lambda E_{B-V}} \sum_{i,j} X_i F_{i,j}^{S99}, \quad (1)$$

$$i \equiv 1, 2, 3, 4, 5, 8, 10 \text{ Myr}$$

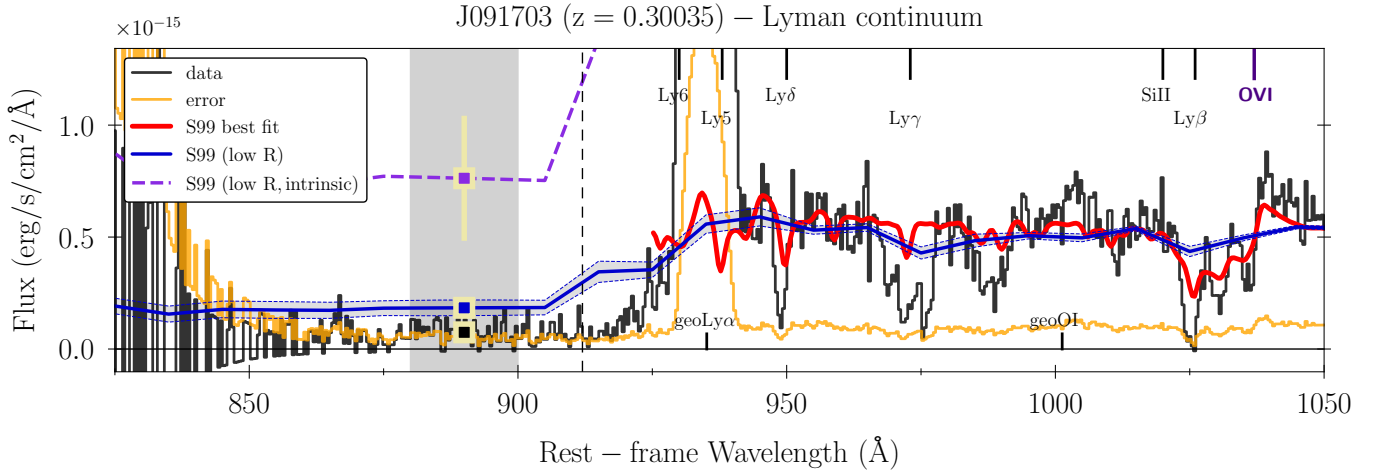
$$j \equiv 0.05, 0.2, 0.4, 1 Z_\odot$$

<sup>1</sup> The high nonGaussianity of the COS line spread function (LSF) does not have a strong impact on the current analysis.

where  $10^{-0.4A_\lambda}$  is the UV attenuation term,  $A_\lambda = k_\lambda E_{B-V}$ , and  $k_\lambda$  is given by the Reddy et al. (2016a) attenuation law (R16 hereafter),  $k_\lambda = 2.191 + 0.974/\lambda(\mu\text{m})$ . We also tested the influence of different attenuation laws on the fitting results, in particular using the Small Magellanic Cloud (SMC) law instead (Prevot et al. 1984; Bouchet et al. 1985). We discuss these results in the next Sect. 4. The light-fraction coefficients ( $X_i$ ) determine the weight of the single-stellar population within the fit, and the best fit is chosen through a nonlinear  $\chi^2$  minimization algorithm with respect to the observed data (LMFIT<sup>2</sup> package, Newville et al. 2016). Errors are derived in a Monte-Carlo way, varying observed pixel fluxes by a Gaussian distribution whose mean is zero and standard deviation is the  $1\sigma$ -error of the flux in the same pixel, and re-fitting the continuum over 500 iterations per spectrum (enough realizations to sample the posterior continuum so that it approaches ‘‘Gaussianity’’ on each pixel). The adopted  $1\sigma$ -errors on the spectra are approximated as the squared-root average of the original asymmetric spectral uncertainties (Flury et al. 2022a).

Figure 1 shows one of the outputs of our fitting routine, the observed and fitted spectrum for *J091703*, with the stellar continuum in red overlaying the observed spectrum in black. The reduced- $\chi^2$  of the fit is  $\chi^2_\nu = 0.93$ , with a median of 1.03 for the whole sample. This galaxy is dominated by a young and low-metallicity population whose average light-weighted age is 4.81 Myr and light-weighted metallicity results in  $0.39 Z_\odot$ . *J091703* is moderately attenuated, with a UV dust-attenuation parameter of  $E_{B-V} = 0.082$  mag. We highlight with dark-blue labels the two main stellar features that help our algorithm to estimate the age and the metallicity of the stellar populations at those wavelengths: the O VI  $\lambda 11032$  and N V  $\lambda 1240$  P Cygni stellar wind profiles (see also the C II  $\lambda 1175$  photospheric line). As, in general, the FUV spectrum is dominated by young massive stars, we do not expect large changes between the STARBURST99 and other models including binaries (e.g., BPASS, Eldridge et al. 2017). Detailed results regarding the stellar populations will be discussed in a forthcoming paper (Chisholm et al., in prep.).

<sup>2</sup> A PYTHON version of the LMFIT package can be found in <https://lmfit.github.io/lmfit-py/>



**Fig. 2.** Stellar continuum modelling in the LyC region for the *J091703* galaxy. The high-resolution stellar continuum fit (down to 925 Å) and the observed *J091703* flux and error are plotted in red, black, and yellow lines, respectively. The stellar continuum is now reproduced with models of lower resolution (blue line and shaded area), allowing us to go below the Lyman limit (<912 Å, dashed vertical line). The intrinsic (unreddened) flux is shown by the purple dashed line. In the LyC window (gray-shaded), the measured, modeled, and modeled-intrinsic mean LyC fluxes are indicated with filled squares and corresponding error bars. The ISM, stellar lines, and geo-coronal lines are displayed with the same color coding as in Fig. 1.

All the data sets are fitted using the same method. To conclude, the main derived quantities of interest for the present paper are the UV dust attenuation parameters ( $E_{B-V}$ ) and the fitted stellar spectral energy distributions (SEDs). The intrinsic SEDs (before UV attenuation) are used below to derive important quantities such as the absolute LyC escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and intrinsic UV luminosities. The UV-fit  $f_{\text{esc}}^{\text{abs}}$  values derived here are compared to other escape fraction estimates in Flury et al. (2022a,b).

**Fiducial photon escape fractions.** Due to the fact that the high-resolution S99 models are only well defined above the Lyman limit (912 Å), we use the low-resolution S99 models ( $R \sim 100$ ) to extend to the LyC region (Leitherer et al. 2011b). These low-resolution models are linearly combined with the same light-fractions and attenuation parameters as the previous fits.

The absolute LyC escape fraction at  $\sim 900$  Å (denoted by  $f_{\text{esc}}^{\text{abs}, 900} \equiv f_{\text{esc}}^{\text{abs}}$  in the following) is then defined as the ratio between the modeled intrinsic LyC flux (i.e., the flux given by the low-resolution models) and the observed (measured, see Sect. 2) LyC flux for every galaxy ( $f^{\text{LyC}}$ ):

$$f_{\text{esc}}^{\text{abs}} = \frac{f^{\text{LyC}}}{\left(\sum_{i,j} X_i F_{i,j}^{\text{S99}}\right) (\text{LyC})}. \quad (2)$$

The modeled LyC flux is computed as the mean flux over the same 20 Å window as the observed LyC flux, which is typically at 850–900 Å, although varying from galaxy to galaxy (Flury et al. 2022a). For galaxies with nondetection in the LyC, we adopt the corresponding  $1\sigma$  COS/G140L upper limit. Error bars on  $f_{\text{esc}}^{\text{abs}}$  come from the simultaneous sampling of  $f^{\text{LyC}}$  and the modeled flux. In the case of LyC detections, these error bars are dominated by the uncertainty in the modeled flux. Figure 2 shows the low-resolution S99 modeled flux for *J091703* (blue shaded region; see the figure caption) and emphasises this approach. The wavelength window over which the LyC flux is measured is shown by the gray region. Dividing the mean

observed flux value in this window (black square) by the modeled intrinsic stellar flux (purple square) yields the absolute photon escape of this source ( $f_{\text{esc}}^{\text{abs}} = 0.161^{+0.073}_{-0.055}$ ).

Throughout this paper, we use the definition of the photon escape given by Eq. (2). However, we note that similar trends and results are obtained if we use other methods to determine  $f_{\text{esc}}^{\text{abs}}$  (see Flury et al. 2022b). The main quantities ( $E_{B-V}$  attenuation, stellar age, metallicity, and LyC escape fraction) derived for our sample are listed in Appendix A.

Together with the LyC flux significance (i.e., the S/N in the LyC flux), the inclusion of the escape fraction makes it possible to distinguish between strong, weak, and nonleakers. As in Flury et al. (2022b), we define strong leakers as those sources with secure LyC detection (significance > 5) and  $f_{\text{esc}}^{\text{abs}} \geq 0.05$ . Weak leakers are considered to be galaxies with fair or marginal LyC detection (significance  $\geq 2$ ), but not falling in the strong category. Finally, nonleakers are defined as undetected galaxies in the LyC (significance < 2). The total sample of 89 galaxies is then divided into 20, 30, and 39 strong, weak, and nonleakers, respectively (see also Appendix A).

## 3.2. Absorption lines

### 3.2.1. Equivalent widths and residual fluxes

In the rest-frame spectral range covered by our observations ( $\sim 800$ – $1500$  Å), several lines of interest are present. Here we choose the following: five Lyman series lines of hydrogen, from Ly $\beta$  to Ly6 (H $\lambda$ 1026, H $\lambda$ 973, H $\lambda$ 950, H $\lambda$ 938, H $\lambda$ 930), and six low-ionization state (LIS) lines (Si II $\lambda$ 989, Si II $\lambda$ 1020, Si II $\lambda$ 1190,1193, Si II $\lambda$ 1260, O I $\lambda$ 1302, C II $\lambda$ 1334). Given the actual COS spectral resolution, some of the mentioned lines are blended with other lines at adjacent rest-frame wavelengths. In particular, the Lyman series lines are blended with O I transitions, and O I $\lambda$ 1302 is also folded with the Si II $\lambda$ 1304 line. Si II $\lambda$ 989 is blended with O I $\lambda$ 988 and N III $\lambda$ 989.

Our goal is to measure the equivalent widths and residual fluxes of these different lines (when present) for all the galaxies in the LzLCS and literature samples. We first find the wavelength of the minimum depth of the line (minimum flux). We



then define a  $\pm 1250 \text{ km s}^{-1}$  integration window in velocity space ( $\Delta\lambda_i$ , chosen by visual inspection) centered on this wavelength, over which the equivalent width  $W_{\lambda_i}$  is computed as:

$$W_{\lambda_i} = \int_{\Delta\lambda_i} (1 - f_{\lambda}/F^*(\lambda)) d\lambda, \quad (3)$$

where  $f_{\lambda}$  is the observed spectral flux density, and  $F^*(\lambda)$  is the modeled stellar continuum. We also compute the residual flux of the line,  $R(\lambda_i) = \langle f_{\lambda}/F^* \rangle$ , as the median over an interval of  $\pm 150 \text{ km s}^{-1}$  around the minimum flux (to avoid biasing towards minimum depth pixels affected by noise). Errors on the observed spectra plus the error on the stellar continuum fitting are incorporated within every absorption line uncertainty. More precisely, we take the median and standard deviation of the  $W_{\lambda_i}$  and  $R_{\lambda_i}$  distributions (over the 500 realizations of the perturbed spectra and subsequent continuum). We consider these as the final measurement and error values of each line.

The drop in the COS/G140L sensitivity redward of  $1600 \text{ \AA}$  in the observed frame reduces the S/N of the measured flux, thus providing poorer constraints on the stellar continuum and lowering the S/N of the LIS lines at wavelengths longer than  $\text{Ly}\alpha\lambda 1216$  ( $\sim 1600 \text{ \AA}$  at  $z = 0.3$ ). Then, and only for the LIS measurements, we prefer to locally model the stellar continuum by a linear function. Equivalent widths and residual fluxes for  $\text{Si II}\lambda 1260$ ,  $\text{O I}\lambda 1302$ , and  $\text{C II}\lambda 1334$  lines are measured assuming this linear continuum, and using the same previous equations.

Ideally, we also want to determine the covering fraction of the lines. In general, the relation between the residual flux and the covering fraction depends on (1) the assumed gas and dust geometries, (2) the different velocity components of the gas clouds that potentially contribute to the line shape, and (3) the degree of saturation of the line, which we discuss in the forthcoming sections.

### 3.2.2. The covering fraction: assumptions about the ISM geometry

The escape of ionizing radiation and the UV absorption features of a galaxy can be explained by the so-called picket-fence model (e.g., Heckman et al. 2001, 2011; Vasei et al. 2016; Reddy et al. 2016b; Gazagnes et al. 2018; Steidel et al. 2018). This model describes a stellar ionizing radiation field surrounded by a nonuniform or patchy ISM, where the neutral and metallic gas are distributed in high-column-density clouds, with low-column-density channels in between. Two gas and dust relative geometries can be assumed. In the first, a uniform foreground dust-screen envelops the whole galaxy (A) while in the second, dust is hosted in the discrete gas clouds of the ISM (B).

A key parameter of the picket-fence model is the geometrical distribution of the gas parameterized by the covering fraction,  $C_f(\lambda_i)$ . This variable represents the fraction of sight lines that are covered by optically thick gas at wavelength  $\lambda_i$ , in contrast to sight lines with low amounts of gas or no gas at all. Observationally, the covering fraction can be related to the residual flux of individual absorption lines ( $R(\lambda_i) = \langle f_{\lambda}/F^* \rangle$ ). The functional form of this dependency changes with the gas and dust geometry (e.g., Vasei et al. 2016; Gazagnes et al. 2018), and always assumes absorption lines are saturated (i.e., the column density of the medium is high enough so that it is optically thick at the  $\lambda_i$  transition). Furthermore, given the limited spectral resolution, we assume that the lines are described by a single gas component or, in other words, that all velocity components of the gas have the same covering fraction.

For geometry A, the gas covering is simply the complement of the residual flux, and can be measured from the line depths as

$$R(\lambda_i) = 1 - C_f(\lambda_i), \quad (4)$$

while for geometry B, the relation depends also on the dust attenuation (see detailed derivation in Vasei et al. 2016; Reddy et al. 2016b; Gazagnes et al. 2018):

$$R(\lambda_i) = \frac{1 - C_f(\lambda_i)}{C_f(\text{HI}) \times 10^{-0.4k_{\lambda_i} E_{\text{B-V}}^{\text{B}}} + (1 - C_f(\text{HI}))}, \quad (5)$$

where  $E_{\text{B-V}}^{\text{B}}$  is the dust attenuation measured according to geometry B, and  $C_f(\text{HI})$  is the neutral gas covering fraction. The HI and the dust have the same covering fraction according to this model. Both the stellar continuum and the residual flux of the lines must be determined consistently with the same geometry (under the same gas and dust distribution), but the actual S/N for most of the spectra does not allow a stellar continuum fitting using geometry B (Gazagnes et al. 2018; Steidel et al. 2018). Therefore, in this paper we adopt geometry A. However, our absorption line calculations will provide insights into the underlying gas and dust geometry and other ISM properties of low- $z$  LyC-emitting galaxies (see Sect. 5).

### 3.2.3. Saturation regime for the HI and LIS lines

Applying Eq. (4) (or Eq. (5)) to compute the gas covering fraction intrinsically assumes that absorption lines are saturated. The condition of saturation is tested using the equivalent width ratios for lines of the same ion (see Appendix B).

The saturation of HI is studied through the ratio of the Lyman series lines  $\text{Ly}\beta$  ( $\text{H I}\lambda 1026$ ) and  $\text{Ly}\delta$  ( $\text{H I}\lambda 950$ ). The  $\text{Ly}\beta$ -to- $\text{Ly}\delta$  equivalent width comparisons reveal that most of the sources fall in the optically thick (saturation) region of the diagram (see Draine 2011), while some of them may be compatible with an optically thin regime, although many of them have large uncertainties at the same time. We caution that some of the discrepancies in this analysis can be due to slight underestimations of the continuum level, or because of the O I contribution, whose lines are blended with the Lyman series. For the metals, we use the Si II as a representative ion, and compare the ratio of Si II  $\lambda 1260$  to Si II  $\lambda 1190, 1193$ . Nevertheless, this comparison is subject to large measurement errors and, although we can be statistically sure about the HI saturation, we can neither discard nor assume saturation for the metallic lines.

Henceforth, we assume that all the HI transitions and metal lines considered are saturated. A more detailed explanation of this reasoning is presented in Appendix B.

### 3.2.4. Estimation of systematic errors

We finally study the possible systematic errors on our measurements. When measuring absorption line properties from observed spectra, the instrumental resolution ( $R$ ) tends to make the lines broader and so residual fluxes can be overestimated. In addition, the observed S/N of every pixel – due to instrumental sensitivity, noise, etc. – also impacts the line shape and can therefore affect our estimates.

To account for these effects, we simulate different absorption lines ( $\text{Ly}\delta$ ,  $\text{Ly}\beta$ , Si II  $\lambda 1260$ , C II  $\lambda 1334$ ) with a standard Voigt profile (Draine 2011) and typical saturated column densities and Doppler broadening in SF galaxies, a single velocity component, and a fixed covering fraction of  $C_f = 0.85$  for all of them. A



foreground dust-screen geometry of the picket-fence model is adopted. We then degrade to the COS/G140L instrumental resolution ( $R \sim 1000$ ) and finally add a set of different random Gaussian S/N per pixel. After that, the residual flux is measured 500 times for every S/N value, and converted to a covering fraction using Eq. (4). The median covering fraction is compared to the original input value to give the relative difference. We conclude that, for the typical S/N measured in our sample at the continuum level close to the lines in question (e.g., close to H I  $\lambda$ 950 the continuum is  $S/N \sim 4$ ) the systematic error on the covering fraction spans between 5% and 20%, with a bias towards lower coverings. However, this value does not dominate over original errors and is within the original error bars we are reporting for single covering fraction measurements (see Sect. 5). In addition, given the actual instrumental resolution, systematic deviations do not change significantly with respect to  $C_f = 0.85$  when other input  $C_f$  is chosen in the simulations. Even in the case when the residual flux of single lines could be slightly overestimated, the general trends in this study will hold, because all the spectra in the working sample have the same resolution. We therefore leave our initial residual flux measurements and errors unaltered. A more exhaustive explanation of the estimation of systematic errors is presented in Appendix C.

### 3.3. Photon escape fraction from the Lyman series and LIS metallic lines

In the above-described picket-fence model, the absolute LyC photon escape fraction is related to the H I covering fraction of saturated lines and the dust attenuation (Heckman et al. 2001, 2011; Vasei et al. 2016; Reddy et al. 2016b; Gazagnes et al. 2018, 2020; Chisholm et al. 2018; Steidel et al. 2018). The expression of the predicted  $f_{\text{esc}}^{\text{abs}}$  depends on the assumed gas and dust geometry. Nonetheless, as demonstrated by Gazagnes et al. (2018), Chisholm et al. (2018), although  $E_{B-V}$  and  $C_f(\text{H I})$  are model dependent, regardless of the use of geometry (A) or (B), the attenuation and covering fraction terms balance each other so that the predicted  $f_{\text{esc}}^{\text{abs}}$  remains almost unaltered (e.g., it does not strongly depend on the assumed geometry). This is because, for the same observed flux, model (B) would require larger dust obscuration but also higher gas covering fraction than model (A) in order to equally match the data.

Consequently, we decide to adopt a uniform dust screen model (A) consistently for the continuum and the  $f_{\text{esc}}^{\text{abs}}$  determination from the absorption lines. Covering fractions are therefore simply given by Eq. (4), and finally the predicted photon escape from H I is (cf., Gazagnes et al. 2018; Chisholm et al. 2018)

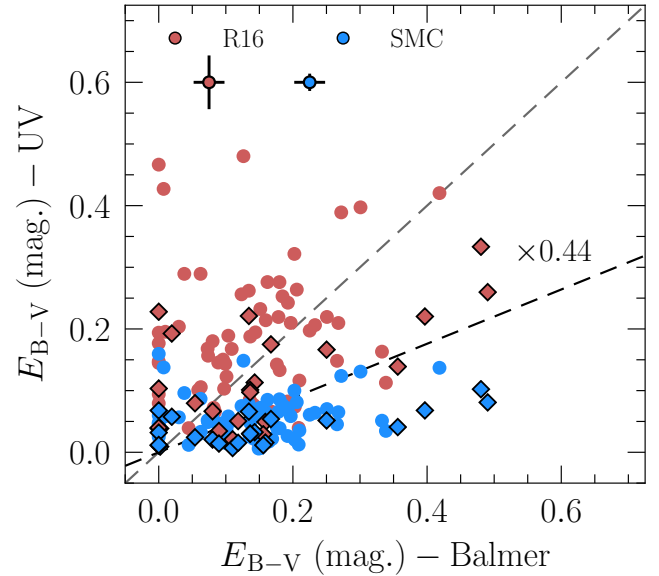
$$f_{\text{esc}}^{\text{abs, HI}} = 10^{-0.4k_{912}E_{B-V}} \times (1 - C_f(\text{H I})), \quad (6)$$

where  $k_{912}$  depends on the assumed attenuation law at 912 Å (e.g.,  $k_{912} \approx 12.87$  for R16).  $C_f(\text{H I})$  is the average H I covering fraction. To compute this, we first use Eq. (4) to figure out individual covering fraction measurements for each line. Then, in order to improve the S/N in single measurements of  $C_f(\lambda_i)$ , we use an inverse-variance-weighted mean definition for the H I covering fraction:

$$C_f(\text{H I}) \pm \Delta C_f(\text{H I}) = \frac{\sum_{i=1}^N C_{fi} \times \frac{1}{\sigma_{C_i}^2}}{\sum_{i=1}^N \frac{1}{\sigma_{C_i}^2}} \pm \sqrt{\frac{1}{\sum_{i=1}^N \sigma_{C_i}^2}}, \quad (7)$$

$$i \equiv \text{H I } \lambda 930, \text{ H I } \lambda 938, \text{ H I } \lambda 950, \text{ H I } \lambda 973, \text{ H I } \lambda 1026$$

where  $N \leq 5$  is the number of Lyman lines that can be measured on a single spectra, and  $\sigma_{C_i}$  is the measurement error (from



**Fig. 3.** Comparison between the optical attenuation parameter  $E_{B-V}$  (color excess, in mag) obtained from the Balmer decrements (CCM attenuation law) and the values derived from the UV-SED fits to the COS spectra. Here we explore two different attenuation laws: R16 (red) and SMC (blue symbols). Median error bars are plotted in the upper left. Circles represent the LzLCS galaxies while diamonds are the Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) sources. The theoretical  $\times 0.44$  shift between the stellar and nebular attenuation, which assumes a foreground screen of dust (Calzetti et al. 2000), is also shown.

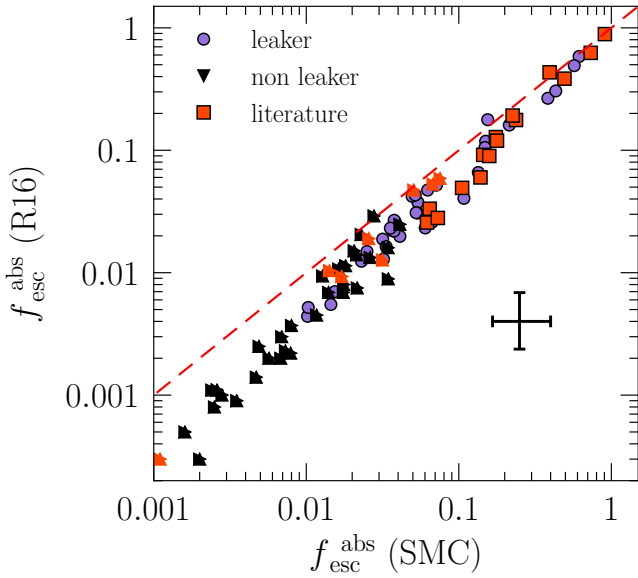
Monte-Carlo sampling of the spectra+continuum, see previous section). These methods are tested against the fiducial photon escape in Sect. 7.

## 4. UV attenuation and absolute photon escape of the LzLCS galaxies

From the direct output parameters of the spectral fits we obtain information on the light-weighted stellar age and metallicity of the stellar population, and the UV dust attenuation. The UV dust attenuation significantly impacts the derived absolute escape fractions, and so here we mostly focus on the relation between these two variables. The stellar population properties of the LzLCS sample will be discussed in a separate paper (Chisholm et al., in prep.).

The fitted  $E_{B-V}$  ranges between  $\sim 0.02$  and 0.45, with a mean value of  $\sim 0.15$  (0.07 and 0.24 as the 0.16 and 0.64 quantiles). Figure 3 shows the  $E_{B-V}$  derived from the UV-SED fits and a comparison to the ones determined from the Balmer decrement measurements using the optical SDSS spectra and the Cardelli et al. (1989) attenuation law (CCM). For the  $E_{B-V}$  from the UV fits, we examined two different attenuation laws: R16 and SMC, chosen to bracket the expected UV attenuation in SF galaxies<sup>3</sup> (Shivaei et al. 2020). A direct comparison between  $E_{B-V}$  values obtained from SMC and CCM is straightforward, because both laws have identical  $k_\lambda$  at optical wavelengths. In the case of R16, we used the Reddy et al. (2015) parametrization at  $\lambda > 1500$  Å

<sup>3</sup> Using the dust-attenuation law of Calzetti et al. (2000), our results do not change significantly with respect to the R16 law, because both attenuation laws have a similar wavelength dependency of  $k_\lambda$ , implying a similar absolute UV attenuation.



**Fig. 4.** 1:1 comparison between the absolute LyC photon escape fractions ( $f_{\text{esc}}^{\text{abs}}$ ) from the UV-SED fits using R16 and SMC dust-attenuation laws. LzLCS leakers and nonleakers are indicated by blue points and black arrows (upper limits), respectively. Red squares are the Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) sources. Median relative error bars are shown in the lower right corner.

converted to CCM law via the usual Balmer Decrement formulae<sup>4</sup>. When R16 is used, the overall  $E_{B-V}$  derived from the UV spectra is higher than that from the CCM law typically by a factor of  $\sim 1.2$ , while it is a factor of  $\sim 2.6$  lower when SMC is chosen.

A systematic shift between the nebular (obtained from emission line ratios) and the stellar attenuation (from the slope of the continuum) may be expected when a foreground dust-screen distribution is in play (see Calzetti et al. 2000). The stellar attenuation is typically a factor of 0.44 lower than the nebular attenuation (see Fig. 3). We only find this trend when the SMC law is used; R16 derived values agree better with the  $E_{B-V}$  from the Balmer decrements. Moreover, an SMC-like attenuation law has appeared more suitable for high- $z$  SF galaxies (Salim et al. 2018) and low-metallicity starburst (Shivvaei et al. 2020), requiring even steeper laws for leakers (Izotov et al. 2016b).

Distinguishing the attenuation law from the data used here is not possible, because the fit quality (reduced- $\chi^2$ ) is similar in both cases. Therefore, and as we mentioned in Sect. 3, we use the R16 attenuation law as a default in this study, motivated by the fact that this law is defined below  $1000 \text{ \AA}$  by a significant number of galaxies. This translates to a factor two lower attenuation at  $900 \text{ \AA}$  than estimates derived from extrapolations of dust attenuation curves such as the SMC, and will impact the determination of the escape of ionizing photons from SED fitting, as we explore below (see also Reddy et al. 2016b; Chisholm et al. 2019). Finally, to avoid a strong dependence on the assumed attenuation law, we refer to the UV attenuation defined as  $A_\lambda = k_\lambda E_{B-V}$  (Eq. (1)) instead of  $E_{B-V}$ . Still, differences in  $A_\lambda$  are found for the two laws at all wavelengths, being systematically higher for R16. This reveals that the fitted stellar parameters ( $X_i$ ) are slightly lower for the SMC law, meaning that both fits can match the observations similarly. In any case, both

<sup>4</sup>  $\frac{E_{B-V}^{\text{R16}}}{E_{B-V}^{\text{CCM}}} = \frac{(k_{\lambda_1} - k_{\lambda_2})^{\text{R16}}}{(k_{\lambda_1} - k_{\lambda_2})^{\text{CCM}}} = 0.93$ ,  $\lambda_1 = H\beta\lambda 4863$ ,  $\lambda_2 = H\gamma\lambda 4370$ .

approaches yield the same light-averaged ages and metallicities (further investigations in Chisholm et al., in prep.).

In Fig. 4 we show the derived  $f_{\text{esc}}^{\text{abs}}$  values (Eq. (2)), assuming both the R16 and the SMC law (see also Flury et al. 2022a). The derived LyC escape fractions vary over a wide range, with values up to 60% (85%) for the LzLCS (literature) sample. Typically, upper limits are below 0.01. Clearly, the choice of the attenuation law has some impact on the derived  $f_{\text{esc}}^{\text{abs}}$ : the relative difference between the two  $f_{\text{esc}}^{\text{abs}}$  versions ranges from 2% to 40%  $\times f_{\text{esc}}^{\text{abs}}$ , with systematically higher values obtained for the SMC law. This systematic shift is smaller for high values of the absolute escape fraction, which are found in galaxies with relatively low UV attenuation (Sect. 6.4). The shift to higher  $f_{\text{esc}}^{\text{abs}}$  for the SMC law is due to fact that this law is steeper than R16, which means that a smaller UV attenuation is needed to reproduce the overall, reddened observed spectral shape. Therefore, the intrinsic stellar flux (before attenuation by dust) is lower with the SMC law, and therefore  $f_{\text{esc}}^{\text{abs}}$  is higher.

Finally, our fitting algorithm is not devoid of degeneracies between stellar age, metallicity, and the  $E_{B-V}$  (see discussion in Gazagnes et al. 2020). The S/N of the original spectra impacts the ability of the fitting code at the time of solving this degeneracy, where only moderate (and high) S/N spectra will give a reliable estimation of the stellar population properties (Chisholm et al. 2019).

## 5. Empirical results and inferences from UV absorption lines

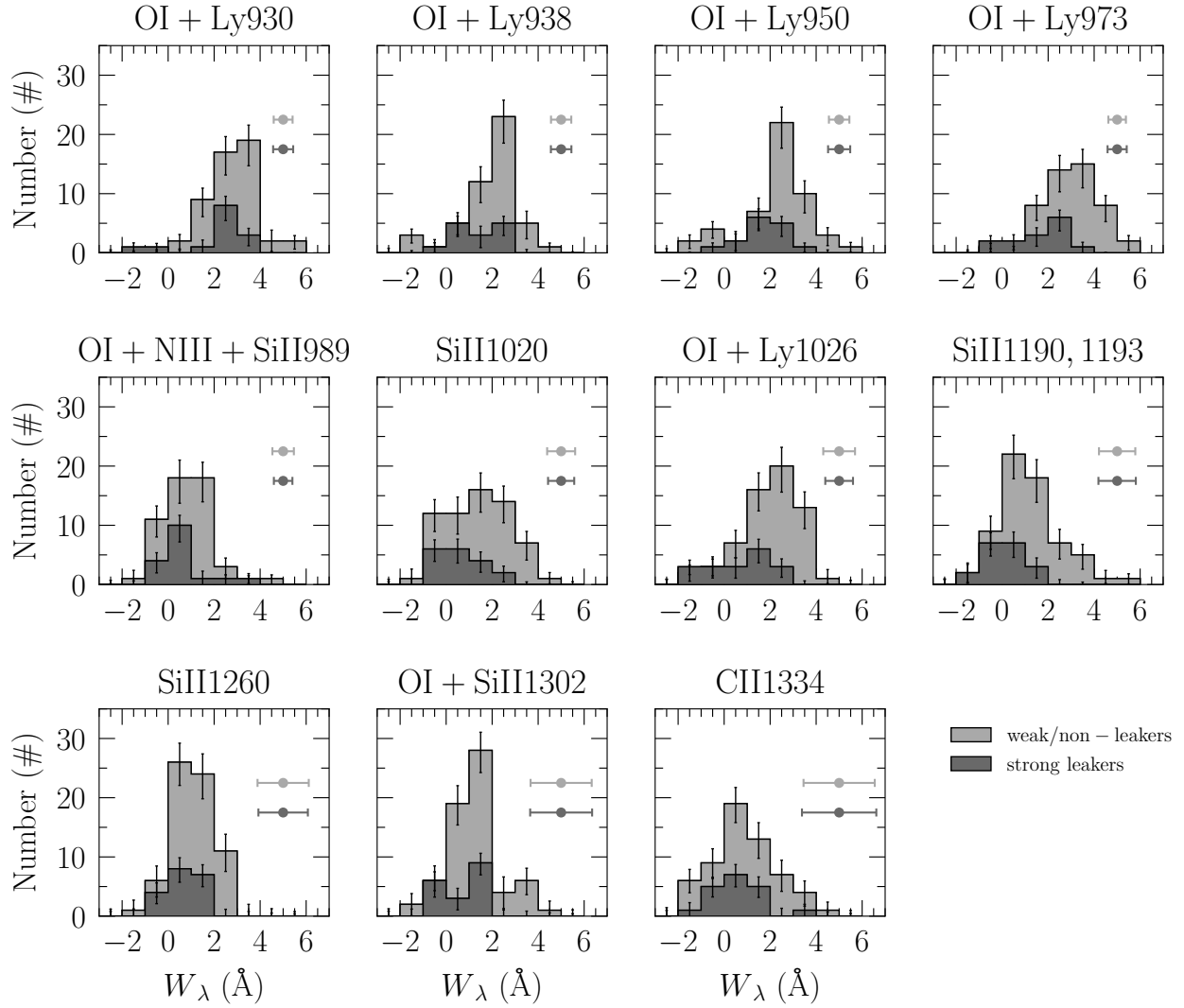
We now search for the empirical relations between the HI/LIS absorption lines, their equivalent widths (Sect. 5.1), residual fluxes, and some galaxy properties (Sect. 5.2). We also compare the observed trends with those in high- $z$  galaxies (Sect. 5.3).

### 5.1. LyC leakers have weak absorption lines

The equivalent width ( $W_\lambda$ ) distributions of the major hydrogen and metal absorption lines are shown in Fig. 5, where the strong leakers (dark gray) are separated from the weak and nonleaker distributions (light gray). The error bars on the histograms come from a proper estimation of the Poissonian confidence intervals for every  $W_\lambda$  bin, taking into account the uncertainty on single measurements.

Overall, the equivalent widths reach up to 4–6  $\text{\AA}$  for the strongest lines. The number of detections (defined here as value/error  $> 2$ ) out of the total of 89 galaxies in the sample are 62, 51, 54, 56, and 50 for the Lyman series, and 32, 35, 23, 20, 13, and 9 for the LIS lines, respectively. The main cause for nondetection is the contamination by geo-coronal emission in the Lyman series range and the decrease in S/N towards longer wavelengths. The presence of negative equivalent widths can be due to different reasons. For some HI lines, such as H $1\lambda 1026$ , the continuum near the line is complex to model because of contributions from a stellar OVI P Cygni profile as well as numerous adjacent ISM absorption lines (in particular OVI, CII, OI). Slight underestimations of the continuum level would lead to negative equivalent widths. In the case of the LIS lines, the low S/N and overlap of some fluorescence transitions within the integration window would also produce some negative  $W_\lambda$  values.

The equivalent width distributions of leakers peaks at lower values than the nonleaking distributions in general, suggesting that leakers have weaker HI and LIS absorption lines than nonleakers. Table 1 shows the median of the  $W_\lambda$



**Fig. 5.** Equivalent width distributions ( $W_\lambda$ , in  $\text{\AA}$ ) for the measured absorption lines. Dark-gray histograms correspond to the strongest leakers (20) in the LzLCS plus published joint samples, while light-gray histograms represent the weak leakers (30) and nondetections in the LyC window (39), out of the total of 89 galaxies. Strong leakers are defined as galaxies with significant detection in the LyC and  $f_{\text{esc}}^{\text{abs}} \geq 0.05$ . Weak leakers also have significant LyC detection but  $f_{\text{esc}}^{\text{abs}} < 0.05$ . Nonleakers are not detected in the LyC. Median equivalent-width errors are indicated in all the panels and for each source type.

distributions for the entire sample (89 galaxies), strong leakers (20), and weak+nonleakers only (69). For example, for H I  $\lambda$ 1026 (Si II  $\lambda$ 1190,1193), the median equivalent width value for leakers is  $0.58 \text{ \AA}$  ( $0.17 \text{ \AA}$ ), while nonleakers show median equivalent widths of  $1.97 \text{ \AA}$  ( $1.21 \text{ \AA}$ ), respectively. To test whether the two distributions differ or not, one can perform the Anderson-Darling test. For the previous lines, the null hypothesis that assumes leaker and nonleaker populations to come from the same parent distribution can be rejected at 97.5% level, with a significance of  $p_{\text{val}} \leq 10^{-2}$ . Based on smaller samples Chisholm et al. (2017) already noted that LyC emitters show weaker H I and LIS lines (see also Mauerhofer et al. 2021, in simulations).

## 5.2. H I and LIS absorption line strengths and UV attenuation are driven by the gas covering fraction

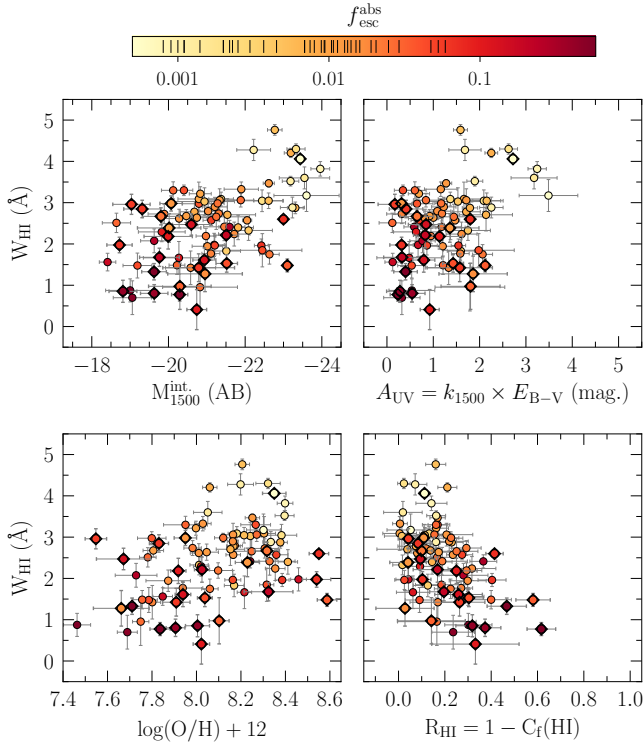
The observed differences in the H I and LIS lines between LyC emitters and nonemitters suggests a connection between the mechanisms that favor the LyC leakage and those that lead to

**Table 1.** Median line equivalent widths (in  $\text{\AA}$ ) for the whole working sample (all, 89 galaxies), strong leakers only (20 galaxies), and weak+nonleakers (69 galaxies).

Ion + $\lambda_{rf}$ ( $\text{\AA}$ )	$W_\lambda$ ( $\text{\AA}$ ) all (89)	$W_\lambda$ ( $\text{\AA}$ ) strong leakers (20)	$W_\lambda$ ( $\text{\AA}$ ) weak+nonleakers (69)
O I+Ly930	2.75	2.05	2.94
O I+Ly938	1.71	1.41	1.80
O I+Ly950	2.09	1.62	2.24
O I+Ly973	2.78	1.58	3.15
O I+N III+Si II989	0.77	0.60	0.83
Si II1020	1.13	0.57	1.32
O I+Ly1026	1.64	0.58	1.97
Si II1190,1193	0.96	0.17	1.21
Si II1260	0.96	0.56	1.08
O I+Si II1302	1.08	0.64	1.21
C II1334	0.79	0.52	0.88

**Notes.** Leakers have lower median equivalent widths than nonleaking galaxies in most cases.





**Fig. 6.** Empirical trends between the H I equivalent width ( $W_{\text{HI}}$ ) and different observational properties (from *top* to *bottom* and *left* to *right*): intrinsic absolute magnitude at 1500 Å ( $M_{1500}^{\text{int}}$ ), UV attenuation ( $A_{\text{UV}}$ ) at 1500 Å, measured gas-phase metallicity ( $12 + \log(\text{O}/\text{H})$ ) and the average H I residual flux ( $R(\text{H I})$ ). Points are color-coded by their absolute LyC photon escape fraction,  $f_{\text{esc}}^{\text{abs}}$  (and upper limits values on  $f_{\text{esc}}^{\text{abs}}$  are indicated inside the color bar for reference). The *Izotov et al. (2016a,b, 2018a,b, 2021)* and *Wang et al. (2019)* sources are displayed with diamonds and also considered in the correlations (see text).

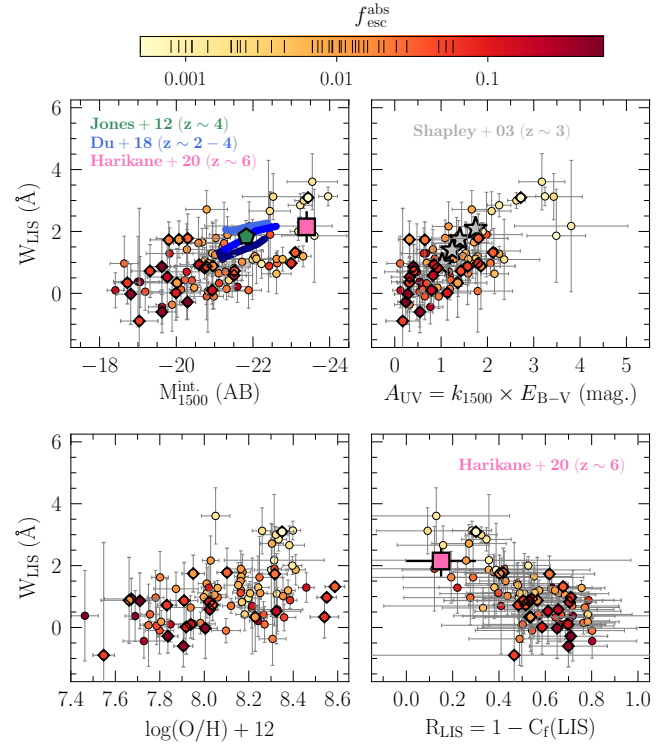
changes in the absorption lines profiles. Therefore, we explore several empirical correlations between H I and LIS equivalent widths and other galaxy properties, such as the intrinsic absolute magnitude at 1500 Å ( $M_{1500}^{\text{int}}$ ), the UV attenuation ( $A_{\text{UV}}$ ), the measured gas-phase metallicity ( $12 + \log(\text{O}/\text{H})$ ), and the average H I and LIS residual flux ( $R(\text{H I})$ ,  $R(\text{LIS})$ ).  $M_{1500}^{\text{int}}$  is produced by taking the mean of the dust-free S99 modeled flux over a 100 Å window around 1500 Å, and converting this into an AB absolute magnitude given the redshift and the distance modulus.  $A_{\text{UV}}$  is taken from the fits,  $A_{\text{UV}} = k_{\text{UV}} \times E_{\text{B-V}}$  – at 1500 Å –, and finally residual fluxes are computed using Eq. (4) and subsequently Eq. (7) for H I and similarly for the LIS lines (Si II  $\lambda\lambda$ 1190,1193, Si II  $\lambda$ 1260, O I  $\lambda$ 1302, C II  $\lambda$ 1334).

For the equivalent widths, we take the inverse-variance-weighted average of the single H I and LIS line widths ( $W_{\text{H I}}$ ,  $W_{\text{LIS}}$ ):

$$W_K \pm \Delta W_K = \frac{\sum_{i=1}^N W_{\lambda i} \times \frac{1}{\sigma_{W_i}^2}}{\sum_{i=1}^N \frac{1}{\sigma_{W_i}^2}} \pm \sqrt{\frac{1}{\sum_{i=1}^N \sigma_{W_i}^2}},$$

$K \equiv \text{H I}$   
 $i \equiv \text{H I } \lambda 930, \text{H I } \lambda 938, \text{H I } \lambda 950, \text{H I } \lambda 973, \text{H I } \lambda 1026$   
 $K \equiv \text{LIS}$   
 $i \equiv \text{Si II } \lambda\lambda 1190, 1193, \text{Si II } \lambda 1260, \text{O I } \lambda 1302, \text{C II } \lambda 1334,$

where  $N$  is the number of H I or LIS lines that can be measured on a single spectra and  $\sigma_{W_i}$  represents the error on the



**Fig. 7.** Empirical trends between the LIS equivalent width ( $W_{\text{LIS}}$ ) and different observational properties:  $M_{1500}^{\text{int}}$ ,  $A_{\text{UV}}$ ,  $12 + \log(\text{O}/\text{H})$  and  $R(\text{LIS})$ , with sources displayed and color-coded as in Fig. 6. Additionally, results from *Shapley et al. (2003,  $z \sim 3$ )* and *Harikane et al. (2020,  $z \sim 6$ )* composites are overplotted. The light-blue, blue, and dark-blue lines correspond to the different redshifts between  $z = 2 - 4$  from *Du et al. (2018)*.

individual equivalent widths measurements (Monte-Carlo sampling). At this point it becomes necessary to take into account the following aspects. Some lines, such as H I  $\lambda$ 973 and Si II  $\lambda$ 989, are contaminated by geo-coronal emission in many cases (those values are removed from the average). We exclude Si II  $\lambda$ 1020 from the average because this line has a very low oscillator strength and is usually very weak and likely optically thin. We also exclude Si II  $\lambda$ 989 because of the geo-coronal contamination and the blending with the high-ionization-state N III  $\lambda$ 989 line.

Averaging the H I and LIS lines of the same ion is reasonable owing to the fact that most of those lines have similar oscillator strengths or are likely saturated, and so their equivalent widths are similar. Even LIS transitions for very different ions (e.g., Si II versus O I) will only differ by a factor of two in SF galaxies (e.g., *Chisholm et al. 2016*). Averaging LIS line measurements has also frequently been reported for high-redshift galaxies (e.g., *Shapley et al. 2003; Du et al. 2018, 2021; Trainor et al. 2019*) to improve the S/N of single line measurements. As mentioned above, in our case the average is strictly necessary for the LIS lines redder than 1200 Å rest-frame (or below 1600 Å in the observed frame), where the S/N is usually very low due to the decrease in the COS instrument sensitivity.

In Figs. 6 and 7, the H I and LIS equivalent widths are plotted against the four quantities described above:  $M_{1500}^{\text{int}}$ ,  $A_{\text{UV}}$ ,  $12 + \log(\text{O}/\text{H})$ ,  $R(\text{H I})$ , and  $R(\text{LIS})$ . LzLCS sources are represented by circles while the published sample is shown using diamonds. The sources are color coded by our  $f_{\text{esc}}^{\text{abs}}$  determined from the UV spectral fits. Overall, strong absorption lines are only found at high  $A_{\text{UV}}$ , high metallicity, and in UV-luminous sources

(where  $M_{1500}^{\text{int}}$  is dust corrected). In other words,  $W_{\text{HI}}$  and  $W_{\text{LIS}}$  both correlate with the UV attenuation and metallicity, but anticorrelate with the intrinsic UV absolute magnitude and residual fluxes. High  $f_{\text{esc}}^{\text{abs}}$  sources are located at the lowest HI/LIS equivalent-width regime (see also Sect. 6.2). The variations in the absorption line properties with the LyC escape fraction are discussed further below.

In order to quantify the above-mentioned trends, we computed the Kendall correlation coefficient for every pair of variables described above ( $\tau$ , see Table 2). For a sample of 89 objects, we consider correlations to be significant if  $p_{\text{val}} \lesssim 2.275 \times 10^{-2}$  (i.e.,  $2\sigma$ ) and strong if  $|\tau| \gtrsim 0.2$ . For  $W_{\text{HI}}$ , the strongest and most significant correlation is found to be with the UV intrinsic absolute magnitude  $M_{1500}^{\text{int}}$ . On the other hand,  $W_{\text{LIS}}$  shows the strongest trend with the LIS residual flux.

Generally, three different mechanisms can change the HI/LIS equivalent widths – if we assume the lines are formed in single gas clouds, with a single velocity and gas column density: (1) the HI/metal column densities and metallicity, (2) the Doppler broadening (velocity dispersion of the cloud), and (3) the HI/LIS covering fraction. Doppler broadening (2) is challenging to address, given our spectral resolution. This said, and based on visual inspection, we do not see any evidence for the highest equivalent widths being among the widest lines, but they are among the deepest. The effect of the covering fraction (3) is clearly visible in the bottom right panels of Figs. 6 and 7, where the residual flux is linked to the covering fraction – for a uniform dust geometry of the picket-fence model – and so the equivalent width is directly related to the line depth. It remains to be seen whether or not other physical effects can also play a role.

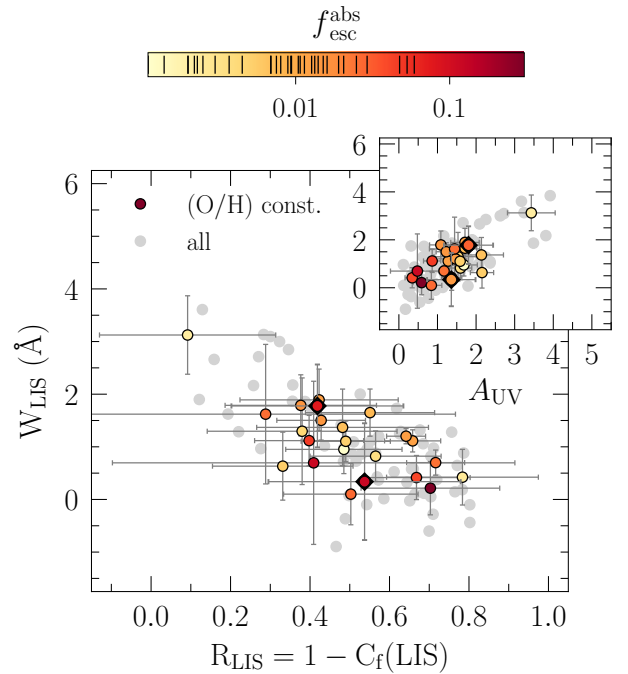
For example, in a study of nearby SF galaxies, Chisholm et al. (2017) suggested that the decrease in the  $W_{\text{LIS}}$  is set by a combination of small HI column densities and low metallicity. The relation between the HI column density ( $N_{\text{HI}}$ ) and the column density of metals ( $N_{\text{Z}}$ ) scales with metallicity ( $Z$ ) as  $N_{\text{Z}} \sim Z \times N_{\text{HI}}$ . A decrease in metallicity ( $Z$ ) would produce a weaker  $W_{\text{LIS}}$ , leading to a natural  $W_{\text{LIS}} - A_{\text{UV}}$  dependency. However, a change in metallicity cannot alter the HI line widths and thus explain the observed  $W_{\text{HI}} - A_{\text{UV}}$  trend. Furthermore, as the HI lines must be saturated – based on the ratio of equivalent widths of different HI and Si II transitions (Appendix B) – column density must be secondary. The observations of the Lyman series lines and the fact that they behave in a very similar fashion to those of the LIS lines therefore exclude option (1) to explain the trends shown in Figs. 6 and 7.

The effect of metallicity is examined in more detail by restricting our analysis to the sources at a constant gas phase metallicity of  $12 + \log(\text{O}/\text{H}) = 8.2 \pm 0.1$  (25 galaxies). Doing so, the trends seen in Figs. 6 and 7 remain the same. As an example, in Fig. 8 we show the resulting  $W_{\text{LIS}} - \text{R}(\text{LIS})$  behavior when the sample is limited in metallicity (colored points) against the whole LzLCS+literature sample (gray points in the background). Quantitatively, a Kendall test performed over the  $W_{\text{LIS}} - A_{\text{UV}}$  relation for example (see inset in Fig. 8), still reveals a similar strong and significant correlation when the sample is restricted in metallicity ( $\tau \sim 0.25$ ,  $p_{\text{val}} \sim 10^{-3}$ ). On the other hand, when restricting to constant  $\text{R}(\text{LIS})$ , the statistical connection between  $W_{\text{LIS}}$  and  $A_{\text{UV}}$  disappears, and similarly for the other variables. This excludes the metallicity (1) as the parameter driving the LIS correlations. We therefore conclude that variations of the geometric covering fraction (3) are the most likely parameter driving the observed trends both for the HI and LIS lines, and also for the UV attenuation (with a possible cloud-to-cloud velocity dispersion, and the different column densities in front of the

**Table 2.** Kendall ( $\tau$ ) correlation coefficients (see Akritas & Siebert 1996) and  $p$ -values for HI and LIS equivalent widths ( $W_{\text{HI}}$ ,  $W_{\text{LIS}}$ ) versus diverse galaxy properties.

	$W_{\text{HI}}$		$W_{\text{LIS}}$	
	$\tau$	$p_{\text{val}}$	$\tau$	$p_{\text{val}}$
$M_{1500}^{\text{int}}$	-0.342	$4.123 \times 10^{-6}$	-0.383	$1.493 \times 10^{-7}$
$A_{\text{UV}}$	+0.220	$3.009 \times 10^{-3}$	+0.447	$9.011 \times 10^{-10}$
$12 + \log(\text{O}/\text{H})$	+0.274	$2.284 \times 10^{-4}$	+0.297	$4.641 \times 10^{-5}$
$\text{R}(\text{HI})$	-0.331	$8.576 \times 10^{-6}$	–	–
$\text{R}(\text{LIS})$	–	–	-0.462	$2.431 \times 10^{-10}$
$-W_{\text{Ly}\alpha}$	-0.535	$5.819 \times 10^{-13}$	-0.302	$4.745 \times 10^{-5}$

**Notes.** Intrinsic (dust-free) absolute magnitude at 1500 Å ( $M_{1500}^{\text{int}}$ ), UV attenuation ( $A_{\text{UV}}$ ), measured gas-phase metallicity ( $12 + \log(\text{O}/\text{H})$ ), the average HI and LIS residual flux ( $\text{R}(\text{HI})$ ,  $\text{R}(\text{LIS})$ ) and Ly $\alpha$  equivalent widths ( $W_{\text{Ly}\alpha}$ ).



**Fig. 8.** LIS equivalent widths ( $W_{\text{LIS}}$ ) versus residual fluxes ( $\text{R}(\text{LIS})$ , main panel) and UV attenuation ( $A_{\text{UV}}$ , inset), when the working sample is restricted to a constant metallicity of  $12 + \log(\text{O}/\text{H}) = 8.2 \pm 0.1$  (color-coded as in Fig. 6). The global trend including all the sources is shown with gray background symbols.

stars contributing to the scatter). These findings are similar to the conclusions reached in (Du et al. 2021) from the analysis of UV absorption lines and attenuation in LBGs at  $z \sim 2-3$ . Du et al. (2021) find that metallicity only affects the LIS equivalent widths when the  $W_{\text{Ly}\alpha}$  is fixed at a constant value or, in other words, when the HI covering fraction is also roughly constant.

### 5.3. Comparison with high-redshift galaxies

Few studies have explicitly shown high-order Lyman series trends with other observables (Henry et al. 2015; Reddy et al. 2016b; Gazagnes et al. 2018, 2020; Steidel et al. 2018; McKinney et al. 2019), because these lines are in the extreme FUV and hard to detect, often only accessible in the nearby Universe (e.g., see Lecavelier des Etangs et al. 2004;

Lebouteiller et al. 2006; Grimes et al. 2009, using the FUSE satellite). In general, our HI and LIS equivalent width lines correlate similarly with the other galaxy properties, supporting a direct link between neutral gas and metallic gas distributions (see next Sect. 6.2), and the claim that the neutral gas content of galaxies is well traced by LIS species (Shapley et al. 2003). The similarity in the behavior of HI and LIS lines also suggests that metallicity cannot be the major contributor to these trends, whose effect would no longer be visible if the lines were saturated.

In Fig. 7, we compare the observed  $W_{\text{LIS}}$  and UV magnitudes with results from Jones et al. (2012) and Harikane et al. (2020), who used LBG composite spectra at  $z \sim 4$  and  $z \sim 6$ , respectively. Here, Jones et al. (2012) and Harikane et al. (2020) reported values are not dust-corrected, i.e., they are observed AB absolute magnitudes. As seen in Fig. 7 (top left), the agreement is fair with the high- $z$  literature results, suggesting that the ISM structure and properties of the low- $z$  galaxies might be similar to those at high redshift.

Continuing, similar empirical relations between  $W_{\text{LIS}}$  and  $A_{\text{UV}}$  (or  $E_{\text{B-V}}$ ) were found by Shapley et al. (2003), Du et al. (2018) and Pahl et al. (2020), although the relations are generally shifted to lower reddening values. In the same Fig. 7, we also show the results from Shapley et al. (2003)<sup>5</sup>. The continuum sample selection in Shapley et al. (2003) against our continuum plus emission line selected sample could be a possible reason for higher  $W_{\lambda}$  at fixed  $A_{\text{UV}}$ . Moreover, the use of a broad-band SED fitting approach to infer the reddening or the spectral stacking could lead to the current differences and, less importantly, the use of different attenuation laws. Leitherer et al. (2011b) and Grimes et al. (2009) show similar trends with the  $12 + \log(\text{O}/\text{H})$ , but shifted to higher metallicities (typical of local starburst galaxies). The trend between the  $W_{\text{LIS}}$  and LIS residual fluxes (bottom right panel) is compatible with the results of Harikane et al. (2020) who use a  $z \sim 6$  galaxy composite. Finally, another possible reason for the small systematic differences in the observed equivalent widths at low- and high- $z$  could be the different spectral resolution and aperture effects, which would systematically lead to stronger observed lines for the low- $z$  galaxies.

## 6. Relations between the UV absorption lines, Ly $\alpha$ , and LyC escape

Here, we present the empirical relations between the UV absorption line, the Ly $\alpha$  emission line, and the Lyman continuum escape (Sect. 6.1). We then investigate the connection between HI and LIS covering fractions (Sect. 6.2). Finally, we study the effect of the neutral and low-ionized gas covering fraction on the leakage of ionizing radiation (Sect. 6.3) and the influence of dust (Sect. 6.4).

### 6.1. Empirical correlations between the HI and LIS absorption lines, Ly $\alpha$ and LyC

Although Ly $\alpha$  and LyC photons interact with the same neutral gas, gas kinematics can influence the Ly $\alpha$  output, while it has almost no effect (directly) on LyC emission. The interplay between the Ly $\alpha$  emission and the LyC leakage has been widely studied (Rivera-Thorsen et al. 2015, 2017; Verhamme et al. 2015, 2017; Dijkstra et al. 2016; Marchi et al. 2018; Izotov et al. 2020, 2021; Gazagnes et al. 2020). Hence, it is a natural

<sup>5</sup> For a proper comparison we adopted the appropriate  $k_{\text{UV}}$  values, depending on the attenuation law used in each study.

step to also study the behavior of the HI/LIS lines with the Ly $\alpha$  line. We use the Ly $\alpha$  equivalent widths ( $W_{\text{Ly}\alpha}$ ) measured by the LzLCS team from the COS spectra (Flury et al. 2022a).

In Fig. 9, the  $W_{\text{Ly}\alpha}$  is compared to the different HI and LIS line properties. Very strong and significant correlations (first column) are found between  $W_{\text{Ly}\alpha}$  and  $W_{\text{HI}}$ , and with  $W_{\text{LIS}}$  (see Table 2).  $W_{\text{Ly}\alpha}$  and  $R(\text{HI})$  are also moderately correlated (however they have larger uncertainties). In summary, larger Ly $\alpha$  equivalent widths are found in galaxies where the weaker HI and LIS absorption lines are, that is,  $W_{\text{Ly}\alpha}$  is anti-correlated with  $W_{\text{HI}}$  and  $W_{\text{LIS}}$  but correlates with  $R(\text{HI})$  and  $R(\text{LIS})$ . The color-coding demonstrates the correlation between  $W_{\text{Ly}\alpha}$  and  $f_{\text{esc}}^{\text{abs}}$ . Galaxies with higher  $f_{\text{esc}}^{\text{abs}}$  are statistically the strongest Ly $\alpha$  emitters (LAEs), that is, they show large equivalent widths (Flury et al. 2022b). Additionally, larger scatter at the high- $W_{\text{Ly}\alpha}$  end at fixed  $W_{\text{HI}}$  and  $W_{\text{LIS}}$  may result from boosted (or at least variations in) intrinsic Ly $\alpha$  production.

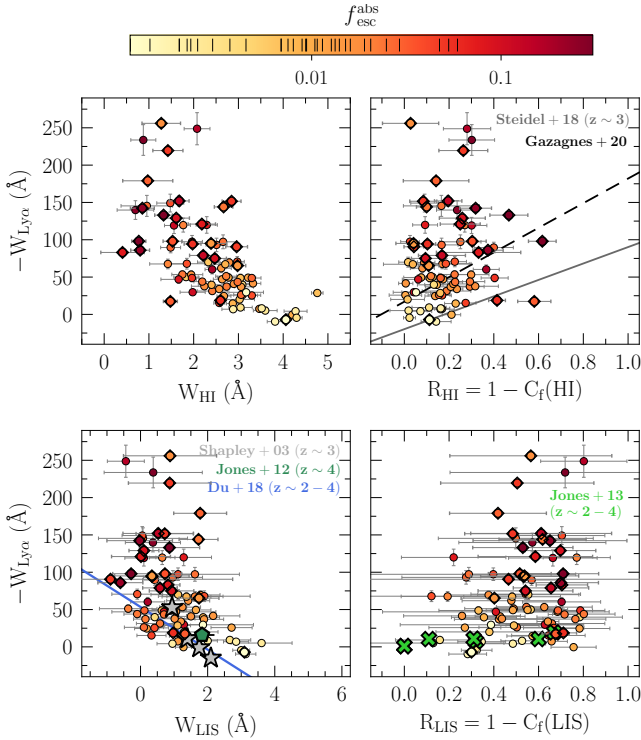
Steidel et al. (2018) and Gazagnes et al. (2020) studied the connection between the Ly $\alpha$  equivalent width and the HI residual flux. In Fig. 9 (top right) we show how their fitted linear relations compare to our data (gray and black lines). Both correlations are less steep than ours, but a more prominent difference exists when comparing with Steidel et al. (2018). As discussed in Gazagnes et al. (2020), sample selection is likely the key to understand these differences. Steidel et al. (2018) photometrically selected LBGs with rest-frame UV bands, meaning that they are generally brighter, and have higher masses and/or higher SFRs than the low-redshift galaxies studied by Gazagnes et al. (2020) and here. Another possible explanation is the scattering of the blue side of the Ly $\alpha$  line by the IGM, which would produce lower  $W_{\text{Ly}\alpha}$  in the Steidel et al. (2018) sample. Additionally, flux aperture losses would also lead to a decrease in the  $W_{\text{Ly}\alpha}$  of the former sample with respect to the LzLCS.

Finally, important to note is the bias introduced by the redshifts of the two samples: from nearby  $z \sim 0.3$  to distant  $z \sim 3$  galaxies, with intrinsic different properties. After all, the current sample may include sources with larger  $W_{\text{Ly}\alpha}$  than the bulk of the LBGs at  $z \sim 3$ , whereas in Steidel et al. (2018), sources were selected to have brighter continua.

The Ly $\alpha$  and LIS connection has been widely studied by stacking LBGs at  $z \sim 3-5$  (Shapley et al. 2003; Jones et al. 2012; Du et al. 2018; Pahl et al. 2020) and also with individual detections at slightly lower redshifts (Trainor et al. 2019; Du et al. 2021, at  $z \sim 2-3$ ). For comparison, we include the linear regression between the Ly $\alpha$  and LIS equivalent widths found by Du et al. (2018) (Fig. 9, bottom left), and data from stacked spectra in Shapley et al. (2003) and Jones et al. (2012). High- $z$  LBG measurements seem to be more consistent with low- $f_{\text{esc}}^{\text{abs}}$  or nonleaking galaxies: strong leakers occupy a different portion of this plot, with the highest  $W_{\text{Ly}\alpha}$ . Similar trends between the Ly $\alpha$  and metal equivalent widths were found in Jaskot et al. (2019) for individual Green Pea-like galaxies.

The  $W_{\text{Ly}\alpha}$  correlation with the metal residual flux is more difficult to interpret – mainly because of the larger error bars on  $R(\text{LIS})$  – but it is still moderately significant. This result may suggest that the trend between Ly $\alpha$  and LIS absorption equivalent width is predominantly due to variations in the covering fraction of neutral hydrogen (see below), which modulates the Ly $\alpha$  profile and is directly connected with the LIS covering fraction (see Sect. 6.2). The analysis of Gazagnes et al. (2020) indicates that the gas kinematics could also explain part of the observed scatter, as well as the distribution of different gas column densities in front of the stars. For comparison, we include the  $W_{\text{Ly}\alpha}$ – $R(\text{LIS})$  correlation from the Jones et al. (2013)





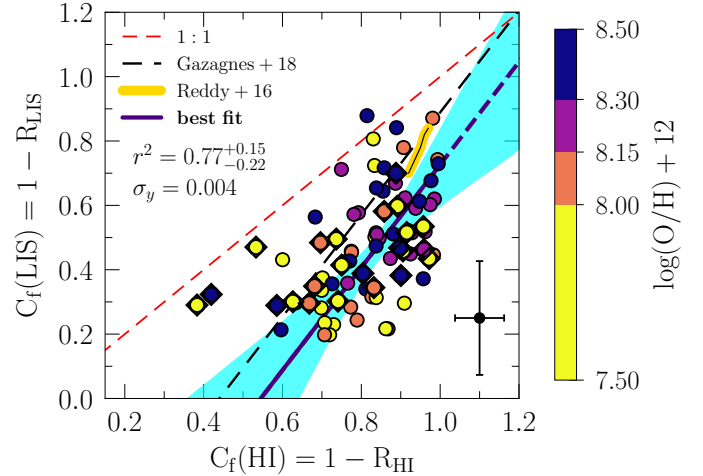
**Fig. 9.** Relation between the Ly $\alpha$  ( $W_{\text{Ly}\alpha}$ ), HI (*top*), and LIS line (*bottom*) equivalent widths and residual fluxes, respectively. Sources are displayed and color-coded as in Fig. 6. Solid and dashed lines represent the linear fits to the relations found in Steidel et al. (2018), Du et al. (2018), and Gazagnes et al. (2020) for comparison. Gray stars, the green polygon, and green crosses correspond to Shapley et al. (2003), Jones et al. (2012, 2013) results.

compilation of  $z \sim 2-4$  galaxies, although the latter sample lacks strong LAEs, as opposed to ours, which explains the differences.

As the LIS absorption lines are a probe of neutral gas, there is a direct link between the Ly $\alpha$  profile and LIS lines (Shapley et al. 2003). From the apparent nonevolution of the  $W_{\text{Ly}\alpha}-W_{\text{LIS}}$  relation with redshift, Du et al. (2018) and Pahl et al. (2020) suggest that Ly $\alpha$  and the LIS absorption features are fundamentally related, indicating that the gas giving rise to LIS absorption modulates the radiative transfer of Ly $\alpha$  photons. We now know that, rather than metallicity, variations of the line covering fractions are the main cause for changes in  $W_{\lambda}$ . In order to examine the role of the gas covering fraction in  $W_{\text{Ly}\alpha}$  from the low-redshift sample, we again restrict our sample to the sources at a constant gas phase metallicity of  $12 + \log(\text{O}/\text{H}) = 8.2 \pm 0.1$  (25 galaxies). With this restriction, we find the same trends as in Fig. 9, confirming that the observed  $W_{\text{Ly}\alpha}-W_{\text{LIS}}$  relation is not primarily due to metallicity variations, but it is ultimately driven by variations in the HI covering fraction (with the different column densities in front of the ionizing sources contributing to the scatter). Given the locus of the strongest leakers in the  $W_{\text{Ly}\alpha}-W_{\text{LIS}}$  plane, having large Ly $\alpha$  equivalent widths but weak LIS lines, this diagram can be used to select potential LyC emitters in high-redshift surveys.

## 6.2. HI and metals are not evenly distributed, but trace each other

Although with different ionization potentials, it is well-known that HI and LIS absorption lines in SF galaxies generally show similar kinematics (line profiles, central velocities, etc.; see



**Fig. 10.** Empirical correlation between the LIS and HI covering fractions. Red, yellow, black, and dark blue lines correspond to the 1:1, Reddy et al. (2016b), and Gazagnes et al. (2018) relations and our Bayesian linear fit (LINMIX), respectively. The points are color-coded according to the gas-phase metallicity of each source,  $12 + \log(\text{O}/\text{H})$ , and typical error bars are shown in the bottom-right corner. The blue shaded area represents the  $2\sigma$  confidence interval including the intrinsic scatter of the fit. The linear correlation coefficient ( $r^2$ ) and intrinsic scatter over the fit ( $\sigma_y$ ) are given. The Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) sources are displayed in diamonds, and also included in the fit.

Dessauges-Zavadsky et al. 2010; Trainor et al. 2015; Chisholm et al. 2016; Reddy et al. 2016b). This indicates that the LIS and neutral gas are comoving and trace similar gas content. Here, we proceed to study the relation between the observed neutral and metallic gas geometrical distributions through the covering fraction.

Covering fractions for every line were derived from the corresponding line residual fluxes using Eq. (4), and applying Eq. (7) for HI and similarly for the LIS lines. This assumes a uniform dust-screen geometry in the picket-fence model, a single velocity component, and saturated lines. In Fig. 10, the LIS covering  $C_f(\text{LIS})$  is compared to the neutral gas covering fraction obtained from the set of Lyman series lines  $C_f(\text{HI})$ . Generally, the LIS covering fraction is lower, typically by a factor of two, compared to that of the HI lines. The median covering fraction is 0.82 for HI and 0.46 for the LIS. A linear relation between the two is clear from Fig. 10, such that the LIS covering fraction increases with increasing HI covering fraction.

To quantify this, we fit a linear regression model to the data using the Bayesian fitting code LINMIX<sup>6</sup> (Kelly 2007), which accounts for errors on both variables, also including censored data (upper or lower limits). Further details are provided in Appendix D. We obtain

$$C_f(\text{HI}) = (0.63 \pm 0.19) \times C_f(\text{LIS}) + (0.54 \pm 0.09). \quad (9)$$

The resulting fit is quite significant, with a linear correlation coefficient of  $r^2 = 0.77 (+0.15, -0.22)$ , and a negligible estimated intrinsic scatter of  $\sigma_y = 0.004$ .

Our result is compared to the Gazagnes et al. (2018) fit, where a similar approach was followed but for a small subset of the LyC emitters (dashed black line). Both linear relations have similar slopes, but slightly different normalizations. Possible differences could be due to the use of the mean of different LIS lines – in

<sup>6</sup> A PYTHON version of the LINMIX package can be found in <https://linmix.readthedocs.io/en/latest/index.html>

our case – instead of single Si II  $\lambda$ 1260 line in the previous work, or due to the effect of the resolution, which was corrected for by Gazagnes et al. (2018). The measured residual fluxes can be affected by the spectral resolution, and therefore the actual linear relation also depends on the resolution. However, as discussed in Sect. 3.2, the resolution effect is subdominant in this sample. Our results are also compatible with those of Reddy et al. (2016b), although spanning a wider range of  $C_f$  values. As sketched in the previous paragraph, the strong leakers are located not only with the lowest HI but also the lowest LIS covering fractions, having 0.77 (HI) and 0.41 (LIS) as median values.

Multiple explanations can lead to these covering fraction differences. As explained in Reddy et al. (2016b), gas kinematics can influence the measured residual fluxes of metal lines. If most of the metals reside in narrow and unresolved absorption regions, the limited spectral resolution could underpredict the gas covering fraction estimated from the residual intensity of the lines. On the other hand, if the metal-enriched gas has a large velocity gradient in the line of sight, the derived covering fractions may be lower limits (see discussion in Mauerhofer et al. 2021). Nevertheless, the similar blueshifted velocity and spread of HI and LIS lines suggests that both phases are not separated, but coupled in the ISM, although once again, the limited spectral resolution does not allow us to confirm this definitely. Orientation effects for lines created in the circumgalactic medium (CGM) instead, where HI and LIS are not particularly well-coupled, could also contribute to the scatter in the relation. Additionally, LIS lines could also be affected by scattering (Mauerhofer et al. 2021, in simulations) and fluorescence of LIS emission in-filling (Scarlata & Panagia 2015), although this effect is very difficult to estimate.

Gazagnes et al. (2018) suggested that a different distribution of the metals, which would also depend on metallicity, could explain the current relation. In an ISM where HI is not fully mixed with the metals, HI would probe high and low-metallicity regions, whereas metals would only probe the higher metallicity areas. Therefore, the HI covering fraction would be higher than that of the metals. To test this scenario, we color-coded the points in Fig. 10 according to the 0.25, 0.5, and 0.75 gas-phase metallicity ( $12 + \log(\text{O}/\text{H})$ ) quantiles. Overall, sources with low HI and LIS covering fractions have lower metallicities, although the scatter and error bars blur this behavior. If we divide the sample into high- ( $12 + \log(\text{O}/\text{H}) > 8.1$ ) and low-metallicity ( $12 + \log(\text{O}/\text{H}) \leq 8.1$ ) subsamples, we find average covering values of  $C_f(\text{HI}) = 0.85$  (0.78) and  $C_f(\text{LIS}) = 0.52$  (0.43) for high (low) metallicities, confirming the previous hypothesis. Nevertheless, high- and low-metallicity subsamples approximately follow the same linear relation predicted by Eq. (9) (similar slopes and intercepts). This excludes a significant dependence of the previous  $C_f(\text{LIS}) - C_f(\text{HI})$  relation on metallicity, but as the measured metallicities are integrated over the whole galaxy, we cannot exclude the inhomogeneous mixing scenario.

Finally, there is still the possibility that the lower end of this  $C_f(\text{HI}) - C_f(\text{LIS})$  correlation is mainly driven by unsaturated lines. There is no simple way to address this question rather than the analysis already made in Appendix B. In there, we conclude that the number of unsaturated lines sources must be minimal, but without certainty for the LIS lines. For example, for galaxies with  $12 + \log(\text{O}/\text{H}) = 7.8$  (among our lower metallicities) and assuming that Si II roughly traces the HI fraction, one has  $\log(\text{Si}/\text{H}) \sim -5.7$ , adopting the nearly constant<sup>7</sup> observational relation of  $\log(\text{Si}/\text{O}) \sim -1.5$  (see Izotov & Thuan 1999).

<sup>7</sup> A constant Si/O ratio is expected because both Si and O are  $\alpha$ -elements sharing similar formation mechanisms.

According to a curve-of-growth analysis (Draine 2011), the LIS lines are likely saturated at  $\log(N_{\text{Si}}/\text{cm}^{-2}) > 14$ , roughly corresponding to  $\log(N_{\text{HI}}/\text{cm}^{-2}) > 19.7$ , which can be accomplished for most of our sources but not for all. Therefore, at neutral column densities below  $\log(N_{\text{HI}}/\text{cm}^{-2}) \sim 19.7$ , where HI is still optically thick, some of the LIS lines could be optically thin. For these unsaturated metal lines, Eq. (4) no longer applies, and the covering fraction cannot be derived from the residual flux of the lines.

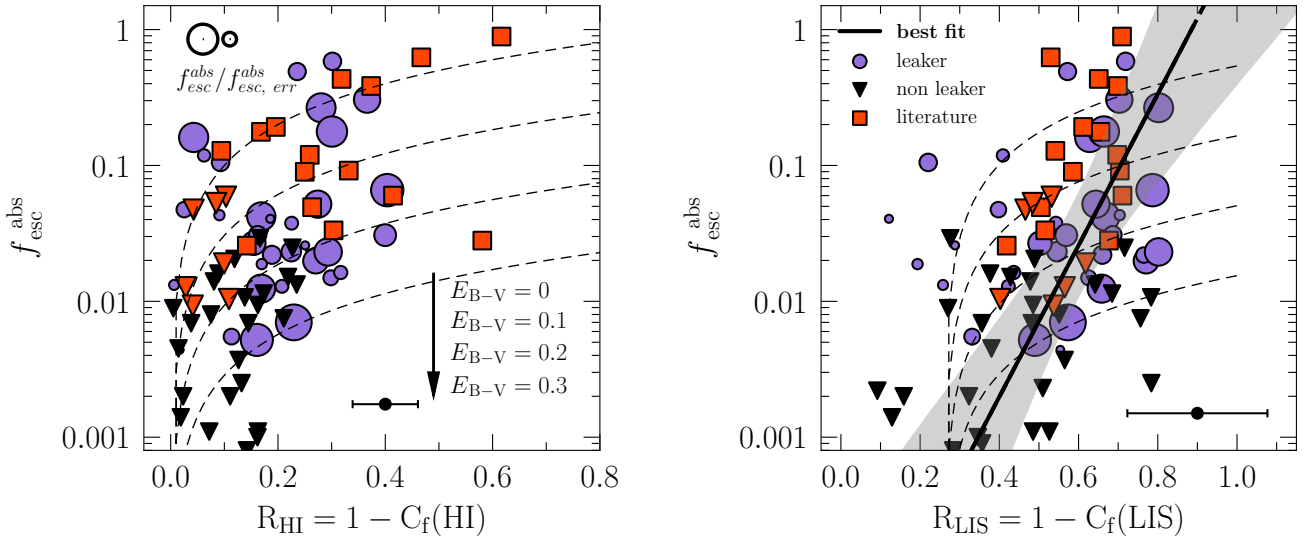
In any case, the HI residual fluxes can still be roughly predicted from the LIS line residual fluxes using Eq. (9), a trend which is also found in simulations (Mauerhofer et al. 2021). This is important at high  $z$ , where the Lyman series is not easily observable due to the increase in IGM opacity, and so the HI residual fluxes are not measurable and the  $f_{\text{esc}}^{\text{abs}}$  cannot be obtained from them. It should be finally considered that the relation between HI and LIS residual fluxes (Eq. (9)) may not be universal. However, so far the available data at high  $z$  are compatible with our findings.

### 6.3. The porosity of the ISM favors the escape of LyC photons

As discussed in Sect. 3.2, HI lines are likely saturated which means high neutral gas column densities of  $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15-16$ . Gazagnes et al. (2018) measured the HI column density of 6 out of 11 of our Izotov et al. (2016a,b, 2018a,b) leakers using the weak O I  $\lambda$ 1039 line and found column densities of  $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 18$ , which means that the neutral gas medium is also opaque to the LyC photons. Therefore, a density-bounded leaking scenario as proposed by Zackrisson et al. (2013), Jaskot & Oey (2013), Nakajima & Ouchi (2014), for example, can be excluded in principle. In contrast, the existence of saturated but nonblack HI lines suggests a nonunity HI covering fraction, where the radiation can escape through holes in the ISM, although these channels may not be free of gas and dust, and therefore may not be fully optically thin (Gazagnes et al. 2018; Steidel et al. 2018). Such a configuration can also explain the high ionization parameters and other properties associated with Green Pea galaxies (Jaskot et al. 2019), and is needed to explain the observed emission line ratios of LyC emitters (Ramambason et al. 2020).

The relationship – already seen in previous figures – between the LyC absolute escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and the residual flux,  $R(\text{HI})$  or  $R(\text{LIS})$ , is shown in Fig. 11. We performed a survival Kendall correlation analysis in order to take into account the upper limits on the escape fraction and the lower limits on the residual fluxes (see Akritas & Siebert 1996; Flury et al. 2022b). A strong ( $\tau = -0.358$ ) and significant correlation ( $p_{\text{val}} = 1.428 \times 10^{-6}$ ) between  $f_{\text{esc}}^{\text{abs}}$  and  $R(\text{HI})$  is found (see Table 3). The strongest leakers tend to have the highest residual fluxes or, conversely, the lowest HI covering fractions. Such behavior is expected from the picket-fence model, where the neutral gas is distributed in clumps around the light sources.

Clearly, HI residual flux overestimates the photon escape in most of the cases:  $f_{\text{esc}}^{\text{abs}} \leq 1 - C_f(\text{HI})$ . This is expected, because the presence of dust needs to be accounted for when inferring the  $f_{\text{esc}}^{\text{abs}}$  (Gazagnes et al. 2018; Chisholm et al. 2018; Steidel et al. 2018). In the same figure, we plot the predictions of the Eq. (6) for  $E_{\text{B-V}} = 0, 0.1, 0.2,$  and  $0.3$  (R16 law). Visually, one can see how our data points follow the curves on top given by the picket-fence scenario (uniform dust-screen). At a given  $R(\text{HI})$ , variations in the LyC photon escape can be explained by different amounts of dust reddening.



**Fig. 11.** Correlation between the LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and the HI (*left*) and LIS line (*right*) residual fluxes. Purple circles represent the leakers, while the black triangles are the nondetections – and for which the  $f_{\text{esc}}^{\text{abs}}$  is an upper limit. The size of the points corresponds to the relative uncertainty on the escape fraction ( $f_{\text{esc}}^{\text{abs}}/f_{\text{esc, err}}^{\text{abs}}$ ), and the median  $1\sigma$  uncertainty of the residual flux is placed in the bottom-right corner of each panel. Red squares correspond to the Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) galaxies (see legend). Overplotted are the curves predicted by Eq. (6) (*right*) and Eq. (12) (*left*) for  $E_{\text{B}-\text{V}} = 0, 0.1, 0.2,$  and  $0.3$  mag, respectively (dashed lines). The black solid line in the right panel corresponds to the best linear fit, and the gray-shaded area is the  $2\sigma$  confidence interval including intrinsic scatter (Eq. (10), see text).

The LyC photon escape fraction behaves similarly to  $R(\text{LIS})$  ( $f_{\text{esc}}^{\text{abs}} < 1 - C_f(\text{LIS})$ ), and shows a strong correlation too (Fig. 11,  $\tau = 0.236$ ,  $p_{\text{val.}} = 1.457 \times 10^{-3}$ ), where leakers have higher residual LIS flux than nonleakers. The differences between the two plots is principally introduced by the  $C_f(\text{HI}) - C_f(\text{LIS})$  relation (Sect. 6.2). There are, on the other hand, several possible scenarios in which the escape of ionizing photons could differ from the measured residual flux. First, as soon as the ISM conditions move from optically thick (saturated lines) to optically thin regimes (linear), the optical depths of the gas clouds are close to one, and so the covering fraction is not well-defined. The presence of residual gas and dust in the low column density channels can also contribute to the same effect (see discussion in Gazagnes et al. 2020). Secondly, the distribution of sources emitting ionizing photons and nonionizing UV continuum may not be the same. From detailed radiation transfer simulations, Mauerhofer et al. (2021) found that around 15%–30% of the light at for example  $1020 \text{ \AA}$  is produced by stars that do not contribute to the budget of ionizing radiation. Hence, these continuum photons would be affected by screens of gas and/or dust that do not influence the LyC escape fraction. Thirdly, as already mentioned in the previous section, the gas kinematics and turbulence can spread the absorption line profiles, increasing the residual fluxes and causing the covering fractions to be underestimated (also Mauerhofer et al. 2021). All these reasons would contribute to the scatter and the decreased ability to infer the  $f_{\text{esc}}^{\text{abs}}$  of galaxies (see Sect. 7.1) using the HI/LIS lines. More generally, a relation between the LyC escape and the residual fluxes of the lines, interpreted as the escape fractions of the lines in question, can always be expected (Mauerhofer et al. 2021), even when  $f_{\text{esc}}^{\text{abs}}$  cannot be analytically predicted via the picket-fence model (see Sect. 7.4).

Given the strong correlation between  $f_{\text{esc}}^{\text{abs}}$  and the residual line flux, we can also expect correlations with HI or LIS equivalent widths. These relations are displayed in Fig. 12. The  $W_{\text{HI}}$  is significantly scattered (left panel), mainly driven by the non-leaker population ( $\tau = -0.425$ ,  $p_{\text{val.}} = 1.082 \times 10^{-8}$ ). On the

LIS side (right panel), a tight correlation between the  $f_{\text{esc}}^{\text{abs}}$  and the  $W_{\text{LIS}}$  exists, with  $\tau = -0.331$  and a  $p$ -value of  $p_{\text{val.}} = 8.273 \times 10^{-6}$ .

It is of great interest to derive empirical relations using the LIS lines in order to predict the LyC escape fraction of galaxies. Nonetheless, caution is required, and all the previous limitations of the method and the current scatter in the correlations must be considered (see Figs. 11 and 12). For the LIS lines, a Bayesian linear fit to the data in log-linear space was performed using LINMIX (Kelly 2007), which provides a proper treatment of the  $f_{\text{esc}}^{\text{abs}}$  upper limits. The resulting relation between the absolute LyC escape fraction and the LIS residual flux is

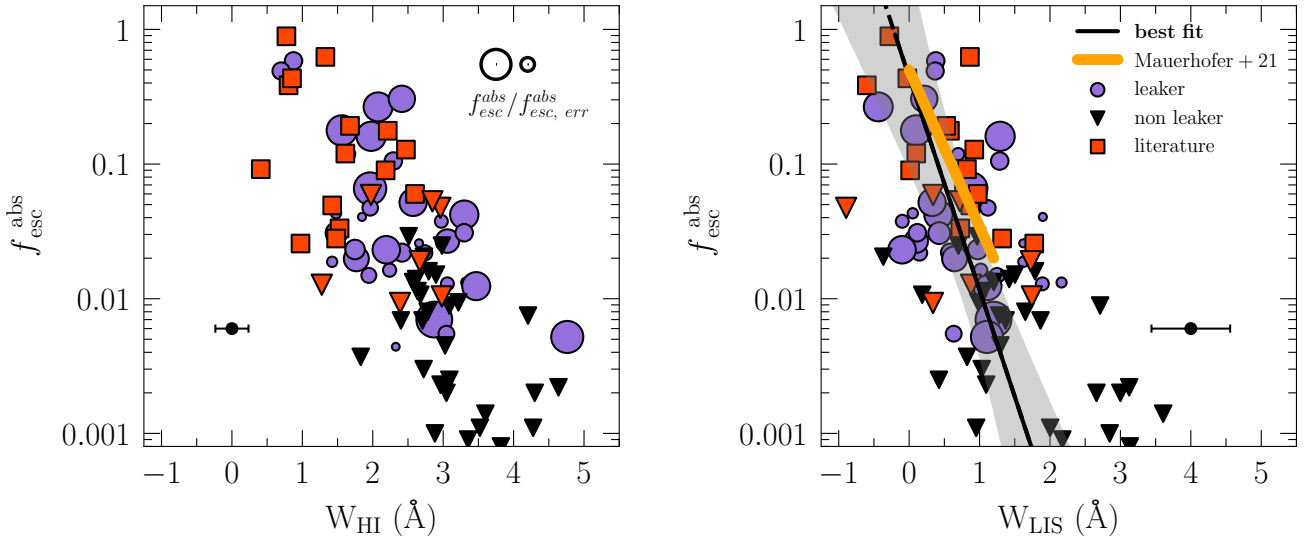
$$\log f_{\text{esc}}^{\text{abs}} = (-4.93 \pm 0.65) + (5.57 \pm 1.11) \times R(\text{LIS}), \quad (10)$$

where the linear correlation coefficient of the fit is  $r^2 = 0.69$  (+0.18, -0.21), with an important intrinsic scatter of  $\sigma_y = 0.20$ . For the  $f_{\text{esc}}^{\text{abs}}$  to  $W_{\text{LIS}}$  relation, and applying once again the LINMIX code to the log-linear data space, we get

$$\log f_{\text{esc}}^{\text{abs}} = (-0.35 \pm 0.32) - (1.58 \pm 0.34) \times W_{\text{LIS}}/\text{\AA}, \quad (11)$$

with  $r^2 = 0.88$  (+0.06, -0.09) and  $\sigma_y = 0.22$ . In Fig. 12 (right), our results are compared to the ones derived by Mauerhofer et al. (2021) using the  $\text{C II}\lambda 1334$  line. These latter authors simulated a low-mass high- $z$  galaxy and generated mock  $\text{C II}\lambda 1334$  observations along different sight lines. The  $\text{C II}\lambda 1334$  equivalent width ( $W_{\text{C II}}$ ) was then computed together with the escape fraction for every line of sight. Mauerhofer et al. (2021) concluded that the equivalent width can be treated as an upper limit to the absolute photon escape as:  $f_{\text{esc}}^{\text{abs}} < 0.5 - 0.4 \times W_{\text{C II}}$ . Compared to our linear fit, we can see that both results are similar, although the Mauerhofer et al. (2021) relation follows a slightly different slope. In the latter work, the absorption line measurements were carried out in a similar way to that described in the present paper but only for a single line. Also, the shift could be due to the modeling of the lines, the dust location within the simulated galaxy (the density of dust is proportional to the metallicity located in





**Fig. 12.** The correlation between the LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and the HI (*left*) and LIS lines (*right*) equivalent widths. The sizes and colors of the points have the same meaning as in Fig. 11. The median  $1\sigma$  uncertainty on the equivalent width is placed at the bottom of every panel. The black solid line in the right panel corresponds to the best linear fit (Eq. (11)), and the gray-shaded area represents the  $2\sigma$  confidence interval, including intrinsic scatter. The upper limit prediction from Mauerhofer et al. (2021) using the C II $\lambda$ 1334 line is plotted in orange.

the highest density HI clouds), or the different intrinsic properties (SFR, ionizing emissivity, metallicity, etc.) between our low- $z$  sample and the  $z \sim 3$  simulated galaxy. Kendall survival statistics quantifying the correlations of  $f_{\text{esc}}^{\text{abs}}$  versus HI/LIS equivalent widths and residual fluxes are computed and presented in Table 3.

#### 6.4. Dust attenuation and LyC leakage

The impact of dust extinction on the escape of ionizing photons is not well understood. Based on simulations, Kimm et al. (2019) and Mauerhofer et al. (2021) suggested that the absorption due to dust is subdominant compared to that due to neutral hydrogen. Physically, LyC photons are destroyed in the same way as nonionizing photons at slightly longer wavelengths, but if the dust resides only in regions of high neutral gas density (usually assumed in the simulations), the high ionization cross-section and large abundance of the hydrogen atom is such that HI likely dominates the absorption of LyC photons in galaxies more than dust does. However, if the dust is still present in the ionized regions or channels (as in the uniform dust-screen scenario), LyC photons would significantly attenuate before escaping the galaxy.

Chisholm et al. (2018) observationally demonstrated that dust needs to be accounted for in order to infer the absolute escape from the Lyman series. It is essential to correct for the UV attenuation in order to derive the intrinsic LyC photon production, which yields the absolute escape fraction by comparison to the observed LyC flux. Therefore, the “effective” UV attenuation  $A_{\text{UV}}$  enters any derivation of the absolute escape fraction, whereas the relative escape fraction only accounts for the differential attenuation between  $912 \text{ \AA}$  and nonionizing flux (see e.g., Siana et al. 2015; Steidel et al. 2018).

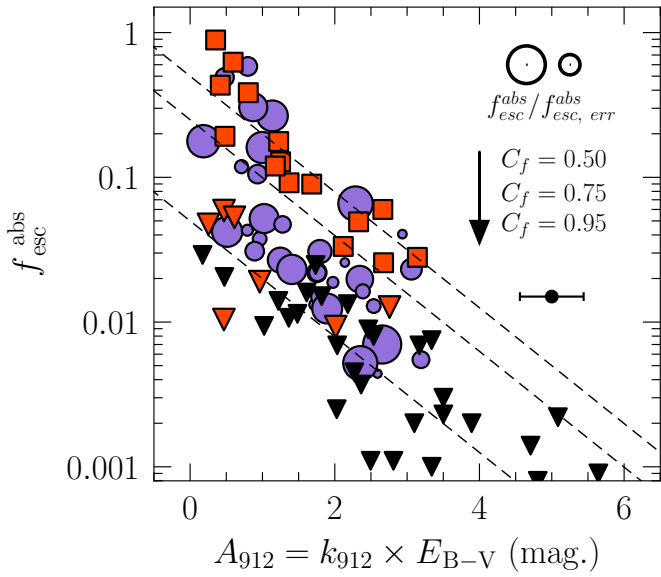
To investigate the role of the dust in the escape fraction in the LzLCS sample, in Fig. 13 we plot  $f_{\text{esc}}^{\text{abs}}$  against the dust attenuation at  $912 \text{ \AA}$ :  $A_{912} = k_{912} \times E_{\text{B-V}}$  (R16). A strong correlation of  $\tau = -0.367$  is found between the two variables ( $p_{\text{val}} = 7.624 \times 10^{-7}$ ). Most of the sources are below the  $10^{-0.4A_{\text{UV}}}$  curve (i.e.,  $f_{\text{esc}}^{\text{abs}} \leq 10^{-0.4A_{\text{UV}}}$ ,  $C_f(\text{HI}) = 0$  following Eq. (6))

which indicates, as mentioned in the previous section, that the covering fraction and dust both overestimate the LyC photon escape of galaxies when considered independently (Heckman et al. 2001, 2011; Chisholm et al. 2018; Gazagnes et al. 2020).

In Fig. 13, the results given by Eq. (6) for  $C_f(\text{HI}) = 0.50, 0.75$ , and  $0.95$  are also shown. Clearly, the spread in the  $f_{\text{esc}}^{\text{abs}}$ -dust plane can be explained by joint changes of  $E_{\text{B-V}}$  and the HI covering fraction of galaxies. We use this expression below to predict the  $f_{\text{esc}}^{\text{abs}}$  values of the LzLCS and high- $z$  galaxies. Simultaneously considering the neutral gas covering (given by the depths of the Lyman lines) and the dust attenuation (from the UV spectral fits) provides the most accurate photon escape prediction of this work.

We note that a correlation between  $f_{\text{esc}}^{\text{abs}}$  and  $E_{\text{B-V}}$  is expected because the  $f_{\text{esc}}^{\text{abs}}$  values come from de-reddening the S99 models (see Eqs. (1) and (2)) and comparing the predicted LyC flux to the observed one. Ideally, an independent measurement of the UV attenuation, which is here obtained from fitting the UV spectra, would test if this correlation is indeed as strong as shown here. We also note that the exact value of the UV attenuation (or  $E_{\text{B-V}}$ ) depends in principle on the assumed geometry and the attenuation law, for which we have only considered two simple cases (see Sect. 4). Probably the most direct verification of the robustness of these relations comes from the comparison of the absolute escape fractions derived using two independent methods, namely the one presented here, which uses the stellar UV spectrum, and the method used by Flury et al. (2022a) (see also Izotov et al. 2016b). In the latter study,  $f_{\text{esc}}^{\text{abs}}$  is computed from the observed LyC flux and the extinction-corrected  $\text{H}\beta$  flux, together with  $\text{H}\beta$  equivalent width which is used as an indicator of the age of the underlying ionizing stellar population. As shown in Flury et al. (2022a), the overall agreement between the two methods is quite good, verifying the strong correlation that we see here between the escape and the dust-attenuation.

Finally, it is worth mentioning that the correlations between the absolute LyC escape fraction, residual fluxes, equivalent widths, and the UV attenuation shown in Figs. 11–13 also hold when one considers the observed flux ratio  $f_{\text{LyC}}/f_{1100}$  instead of  $f_{\text{esc}}^{\text{abs}}$ . In particular, the fact that both  $f_{\text{LyC}}/f_{1100}$  and  $f_{\text{esc}}^{\text{abs}}$  correlate



**Fig. 13.** Correlation between the LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and the UV dust-attenuation:  $A_{912} = k_{912} \times E_{B-V}$ . Labeling is as in Fig. 11. The median  $1\sigma$  uncertainty on  $A_{UV}$  is plotted on the right of the plot. Overplotted are the three lines predicted by Eq. (6) for  $C_f(\text{HI}) = 0.5, 0.75,$  and  $0.95,$  respectively (dashed lines), using  $k_{912} = 12.87$  (R16 law).

**Table 3.** Survival Kendall ( $\tau$ ) correlation test results for  $f_{\text{esc}}^{\text{abs}}$  versus different galaxy properties.

	$f_{\text{esc}}^{\text{abs}}$	
	$\tau$	$p_{\text{val}}$
R(HI)	+0.358	$1.428 \times 10^{-6}$
R(LIS)	+0.236	$1.457 \times 10^{-3}$
$W_{\text{HI}}$	-0.425	$1.082 \times 10^{-8}$
$W_{\text{LIS}}$	-0.331	$8.273 \times 10^{-6}$
$A_{UV}$	-0.367	$7.624 \times 10^{-7}$

**Notes.** HI and LIS equivalent widths ( $W_{\text{HI}}, W_{\text{LIS}}$ ), HI and LIS line residual fluxes (R(HI), R(LIS)) and UV attenuation ( $A_{UV}$ ) accounting for censored data (upper or lower limits).

**References.** Akritas & Siebert (1996).

with the UV attenuation indicates that dust is primarily in the clouds, as shown in Sect. 5.2 (see also Sect. 7.4), and hence the amount of attenuation increases with the covering fraction.

For easier comparison with future studies, and to better illustrate the behavior of the LyC escape with the previous parameters, a summary plot of Figs. 11–13 can be seen in Appendix E (Fig. E.1). To this end, we computed the median value of R(HI), R(LIS),  $W_{\text{HI}}$ ,  $W_{\text{LIS}}$  and  $E_{B-V}$  in three  $f_{\text{esc}}^{\text{abs}}$  intervals, specifically at  $f_{\text{esc}}^{\text{abs}} \in [0-5\%), [5-10\%),$  and  $[10-100\%).$  Table E.1 emphasizes these results, and a quantitative analysis of these trends, comprising the sample median values, will be further investigated in future LzLCS publications using stacked spectra.

## 7. Discussion

Having established relations between the UV absorption lines and LyC escape, we now examine the power of these lines to predict the ionizing photon escape fraction of galaxies (Sect. 7.1). We also present a physical picture of the ISM of low- $z$  LyC

emitters based on our UV absorption lines and  $f_{\text{esc}}^{\text{abs}}$  results as well as earlier results in the literature (Sect. 7.2). We finally discuss some implications and comparisons for the high- $z$  Universe (Sect. 7.3), and the caveats of our modeling approach (Sect. 7.4).

### 7.1. Predicting the Lyman-continuum photon escape with the Lyman series and LIS lines

We want to test whether the HI and, more particularly, measurements of the LIS lines could be used to effectively predict the  $f_{\text{esc}}^{\text{abs}}$  of galaxies. Under the assumption of a foreground screen of dust in the picket-fence model, Chisholm et al. (2018) determined that the parameters that contribute the most to predicting the photon escape are the dust attenuation and the HI covering fraction, implying that the column density negligibly affects the line depths (if the lines are saturated). Therefore, the accurate determination of both  $E_{B-V}$  and  $C_f(\text{HI})$  is necessary. In this section, we carry out a similar exercise as in Chisholm et al. (2018) and Gazagnes et al. (2020), but taking advantage of our larger LzLCS data set, and thus improving the robustness of the method and providing more statistical meaning. First, we take the derived  $E_{B-V}$  from our SED fits as well as the weighted-averaged HI covering fractions,  $C_f(\text{HI})$ . The predicted photon escape from the HI lines is computed from Eq. (6),  $f_{\text{esc}}^{\text{abs, HI}}$ .

Thus, the predicted  $f_{\text{esc}}^{\text{abs, HI}}$  value is compared to the fiducial value derived from the spectral UV fits (Sect. 3.3) in Fig. 14 (top). LINMIX determines a clear correlation between the two, with a correlation coefficient of  $r^2 = 0.97 (+0.02, -0.07)$ . The intrinsic scatter is negligible ( $\sigma_y \sim 10^{-4}$ ), which means that the errors on the variables can fully explain the scatter in the 1:1 relation. The median of the predicted  $f_{\text{esc}}^{\text{abs}}$  for leakers does not change with respect to the fiducial  $f_{\text{esc}}^{\text{abs}}$  distribution, although a small systematic shift to higher predicted escape values exists for the  $f_{\text{esc}}^{\text{abs}} < 10\%$  distribution, where a significant fraction of the sources are upper limits in  $f_{\text{esc}}^{\text{abs}}$ . In conclusion,  $f_{\text{esc}}^{\text{abs}}$  can be well recovered from the Lyman series and dust attenuation.

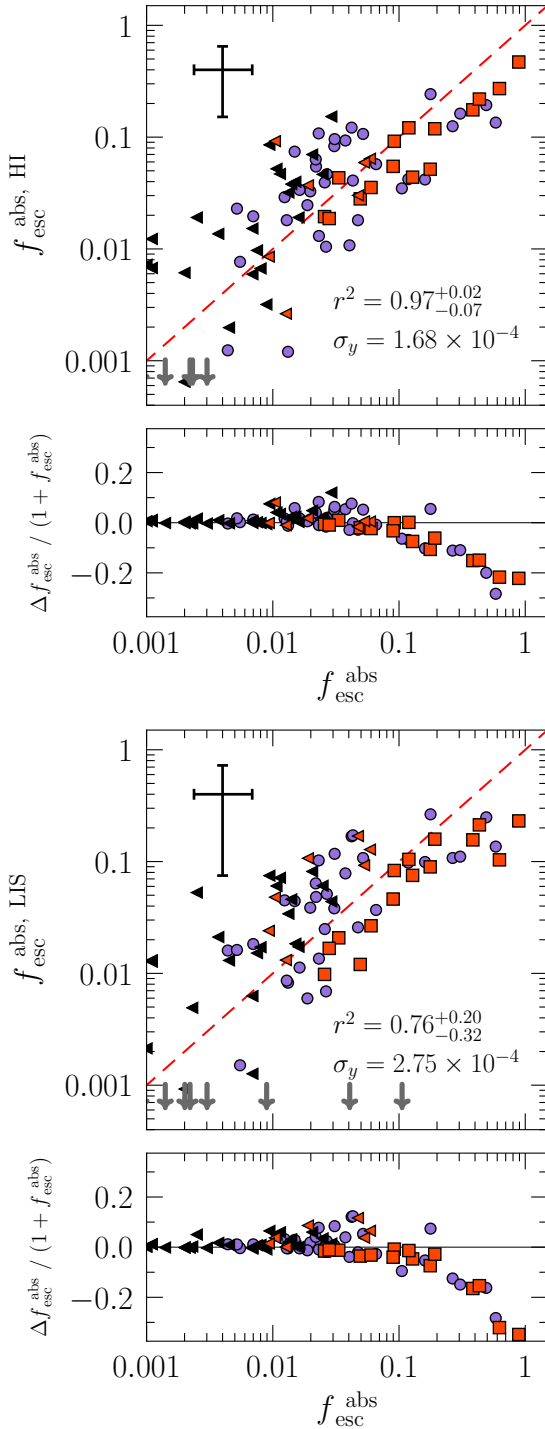
In any case, our original purpose is to decipher whether or not the  $f_{\text{esc}}^{\text{abs}}$  can be predicted using the LIS lines. To approach this, we first take advantage of the intrinsic relation between the residual flux (or covering factor) of HI and metals, which is empirically given by our Eq. (9). The HI is then computed by substituting every measured  $C_f(\text{LIS})$  in that expression, and then the predicted photon escape from LIS,  $f_{\text{esc}}^{\text{abs, LIS}}$ , is determined with a formula similar to that in Eq. (6):

$$f_{\text{esc}}^{\text{abs, LIS}} = 10^{-0.4k_{912}E_{B-V}} \times (1 - [a \times C_f(\text{LIS}) + b]), \quad (12)$$

where the coefficients  $[a, b] = [(0.63 \pm 0.19), (0.54 \pm 0.09)]$  are given by Eq. (9). The results are plotted against the fiducial  $f_{\text{esc}}^{\text{abs}}$  in Fig. 14 (bottom). The gray arrows point towards  $f_{\text{esc}}^{\text{abs, LIS}} \approx 0$ , for those sources with  $C_f(\text{LIS}) \approx 1$  (zero residual flux) and/or very high  $E_{B-V}$ , falling below the axis limits. Still, a strong linear correlation is present when using the LIS lines ( $r^2 = 0.76 (+0.20, -0.32)$ ), with no significant intrinsic scatter.

As discussed earlier (Sect. 6.3), several mechanisms can lead to disagreement between the real covering fraction and the covering derived using the residual flux of the lines: ISM regions close to the optically thin limit, the mismatch between the distribution of UV ionizing and nonionizing sources, the gas kinematics and turbulence, and the ISM geometry can all potentially contribute to the scatter in the relations shown in Fig. 14 (see Mauerhofer et al. 2021).

A more detailed look into Fig. 14 reveals that for the highest  $f_{\text{esc}}^{\text{abs}} \geq 0.1$  sources, the Lyman series and low-ionization covering methods systematically underestimate the photon escape.



**Fig. 14.** Comparison of predictive methods for the LyC absolute escape fraction. *Top:* Lyman series photon escape (Eq. (6)) against the fiducial S99-derived values. *Bottom:* same as top panel but for the photon escape inferred using the LIS line covering fraction (Eq. (9)). Purple points correspond to the LzLCS leakers while the red squares are the Izotov et al. (2016a,b, 2018a,b, 2021) and Wang et al. (2019) galaxies. Black triangles correspond to nonleakers. Median relative error bars are shown in the upper right corners of each panel. The subplots at the bottom illustrate the overall accuracy of the predicted  $f_{\text{esc}}^{\text{abs}}$  (i.e.,  $\Delta f_{\text{esc}}^{\text{abs}} / (1 + f_{\text{esc}}^{\text{abs}})$ ) as a function of the fiducial  $f_{\text{esc}}^{\text{abs}}$ .

The effect is more noticeable when one plots the overall accuracy of the predicted  $f_{\text{esc}}^{\text{abs}}$  ( $\Delta f_{\text{esc}}^{\text{abs}} / (1 + f_{\text{esc}}^{\text{abs}})$ ) as a function of the fiducial  $f_{\text{esc}}^{\text{abs}}$  itself (Fig. 14, bottom panels). This result, which

was already noted by Gazagnes et al. (2020) and is supported by hydrodynamical simulations by Kakiichi & Gronke (2021), suggests that the leakage could be partially due to a density-bounded ISM, where the LyC photons can escape through optically thin  $\log(N_{\text{HI}}) \text{ cm}^{-2} \lesssim 10^{17.2}$  regions to the LyC.

Overall, we propose three different approaches to predict the LyC absolute photon escape from the LIS lines, and demonstrate the limitations on each: (1) using the residual flux of the absorption lines (Eq. (10)), (2) their equivalent widths (Eq. (11)), and (3) a combination of residual flux and dust attenuation (Eq. (12)). When compared directly to  $f_{\text{esc}}^{\text{abs}}$ , the third method is ideally recommended in all the cases, as it presents the most stringent correlation coefficient of the fit and the least intrinsic scatter.

A similar methodology to (3) allowed Chisholm et al. (2018) to predict the  $f_{\text{esc}}^{\text{abs}}$  of five high- $z$  galaxies detected in LyC, and without Lyman-series observations. In particular, for the Cosmic Horseshoe (Belokurov et al. 2007) at  $z = 2.38$ , these authors predicted a  $f_{\text{esc}}^{\text{abs}}$  of around 1%, in agreement with the  $f_{\text{esc}}^{\text{abs}} < 2\%$  limit imposed by HST imaging (Vasei et al. 2016).

At high  $z$ , low-mass galaxies are characterized by hosting a metal- and dust-poor ISM similar to our low- $z$  analogs, where  $E_{B-V} < 0.1$  at  $z \sim 3-6$  (Schaerer et al. 2013). Taking this attenuation as representative, and using the median HI and metal LIS covering fraction for the leakers in our sample (0.70 and 0.38, respectively), this gives  $f_{\text{esc}}^{\text{abs}} \geq 0.07$  for HI and LIS predicted values. This is below but similar to the 10% theoretical lower limit for faint SF galaxies to dominate reionization at  $z \sim 6-9$  (Robertson et al. 2013)<sup>8</sup>, and is in good agreement with the 9% (6%) average value derived in Steidel et al. (2018), Pahl et al. (2021) at the same redshifts; and also Rosdahl et al. (2018) in simulations. However, it should also be stressed that the predictions for  $f_{\text{esc}}^{\text{abs}}$  given by Eq. (12) are only reliable in a statistical sense. Given the scatter in the above-mentioned relations, predictions for any individual object are not very reliable.

We finally point out that higher resolution and higher S/N UV spectra would help to obtain a better measurement of the covering fractions and ionic column densities (fundamental to disentangling the role of the ISM structure and gas kinematics in LyC leakage), which would improve the  $f_{\text{esc}}^{\text{abs}}$  determination from the SED modeling (subject to quite large uncertainties in some cases), and to better constrain the stellar population and dust properties ( $E_{B-V}$ ).

## 7.2. The absorbing ISM of LyC emitters

Low-metallicity (subsolar) galaxies typically have lower masses ( $\lesssim 10^{10} M_{\odot}$ ), higher gas fractions, and less dust than the typical local SF galaxy (Trebitsch et al. 2017). Low-mass systems have lower gravitational potentials, which increases the efficiency of stellar feedback and leads to more bursty star-formation episodes (Muratov et al. 2015). The low-metallicity and burstiness of their star-formation history means that these sources also have high ionizing production efficiencies (intrinsic number of ionizing photons per UV luminosity or mass, equivalently) and moderate ionizing escape fractions (Ma et al. 2020). This, together with their higher number density at all epochs (with respect to more massive galaxies), makes them the most promising analogs of the sources of cosmic reionization (Finkelstein et al. 2019).

Different mechanisms can remove the gas and dust from such galaxies, creating holes in the ISM with appreciably lower

<sup>8</sup> The reader may notice here that various samples have different  $M_{\text{UV}}$  ranges. For instance, Robertson et al. (2013) consider  $M_{\text{UV}} \lesssim -15$  for  $z \sim 7$  galaxies, while our sample spans  $-18 \leq M_{\text{UV}} < -21$  at  $z \sim 0.3$ .



gas and dust column densities. Our observations suggest that LyC photons primarily escape through these channels. These holes are likely created by stellar winds and supernova explosions after starburst episodes, or they are photoionized channels induced by turbulence during the starburst period (Kakiichi & Gronke 2021). They could also be embedded in an intrinsically low-density ISM, which results from catastrophic cooling effects around low-metallicity star clusters (Jaskot et al. 2019).

The picket-fence model describes an ionization-bounded galaxy (Zackrisson et al. 2013; Vasei et al. 2016; Reddy et al. 2016b) where the ionizing photons can partially escape through these holes in the ISM. The fraction of sight lines covered by optically thick neutral gas clouds is the covering fraction. We now summarize a scenario where the HI covering fraction,  $C_f(\text{HI})$ , can explain most of the observed empirical trends found in the present study.

The presence of (likely) saturated but nonblack HI and LIS absorption lines in our spectra (Sect. 5.2) implies the existence of low-column-density channels, and therefore a nonzero covering fraction for most of the galaxies. Measurements of the HI column density of Green-Pea galaxies lead to similar conclusions in Gazagnes et al. (2018) and McKinney et al. (2019), and in the recent, detailed study of the local galaxy Haro11 by Östlin et al. (2021).

The LzLCS sample reveals many strong correlations between the UV absorption lines and other secondary galaxy parameters. We find that the measured HI and LIS equivalent widths scale with (1) the intrinsic absolute magnitudes at 1500 Å,  $M_{1500}^{\text{int}}$ , and with (2) the UV attenuation,  $A_{\text{UV}}$  (or  $E_{\text{B-V}}$ , equivalently), and scale inversely with (3) the HI and LIS residual fluxes (Figs. 6 and 7). We also demonstrated that metallicity plays a minor role in these relations.

The inverse connection between the equivalent width of the lines and the residual flux (3) naturally comes from the definition of a saturated absorption line profile (Draine 2011). In contrast, the  $W_\lambda - A_{\text{UV}}$  relation (2) can be explained by invoking a clumpy gas-to-dust geometry (Shapley et al. 2003; Steidel et al. 2018; Du et al. 2018, 2021; Pahl et al. 2020), where the total reddening is related to the fraction of the total sight lines with optically thick neutral gas (Vasei et al. 2016; Reddy et al. 2016b). A galaxy with a larger fraction of sight lines covered by optically thick neutral gas (a higher  $C_f(\text{HI})$ ) also encounters more dust and the spectrum is more attenuated ( $A_{\text{UV}}$  increases), because the dust resides in the same gas. In other words, the same gas which gives rise to  $W_{\text{LIS}}$  also reddens the continuum (Jones et al. 2012).

Aside from the  $W_\lambda - A_{\text{UV}}$  connection, there are other lines of evidences for a clumpy gas-to-dust geometry in our sample. For instance, in a clumpy geometry, transitions for the same ion at different wavelengths will have different residual fluxes (see Eq. (5) and Gazagnes et al. 2018). We tentatively see this trend when looking at the median residual flux of every Lyman line in the LzLCS. Values vary by as much as 0.71, 0.81, 0.83, and 0.89 from Ly5 to Ly $\beta$ , but typical errors of  $\sim 0.13$  obscure robust conclusions. The hypothesis of a clumpy distribution of gas and dust has been widely proposed in the literature, and it has special relevance for example when trying to explain the observed Ly $\alpha$ -to-H $\alpha$  ratios in SF galaxies (see Scarlata et al. 2009; Jaskot et al. 2019, and references therein), and the Ly $\alpha$  peak-velocity separation (Gronke et al. 2016; Orlitová et al. 2018; Jaskot et al. 2019).

When Ly $\alpha$  varies from weak absorption to strong emission, the strengths of the HI and LIS absorption lines decrease (Fig. 9 in Sect. 6.1, see also Shapley et al. 2003; Jones et al. 2012; Rivera-Thorsen et al. 2015; Du et al. 2018, 2021; Trainor

et al. 2019; Pahl et al. 2020). Invoking our picket-fence model, the higher the covering fraction, the higher the fraction of Ly $\alpha$  photons that are resonantly scattered out of the line of sight and absorbed by dust, resulting in a weaker Ly $\alpha$  emission with stronger HI and LIS profiles (Steidel et al. 2018; Gazagnes et al. 2020). Similarly, using  $z \sim 3-5$  composites, Jones et al. (2012) demonstrated that other host properties, such as the luminosity, have little effect on the  $W_{\text{Ly}\alpha} - W_{\text{LIS}}$  variations. The relationships between  $W_{\text{Ly}\alpha} - W_{\text{LIS}}$  and  $W_{\text{Ly}\alpha} - A_{\text{UV}}$  are presumably invariant across cosmic time, as both are affected by variations in the same HI covering (Du et al. 2018). This also suggests that the Ly $\alpha$  production efficiency does not change drastically with redshift (Trainor et al. 2019; Pahl et al. 2020).

Moreover, the existence of double- or triple-peaked Ly $\alpha$  profiles in leakers (Verhamme et al. 2017; Izotov et al. 2020, 2021) can be associated with the characteristic bimodal distribution of HI column densities in the patchy ISM (Verhamme et al. 2015; Mauerhofer et al. 2021; Kakiichi & Gronke 2021), or the presence of completely ionized channels (Rivera-Thorsen et al. 2017). This may lead to the observed correlations between the escape of Ly $\alpha$  and LyC photons (e.g., Verhamme et al. 2017; Marchi et al. 2018; Steidel et al. 2018; Gazagnes et al. 2020; Pahl et al. 2021).

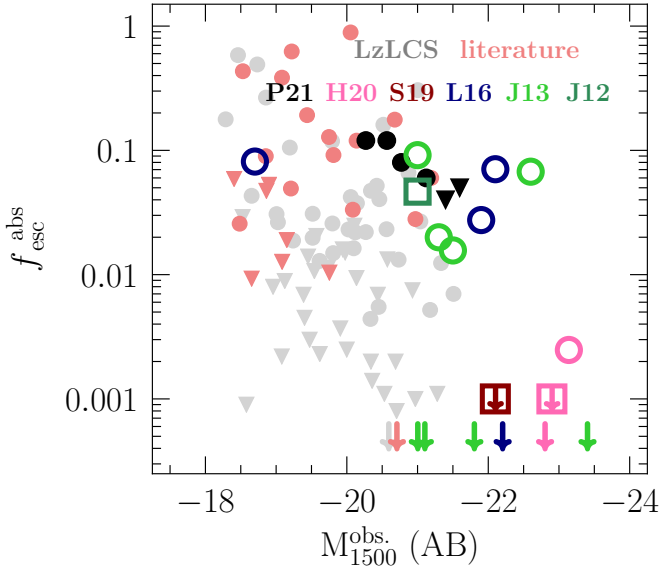
While the picket-fence model allows us to predict (at least on average) the LyC escape fraction relatively well, we find that the absorption lines tend to underestimate the escape fractions for the strongest leakers with  $f_{\text{esc}}^{\text{abs}} \gtrsim 0.1 - 0.2$  (see also Gazagnes et al. 2020). This suggests that there could be a secondary scenario where LyC emission is regulated by a density-bounded ISM in these galaxies, as suggested by several previous studies (Zackrisson et al. 2013; Jaskot & Oey 2013; Nakajima & Ouchi 2014; Jaskot et al. 2019). Indeed, from simulations, Kakiichi & Gronke (2021) show that a switch from an ionization-bounded regime to one that is density-bounded may naturally happen at evolved stages in the history of a galaxy. This transition occurs when turbulence ceases and the ionization radius then increases with time in a more static situation. Indications for the existence of the density-bounded regime for leakers with the highest  $f_{\text{esc}}^{\text{abs}}$  was presented by Gazagnes et al. (2020) and Ramambason et al. (2020), but whether this also quantitatively agrees with the present observations will be examined in future publications.

### 7.3. The LyC escape fraction of high- $z$ galaxies

With the empirically calibrated and theoretically motivated relations between the residual flux, UV attenuation, and the absolute LyC escape fraction (Eq. (12)), we can estimate  $f_{\text{esc}}^{\text{abs}}$  from literature measurements of the LIS absorption lines for high- $z$  galaxies, and compare them to local measurements from the LzLCS.

Figure 15 presents the result of this exercise, showing the predicted  $f_{\text{esc}}^{\text{abs}}$  from individual and stacked galaxy spectra reporting residual flux measurements of the UV LIS lines as a function of the observed absolute UV magnitude (the latter without a correction for UV attenuation). To compute the LyC escape fraction, we use the reported LIS residual flux, and then Eq. (9) to estimate the  $C_f(\text{HI})$ , and the R16 attenuation law to estimate the  $E_{\text{B-V}}$  from the average dust corrections given in Bouwens et al. (2020), as a function of redshift<sup>9</sup>. We also show the  $f_{\text{esc}}^{\text{abs}}$  measurements at low  $z$  from the present paper (gray, red circles and downward point triangles), and the latest update by

<sup>9</sup> More precisely, we adopt  $E_{\text{B-V}} = 0.107, 0.095, 0.086, 0.058$  for  $z = 3, 4, 5, 6$ .

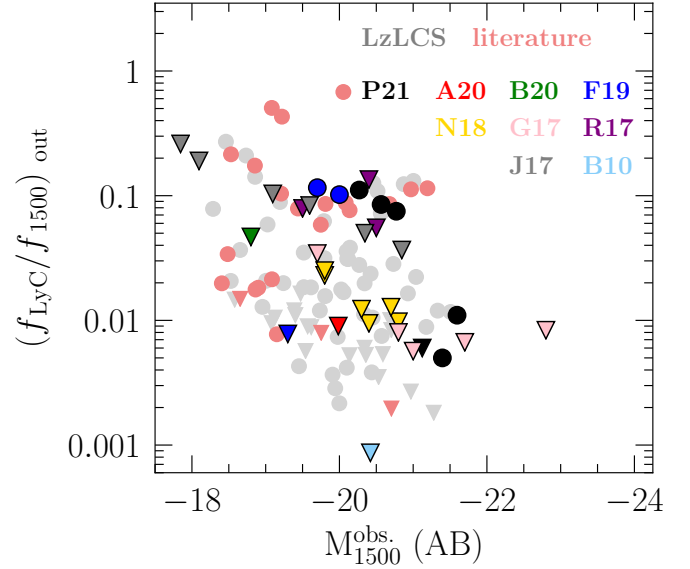


**Fig. 15.** Estimated and measured LyC absolute escape fractions of high- and low- $z$  galaxies as a function of the observed  $M_{1500}^{\text{obs}}$ . Open circles and squares show the  $f_{\text{esc}}^{\text{abs}}$  predictions from the UV LIS lines for individual galaxies and stacks, respectively, taken from [Harikane et al. \(2020\)](#) at  $z \sim 6$ , in pink; [Sugahara et al. \(2019\)](#) at  $z \sim 5$ , in dark red; [Leethochawalit et al. \(2016\)](#) at  $z \sim 4-5$ , in blue; [Jones et al. \(2013\)](#) at  $z \sim 2-4$ , in light green; and [Jones et al. \(2012\)](#) at  $z \sim 4$ , in green. Direct  $f_{\text{esc}}^{\text{abs}}$  measurements from this paper (LzLCS and literature) are plotted in light gray and red;  $z \sim 3$  measurements from [Pahl et al. \(2021\) measurements are displayed with filled black symbols, for comparison. Downward-pointing arrows represent sources with estimated  \$f\_{\text{esc}}^{\text{abs}} \approx 0\$ .](#)

[Pahl et al. \(2021\)](#) of the LyC escape fractions measured directly from stacks of  $z \sim 3$  galaxy spectra from the Keck Lyman Continuum Survey (KLCS, black circles; [Steidel et al. 2018](#)). Downward arrows means predicted  $f_{\text{esc}}^{\text{abs}} \approx 0$ , for those sources with  $C_f(\text{LIS}) \approx 1$  (zero residual flux) and/or very high  $E_{B-V}$ , falling below the axis limits.

As seen in Fig. 15, all our estimates yield absolute LyC escape fractions  $f_{\text{esc}}^{\text{abs}} < 0.1$  at  $z \gtrsim 4$  for relatively bright  $M_{1500}^{\text{obs}} \lesssim -21$  galaxies. Measurements from individual sources show large scatter, as also found at low  $z$ . For the highest redshift available ( $z \sim 6$ ), the stack and the galaxy J0210–0523 of [Harikane et al. \(2020\)](#), with  $M_{1500}^{\text{obs}} \sim -23$ , and our method predict negligible escape fractions,  $f_{\text{esc}}^{\text{abs}} \lesssim 0.003$ . From the  $z \sim 4$  stack of [Jones et al. \(2012\)](#) with  $M_{1500}^{\text{obs}} \sim -21$ , we estimate  $f_{\text{esc}}^{\text{abs}} \sim 0.05$ , which is comparable to the value measured at  $z \sim 3$  for galaxies with similar UV magnitude ([Pahl et al. 2021](#)). We reiterate the fact that our  $f_{\text{esc}}^{\text{abs}}$  values of the high- $z$  galaxies are lower than those estimated in [Jones et al. \(2012\)](#), [Harikane et al. \(2020\)](#), because we account for both the residual flux and the UV attenuation ([Gazagnes et al. 2018](#); [Chisholm et al. 2018](#)). Escape fractions of  $\sim 10\%$  are found at  $M_{1500}^{\text{obs}} \sim -20$  from the KLCS stacks, and [Pahl et al. \(2021\)](#) find decreasing  $f_{\text{esc}}^{\text{abs}}$  towards UV-brighter galaxies, which overall appears consistent with our estimates. However, the available data are still scarce, and whether  $f_{\text{esc}}^{\text{abs}}$  increases further in fainter galaxies must be confirmed with future observations (see [Bian & Fan 2020](#)).

The overlap between the high- and low- $z$  data in observed UV magnitude is quite limited (Fig. 15) and meaningful sample averages remain to be obtained at low redshift. It is therefore difficult to draw conclusions on the possible redshift evolution of  $f_{\text{esc}}^{\text{abs}}$  from this figure, because possible trends may



**Fig. 16.** Ionizing-to-nonionizing observed flux ratio  $(f_{\text{LyC}}/f_{1500})_{\text{out}}$  in units of  $F_{\nu}$  as a function of  $M_{1500}^{\text{obs}}$  for high- and low- $z$  galaxies. At high redshift, we only show measurements determined from stacks. LyC detections from [Pahl et al. \(2021\)](#) and [Fletcher et al. \(2019\)](#) at  $z \sim 3$  are plotted in black and blue symbols. Circles show detections and downward pointing triangles are  $1\sigma$  upper limits that are color coded by sample ([Pahl et al. 2021](#); [Alavi et al. 2020](#); [Bian & Fan 2020](#); [Fletcher et al. 2019](#); [Naidu et al. 2018](#); [Grazian et al. 2017](#); [Rutkowski et al. 2017](#); [Japelj et al. 2017](#)). The high- $z$  data probe  $z \sim 2.5-4$ , except for the study of [Alavi et al. \(2020\)](#) at  $z \sim 1.3$ , and [Bridge et al. \(2010\)](#) at  $z \sim 0.7$ . The low- $z$  data from the present study (LzLCS and literature) are plotted in light gray and red, as in the previous figure.

be a consequence of a general luminosity evolution or due to different detection limits within the high- $z$  samples. Furthermore, the adopted UV attenuations for the high-redshift samples are unconstrained, making the comparison of  $f_{\text{esc}}^{\text{abs}}$  as a function of the intrinsic UV magnitude also difficult to interpret. In any case, we note that both individual galaxies and stacks with  $f_{\text{esc}}^{\text{abs}} \sim 5\%-10\%$  are found over a range of four magnitudes (from  $M_{1500}^{\text{obs}} \sim -19$  to  $-22$ ). Objects with  $f_{\text{esc}}^{\text{abs}} > 0.1$  are (so far) found only at magnitudes fainter than  $M_{1500}^{\text{obs}} < -20$  at low  $z$ . However, some individual galaxies at  $z > 3$  are bright ( $M_{1500}^{\text{obs}} < -21$ ) and have high LyC escape fractions (e.g., Ion2 in [Vanzella et al. 2016](#)). Thus, bright high-redshift galaxies may have markedly different LyC escape properties. Overall, the low- $z$  results concur, albeit with the low LyC escape fractions found in bright galaxies ( $M_{1500}^{\text{obs}} \lesssim -20$ ) by many studies using stacks at high  $z$  (see e.g., [Grazian et al. 2016](#); [Guaita et al. 2016](#); [Marchi et al. 2017](#); [Japelj et al. 2017](#)).

Another quantity to be compared with high- $z$  studies is the observed ‘‘Lyman decrement’’ between the LyC and the nonionizing UV flux, that is, the commonly used  $(f_{\text{LyC}}/f_{1500})_{\text{out}}$  corrected for the IGM attenuation. This quantity can directly lead to the ionizing emissivity of galaxies when integrating over the UV luminosity function (see e.g., [Pahl et al. 2021](#)). In Fig. 16 we compile the available data from high- $z$  stacks, for which we plot the IGM-corrected measurements of  $(f_{\text{LyC}}/f_{1500})_{\text{out}}$ . We use stacked spectra to compensate for differences in the mean IGM transmission. The data comprise results from a variety of surveys mostly carried out at redshift  $z \sim 2.5-4$ , plus the study of [Alavi et al. \(2020\)](#) at  $z \sim 1.3$ , and [Bridge et al. \(2010\)](#) at  $z \sim 0.7$ . We have also added the data from the composite SEDs

of the LAE sample of Fletcher et al. (2019) at  $z = 3.1$ , who report LyC-detected and nondetected subsamples.

Interestingly, the Lyman decrement observed both from the KLCS (Steidel et al. 2018; Pahl et al. 2021) and the LAE sample of Fletcher et al. (2019) at  $z \sim 3$  are very similar, and comparable to the observed values  $(f_{\text{LyC}}/f_{1500})_{\text{out}} \sim 0.1$  at  $z \sim 0.3$  and at same UV magnitudes,  $M_{1500}^{\text{obs}} \sim -19.5$  to  $-21$ . The upper limits found for galaxies with  $M_{1500}^{\text{int}} \lesssim -21$  indicate small Lyman decrements that are compatible with low-redshift observations. On the UV-faint end ( $M_{1500}^{\text{obs}} \gtrsim -20$ ), the high- and low- $z$  observations also appear fairly consistent with each other<sup>10</sup>. Further studies are required to better understand what determines the differences found between the various high- and low- $z$  samples and to determine representative population averages among low-redshift galaxies. In summary, the high- $z$  observations show a consensus of decreasing Lyman decrement towards brighter UV magnitudes, although the overall correlation is weak. This trend is also tentatively shared by the LzLCS sample.

#### 7.4. Main assumptions, limitations, and caveats

Among the possible sources of uncertainty in our analysis are (1) the choice of the dust-attenuation law, (2) the assumed gas and dust geometry, (3) the adopted picket-fence model, and (4) the definition and meaning of the LyC escape fraction.

(1) *Assumptions on the dust-attenuation law.* As previously described (Sect. 4), the choice of the attenuation curve primarily impacts the derived stellar population parameters and UV attenuation. In particular,  $X_i$  coefficients (see Eq. (1)) and the dust-attenuation parameters ( $E_{B-V}$ ) can both change, while average stellar ages and metallicities remain consistent (same light-fractions, see Chisholm et al. 2019).

A steeper attenuation law (like the SMC curve) yields a lower UV attenuation and hence systematically higher escape fractions  $f_{\text{esc}}^{\text{abs}}$  compared to the flatter R16 attenuation law adopted as default here. Typically,  $f_{\text{esc}}^{\text{abs}}$  is found to be higher by a factor of between 1.05 and 1.50 from high to low LyC escape fractions (see Fig. 4). The intrinsic, that is, extinction-corrected UV magnitudes ( $M_{1500}^{\text{int}}$ ) are therefore also fainter for the SMC law. On the other hand, line equivalent widths and residual fluxes are essentially independent of the attenuation law, because the stellar continuum level is similarly modeled for the different laws.

From the present UV data alone, it is difficult to distinguish between the two attenuation laws used here. A joint analysis of the full SED from the UV to the optical, including an independent attenuation measurement such as the Balmer decrement, would be necessary to constrain the attenuation law (see e.g., Shivaiei et al. 2020). The offsets found between the color excess  $E_{B-V}$  derived from the UV spectra and the Balmer decrement shown in Fig. 3 indicate that an attenuation law steeper than the R16 curve is probably appropriate for a fraction of the galaxies in our sample. For the compact, high- $O_{32}$  LyC emitters detected prior to the LzLCS, an SMC-like or even steeper curve in some cases had been found by Izotov et al. (2016b, 2018a,b).

Also, from the analysis of 218  $z = 1.4$ – $2.6$  SF galaxies with both UV rest-frame and Balmer decrement measurements plus

<sup>10</sup> The upper limit of  $(f_{\text{LyC}}/f_{1500})_{\text{out}}$  at  $M_{1500}^{\text{obs}} = -19.3$  from Fletcher et al. (2019) is low by construction, because this represents the composite SED of LAEs with nondetections in the LyC.

known gas-phase metallicities, Shivaiei et al. (2020) suggest that “low-metallicity” galaxies have a steeper, SMC-like attenuation curve, whereas galaxies with  $12 + \log(O/H) \gtrsim 8.5$  are better described by the R16 law. In any case, no consensus exists so far as to the most appropriate UV attenuation law, or even as to the number of dust components that should be considered (Charlot & Fall 2000), and significant variations may exist between individual galaxies, different galaxy types, and so on.

We primarily use the R16 as the default here because it has been observationally constrained down to the Lyman limit, and thus avoids the need for extrapolations of the SMC law to the spectral range covered by the HST/COS spectra used here. Furthermore the R16 law is often used for high-redshift studies (e.g., Reddy et al. 2016b; Steidel et al. 2018), and may thus simplify comparisons.

(2) *Assumptions on the gas and dust geometry.* The so-called picket-fence model (Sect. 3) proposes two simple ways to describe the relative distribution of dust with respect to the gas in the ISM. Either the dust is distributed in a homogeneous foreground layer (uniform dust screen, A), or it resides within the same clouds as the gas (clumpy geometry, B).

Both UV attenuation and the geometrical covering fraction depend on the adopted geometry, as amply discussed in Vasei et al. (2016), Reddy et al. (2016b), Gazagnes et al. (2018), Steidel et al. (2018). Overall, higher values of  $E_{B-V}$  and  $C_f$  are found in the clumpy geometry (geometry B, referred to as the holes scenario by Steidel et al. 2018), as seen for example in Gazagnes et al. (2018), Chisholm et al. (2018), Steidel et al. (2018). This is simply explained by the fact that, in this case, a fraction  $(1 - C_f)$  of UV continuum light is directly transmitted, that is, unaltered, to the observer. To fit the observed UV slope and the depth of the absorption lines, higher  $E_{B-V}$  and  $C_f$  are therefore needed.

Despite the differences for  $E_{B-V}$  and  $C_f$ , Chisholm et al. (2018) and Gazagnes et al. (2020) showed that, to first order, the determination of  $f_{\text{esc}}^{\text{abs}}$  from the UV absorption lines is independent of the assumed dust distribution between these two picket-fence geometries. However, for reasons that are not completely clear, Steidel et al. (2018) obtain somewhat lower average absolute LyC escape fractions when assuming the “holes” geometry.

Our analysis of the HI and low-ionization UV absorption lines in Sect. 5.2 clearly shows that the observed behavior favors a clumpy dust geometry, i.e., geometry B, as also suggested by earlier studies (see e.g., Du et al. 2021). However, for our spectral fits, we assume the foreground dust-screen scenario (A), which is not the favored scenario. This assumption is primarily driven by practical reasons and simplicity, because the fits of the stellar continuum (where ISM absorption lines are masked) can then be done independently of  $C_f$ , and the absorption lines can subsequently be analyzed. Indeed, in this geometry, the covering fraction is simply related to the residual flux of the (saturated) lines via  $R = 1 - C_f$  (Gazagnes et al. 2018).

As long as the two assumptions yield almost identical LyC escape fractions, this apparent inconsistency is not relevant. However, it must be borne in mind that values such as the UV attenuation (quantified by  $E_{B-V}$ ) and the covering fraction depend on geometrical assumptions, and they should be handled with care in comparisons between different studies.

(3) *Validity of the picket-fence model.* In this paper, we make use of the picket-fence model to derive simple analytical solutions of the radiative transfer equations, which can be used to fit the observed galaxy spectra using SEDs of integrated



stellar populations and ISM absorption lines (cf., Sect. 3, Vasei et al. 2016; Reddy et al. 2016b). With the assumptions that the “channels” between the pickets have a negligible column density and the absorption lines are saturated in the ISM clouds (pickets), this model then allows one to determine the geometric covering fraction of the absorbing gas from the residual flux of these lines.

Although this simple model works well to describe the observations and predict the average LyC continuum escape fractions, as shown in this paper and earlier studies (Gazagnes et al. 2018; Steidel et al. 2018; Chisholm et al. 2018), we remind the reader that this model is a drastic over-simplification of the complexity observed in resolved nearby galaxies and predicted by state-of-the-art high-resolution galaxy simulations. In realistic situations, the geometry may differ from the simple assumptions of a single UV-emitting source behind a picket-fence ISM and a uniform dust screen made here. Different star-forming regions, whose emission contributes to the integrated spectra, may be “covered” by varying column densities of gas and dust, and each of those may also show substructure. Furthermore, the kinematics of the gas may be complex. How important these effects are for the diagnostics used in the present paper is nevertheless impossible to estimate. In addition, some galaxies, for example, the most compact ones, may also be better approximated by a picket-fence model than others.

The recent work of Mauerhofer et al. (2021), who simulate a  $z \sim 3$  low-mass galaxy at high resolution, and performed mock observations of UV absorption lines, may provide some guidance towards answering these questions. Indeed, these simulations illustrate the nontrivial definition of the covering fraction and the complex relations between this quantity, the residual line fluxes, and the escape fraction of LyC photons. Despite this, the simulations reveal correlations between residual fluxes and LyC escape, also with significant scatter, similar to those found in our analysis. In any case, going beyond the analytical picket-fence model for the analysis of spectroscopic observations will require the development of new methods and data of higher quality (in S/N and resolution).

(4) *Definition and meaning of the LyC escape fraction.* As described in Sect. 3, the LyC escape fraction  $f_{\text{esc}}^{\text{abs}}$  derived in this paper relies on comparing the observed LyC flux in a limited window shortward of the Lyman edge to the expected intrinsic LyC flux. In our analysis, we measure the LyC typically from 850 to 900 Å in the COS spectra (Flury et al. 2022a), and the intrinsic LyC flux is predicted by the fits to the nonionizing UV spectrum. This method is the most basic and most commonly used method for “direct” determination of the LyC escape fraction, that is, when a direct measurement of the LyC flux is available. It has been used by almost all existing studies of this kind, allowing for meaningful comparisons of  $f_{\text{esc}}^{\text{abs}}$  at all redshifts (Izotov et al. 2016a; Steidel et al. 2018).

The method can be applied to both spectroscopic observations and to imaging where two or more filters are used to probe the LyC flux and the nonionizing UV. When spectra are used, the LyC flux has so far always been measured over a small wavelength range just shortward of 912 Å. Imaging may probe a wider spectral range in the LyC, depending on the filter widths used; for HST/WFC3 imaging of the strong LyC leaker Ion2, e.g., the LyC is probed over  $\sim 730\text{--}890$  Å restframe (Vanzella et al. 2016).

One question that presents itself in regard to this method is to what extent a measure of such a monochromatic escape fraction reflects the total escape fraction of all H-ionizing photons, including those at shorter wavelengths. For example, in

their numerical simulations of escaping LyC radiation, Gnedin et al. (2008) find that the total escape fraction of ionizing photons below 912 Å is approximately a factor of 1.25 higher than the fraction escaping close to the Lyman edge. Furthermore, McCandliss & O’Meara (2017) show the relation between these two escape fractions for slabs of variable H I column density and fixed ionization fraction, finding fairly large enhancements (a factor of 2 and higher for “Lyman edge escape fractions”  $< 0.32$ , comparable to our sample). However, such a simplified geometry and the transmission curves computed in their work cannot be used to predict the resulting spectrum of density-bounded H II regions (cf., below, and Inoue 2010).

Generally, if both (i) the escape fraction of LyC photons is wavelength-independent and (ii) the only emission process at the measured spectral range is stellar emission, the ratio of the observed to intrinsic stellar SED will provide a correct measure of the total escape fraction of LyC photons. Therefore, both of these conditions need to be examined. In fact, the escape fraction is to first order wavelength-independent if the ISM can be described by a picket-fence model, where the “leaking” regions have a negligible H I column density and the total escape fraction is governed by the geometrical covering factor (see also Inoue 2010). The detailed analyses in this paper and earlier studies of LyC-emitting galaxies (Gazagnes et al. 2018; Steidel et al. 2018) have shown clear evidence in favor of the picket-fence model, although a low residual column density in the “escape channels” cannot be excluded. The second condition (ii) might not always be validated, because free-bound emission of hydrogen in ionized nebulae emits shortward of the Lyman edge, and may therefore contribute a non-negligible additional emission source close to the Lyman edge, as discussed for example by Inoue (2010). Theoretically, the strength of this emission primarily depends on the mean energy of the ionizing photons (i.e., on the detailed SED) and on the electron temperature of the ionized gas, and might lead to a “Lyman bump” in extreme cases (Inoue 2010)<sup>11</sup>. To the best of our knowledge, this emission has so far not been identified, although it is expected. If present, our monochromatic determinations of  $f_{\text{esc}}^{\text{abs}}$  could overestimate the true total escape fraction.

A second, independent method to determine the LyC escape fraction has also been used for the LzLCS survey and the literature samples; see Flury et al. (2022a,b). This method uses the de-reddened H $\beta$  flux, which is proportional to the total number of ionizing photons absorbed (i.e., nonescaping) in the galaxy, and compares it to the observed LyC flux, using theoretical SEDs to relate the total ionizing photon flux to the monochromatic flux shortward of the Lyman edge. As discussed by Izotov et al. (2018b), this method therefore measures the total ionizing photon escape fraction. However, if strong free-bound emission is present, this method might also overestimate  $f_{\text{esc}}^{\text{abs}}$ . Comparing this method to the “monochromatic” method, Izotov et al. (2018b) find good agreement overall and no systematic differences (cf., Izotov et al. 2018a,b, 2021). Similarly, for the larger sample analyzed here, we do not find systematic differences between the two methods, although the scatter is fairly large (Flury et al. 2022a), and possibly driven by uncertainties related to the reddening corrections (both in the optical and UV). Indeed, as shown above (see Fig. 4), the  $f_{\text{esc}}^{\text{abs}}$  values derived here solely from the UV spectrum depend quite strongly on the

<sup>11</sup> We note that free-bound emission could also be expected if the escape channels in the ISM contain a low H I column density, as shown by the CLOUDY simulations of Inoue (2010).

assumed attenuation law, making this probably the largest source of uncertainty of our analysis.

Finally, as it is known that theoretical SEDs of star-forming galaxies have a significant lack of ionizing photons at energies above 54 eV, which are required to explain nebular lines of He II frequently observed in low-metallicity SF galaxies (see e.g., Stasińska et al. 2015; Schaerer et al. 2019), one might wonder whether or not these photons could alter our analysis and play a role in the ionizing photon budget of hydrogen. The answer appears to be no, primarily because the number of photons emitted above the He II ionization potential ( $\lambda < 228 \text{ \AA}$ ) is small compared to those ionizing hydrogen. Indeed, their relative ratio,  $Q_2/Q_0$ , can be determined empirically from the observed He II and H recombination lines and is typically  $\sim(0.2-2) \times 10^{-2}$  in low- $z$  galaxies with nebular He II  $\lambda 4686$  detections (see Schaerer et al. 2019). Their contribution to the ionization of hydrogen is therefore negligible. From model SEDs, it is well known that the bulk of the H-ionizing photons are emitted relatively close to the Lyman edge; for example  $\sim 60\%-70\%$  of the photons in the LyC are emitted longward of the He I ionization potential (24.6 eV, or 504  $\text{\AA}$ ) at young ages (see e.g., Nakajima et al. 2018; Xiao et al. 2018). Although all photons with energies  $> 13.6 \text{ eV}$  can ionize hydrogen, the majority are emitted at relatively low energies. In any case, methods using H $\beta$  (or other H recombination lines) to determine the LyC escape fraction inherently probe the total H-ionizing photon budget.

## 8. Summary and conclusions

The Low-Redshift Lyman Continuum Survey (LzLCS, Flury et al. 2022a) recently observed 66 low- $z$  ( $z \sim 0.3$ ) star-forming galaxies with HST/COS to measure the LyC in a large sample of galaxies for the first time. We augmented the LzLCS data, adding 23 previous LyC observations taken with COS (Izotov et al. 2016a,b, 2018a,b, 2021, and Wang et al. 2019), and the available LyC, UV, and optical spectra analyzed in a homogeneous manner. The sample provides unique insight into the LyC emission, the LyC escape fraction, and many related properties; it covers galaxies with a wide range of physical properties (UV magnitude, metallicity, and others) allowing us to empirically study how the escape of LyC emission varies from galaxy-to-galaxy and, in particular, how this phenomenon is related to the ISM properties of these sources.

To study the ISM properties of these low-redshift galaxies, we analyzed the COS UV spectra in the framework of a picket-fence model with a uniform dust screen. First, we fit the UV stellar continuum (see Sect. 3.1) to constrain the massive star content of these galaxies. From our nonparametric fits, we derived direct properties such as the stellar ages and metallicities, UV dust attenuations ( $E_{B-V}$ ), intrinsic UV magnitudes ( $M_{1500}^{\text{int}}$ ), and absolute LyC escape fractions ( $f_{\text{esc}}^{\text{abs}}$ ). The stellar population properties will be discussed in a separate paper (Chisholm et al., in prep.).

The derived LyC escape fractions range up to 60% (85%) for the LzLCS (published) sample. Approximately 45% of the sample is not detected in the LyC, with  $1\sigma$  upper limits of  $f_{\text{esc}}^{\text{abs}} < 0.01$ . The absolute escape fraction strongly depends on the dust-attenuation parameter ( $E_{B-V}$ ), with the strongest LyC emitters only found with low UV attenuation. We also study the impact of the assumed attenuation law on  $f_{\text{esc}}^{\text{abs}}$ : adopting the SMC extinction law instead of the attenuation law from Reddy et al. (2016a) leads to an increase in  $f_{\text{esc}}^{\text{abs}}$  by  $\sim 10\%-50\%$ , and the largest differences appear for galaxies with the lowest LyC escape fractions (Sect. 4).

With the stellar continuum fits in hand, we then measured the equivalent widths ( $W_\lambda$ ) and residual fluxes ( $R_\lambda$ ) of different interstellar absorption lines of HI (Lyman series) and metallic low ionization state (LIS) lines, such as Si II  $\lambda\lambda 1190, 1193$ , Si II  $\lambda 1260$ , O I  $\lambda 1302$ , and C II  $\lambda 1334$  (see Sect. 3.2). With the picket-fence model, the residual fluxes allow us to infer a geometrical covering fraction of the UV source by the gas. Our analysis of the absorption lines and other galaxy properties has led us to the following main conclusions:

- Lyman continuum emitters (leakers) have weak but likely saturated HI and LIS lines (Sect. 5.1), and changes in  $W_\lambda$  are mainly driven by the neutral gas covering fraction ( $C_f(\text{HI})$ ). Metallicity plays a secondary role (Sect. 5.2).
- We find that leakers have high HI and LIS residual fluxes ( $R_\lambda$ ) and low UV attenuations ( $A_{\text{UV}}$ ). This indicates that leakers tend to host a dust-poor ISM, with low HI column density channels intermixed between higher column density regions (Sect. 5.2). Our large sample confirms the finding that these channels enable the escape of ionizing photons, as shown in other recent studies (Reddy et al. 2016b; Gazagnes et al. 2018; Steidel et al. 2018).
- Leakers have large Ly $\alpha$  equivalent widths ( $W_{\text{Ly}\alpha}$ ), and this correlates strongly with the HI and LIS equivalent widths, showing that the absorption lines trace the same gas that regulates the radiative transfer of Ly $\alpha$  photons.  $W_{\text{Ly}\alpha}$  is therefore partly driven by the neutral gas covering fraction  $C_f(\text{HI})$  (Sect. 6.1), as also found in earlier studies (e.g., Gazagnes et al. 2020; Pahl et al. 2020).
- We find HI and LIS covering fractions varying from  $\sim 0.5$  to 1 and  $\sim 0.2$  to 1, respectively. As in Reddy et al. (2016b) and Gazagnes et al. (2018), the LIS covering fraction is on average smaller than that of HI, and the two quantities are strongly correlated. This probably indicates that HI and metals are not evenly distributed within the host galaxy, but that they trace each other (Sect. 6.2).
- The LyC photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) strongly correlates with both HI and LIS equivalent widths and residual fluxes of the lines (Sect. 6.3), and also with the UV attenuation (Sect. 6.4). We derive empirical calibrations to estimate  $f_{\text{esc}}^{\text{abs}}$  from UV absorption line measurements.
- Finally, we show that simultaneous UV absorption line and attenuation measurements can, in general, predict the escape fractions of galaxies (Sect. 7.1). We apply our methods to literature measurements of UV LIS lines in  $z \sim 4-6$  galaxies, estimating modestly low LyC escape fractions of  $f_{\text{esc}}^{\text{abs}} \lesssim 0.1$  for relatively bright,  $M_{1500}^{\text{obs}} \lesssim -21$ , high-redshift galaxies (Sect. 7.3).

In short, our detailed quantitative analysis of the UV spectra of 89 low- $z$  galaxies shows that LyC leakage is a complex phenomenon (see Sect. 7.2), which mainly depends on the gas distribution (covering fractions) and dust content (UV attenuation) of the host galaxy. We demonstrate that the HI Lyman series lines shortward of Ly $\alpha$  and metallic absorption lines can be used to statistically predict the escape fraction of ionizing photons, as proposed by Chisholm et al. (2018). Although several uncertainties and open questions remain, our study demonstrates how LIS UV absorption lines can be used to study and constrain the escape of ionizing photons. These diagnostics will be invaluable and key in upcoming studies of galaxies up to the Epoch of Reionization, with the advent of the James Webb Space Telescope (JWST) and other 30-meter class telescopes.

**Acknowledgements.** A.S.L. and D.S. acknowledge support from Swiss National Science Foundation. S.R.F. and A.E.J. acknowledge support from NASA through science HST-GO-15626 from the Space Telescope Science Institute. R.A. acknowledges support from ANID Fondecyt Regular 1202007. M.T. acknowledges support from the NWO grant 016.VIDI.189.162 (“ODIN”). This research is based on observations made with the NASA/ESA *Hubble* Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. These observations are associated with program GO 15626 (P.I. Jaskot). Additional work was obtained from the data archive at the Space Telescope Science Institute from HST proposals 13744, 14635, 15341, and 15639.

## References

- Akritas, M. G., & Siebert, J. 1996, *MNRAS*, **278**, 919
- Alavi, A., Colbert, J., Teplitz, H. I., et al. 2020, *ApJ*, **904**, 59
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, **93**, 5
- Bassett, R., Ryan-Weber, E. V., Cooke, J., et al. 2019, *MNRAS*, **483**, 5223
- Becker, R. H., Fan, X., White, R. L., et al. 2001, *AJ*, **122**, 2850
- Belokurov, V., Evans, N. W., Moiseev, A., et al. 2007, *ApJ*, **671**, L9
- Bergvall, N., Zackrisson, E., Andersson, B. G., et al. 2006, *A&A*, **448**, 513
- Bian, F., & Fan, X. 2020, *MNRAS*, **493**, L65
- Bian, F., Fan, X., McGreer, I., Cai, Z., & Jiang, L. 2017, *ApJ*, **837**, L12
- Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, *AJ*, **154**, 28
- Borthakur, S., Heckman, T., Tumlinson, J., et al. 2016, *ApJ*, **833**, 259
- Bouchet, P., Lequeux, J., Maurice, E., Prevot, L., & Prevot-Burnichon, M. L. 1985, *A&A*, **149**, 330
- Bouwens, R., González-López, J., Aravena, M., et al. 2020, *ApJ*, **902**, 112
- Bridge, C. R., Teplitz, H. I., Siana, B., et al. 2010, *ApJ*, **720**, 465
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Carr, C., Scarlata, C., Henry, A., & Panagia, N. 2021, *ApJ*, **906**, 104
- Cen, R. 2020, *ApJ*, **889**, L22
- Charlot, S., & Fall, S. M. 2000, *ApJ*, **539**, 718
- Chisholm, J., Tremonti, C. A., Leitherer, C., Chen, Y., & Wofford, A. 2016, *MNRAS*, **457**, 3133
- Chisholm, J., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, **605**, A67
- Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018, *A&A*, **616**, A30
- Chisholm, J., Rigby, J. R., Bayliss, M., et al. 2019, *ApJ*, **882**, 182
- Chisholm, J., Prochaska, J. X., Schaerer, D., Gazagnes, S., & Henry, A. 2020, *MNRAS*, **498**, 2554
- Clarke, C., & Oey, M. S. 2002, *MNRAS*, **337**, 1299
- Dayal, P., Volonteri, M., Choudhury, T. R., et al. 2020, *MNRAS*, **495**, 3065
- de Barros, S., Vanzella, E., Amorín, R., et al. 2016, *A&A*, **585**, A51
- Dessauges-Zavadsky, M., D’Odorico, S., Schaerer, D., et al. 2010, *A&A*, **510**, A26
- Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, *ApJ*, **828**, 71
- Draine, B. T. 2011, *Physics of the Interstellar and Intergalactic Medium* (Princeton: Princeton University Press)
- Du, X., Shapley, A. E., Reddy, N. A., et al. 2018, *ApJ*, **860**, 75
- Du, X., Shapley, A. E., Topping, M. W., et al. 2021, *ApJ*, **920**, 95
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, *PASA*, **34**, e058
- Erb, D. K. 2015, *Nature*, **523**, 169
- Faisst, A. L. 2016, *ApJ*, **829**, 99
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *Rev. Mex. Astron. Astrofis.*, **53**, 385
- Finkelstein, S. L., D’Aloisio, A., Paardekoooper, J.-P., et al. 2019, *ApJ*, **879**, 36
- Fitzpatrick, E. L. 1999, *PASP*, **111**, 63
- Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, *ApJ*, **878**, 87
- Flury, S. R., & Moran, E. C. 2020, *MNRAS*, **496**, 2191
- Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022a, *ApJS*, **260**, 1
- Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022b, *ApJ*, **930**, 126
- Fontanot, F., Cristiani, S., Pfrommer, C., Cupani, G., & Vanzella, E. 2014, *MNRAS*, **438**, 2097
- Gazagnes, S., Chisholm, J., Schaerer, D., et al. 2018, *A&A*, **616**, A29
- Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A&A*, **639**, A85
- Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *A&A*, **578**, A83
- Gnedin, N. Y., Kravtsov, A. V., & Chen, H. 2008, *ApJ*, **672**, 765
- Grazian, A., Giallongo, E., Gerbasi, R., et al. 2016, *A&A*, **585**, A48
- Grazian, A., Giallongo, E., Paris, D., et al. 2017, *A&A*, **602**, A18
- Grazian, A., Giallongo, E., Fiore, F., et al. 2020, *ApJ*, **897**, 94
- Green, G. M. 2018, *J. Open Source Softw.*, **3**, 695
- Grimes, J. P., Heckman, T., Aloisi, A., et al. 2009, *ApJS*, **181**, 272
- Gronke, M., Dijkstra, M., McCourt, M., & Oh, S. P. 2016, *ApJ*, **833**, L26
- Guaita, L., Pentericci, L., Grazian, A., et al. 2016, *A&A*, **587**, A133
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, **142**, 1633
- Guseva, N. G., Izotov, Y. I., Schaerer, D., et al. 2020, *MNRAS*, **497**, 4293
- Harikane, Y., Laporte, N., Ellis, R. S., & Matsuoka, Y. 2020, *ApJ*, **902**, 117
- Heckman, T. M., Sembach, K. R., Meurer, G. R., et al. 2001, *ApJ*, **558**, 56
- Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, *ApJ*, **730**, 5
- Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, *ApJ*, **809**, 19
- Henry, A., Berg, D. A., Scarlata, C., Verhamme, A., & Erb, D. 2018, *ApJ*, **855**, 96
- Hogarth, L., Amorín, R., Vílchez, J. M., et al. 2020, *MNRAS*, **494**, 3541
- Inoue, A. K. 2010, *MNRAS*, **401**, 1325
- Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, *MNRAS*, **442**, 1805
- Izotov, Y. I., & Thuan, T. X. 1999, *ApJ*, **511**, 639
- Izotov, Y. I., Thuan, T. X., & Lipovetsky, V. A. 1994, *ApJ*, **435**, 647
- 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., Thuan, T. X., & Guseva, N. G. 2017, *MNRAS*, **471**, 548
- 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., Schaerer, D., Worseck, G., et al. 2020, *MNRAS*, **491**, 468
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2021, *MNRAS*, **503**, 1734
- Japelj, J., Vanzella, E., Fontanot, F., et al. 2017, *MNRAS*, **468**, 389
- Jaskot, A. E., & Oey, M. S. 2013, *ApJ*, **766**, 91
- Jaskot, A. E., & Oey, M. S. 2014, *ApJ*, **791**, L19
- Jaskot, A. E., Dowd, T., Oey, M. S., Scarlata, C., & McKinney, J. 2019, *ApJ*, **885**, 96
- Ji, Z., Giavalisco, M., Vanzella, E., et al. 2020, *ApJ*, **888**, 109
- Jones, T., Stark, D. P., & Ellis, R. S. 2012, *ApJ*, **751**, 51
- Jones, T. A., Ellis, R. S., Schenker, M. A., & Stark, D. P. 2013, *ApJ*, **779**, 52
- Kakiichi, K., & Gronke, M. 2021, *ApJ*, **908**, 30
- Kelly, B. C. 2007, *ApJ*, **665**, 1489
- Kennicutt, R. C., & De Los Reyes, M. A. C. 2021, *ApJ*, **908**, 61
- Kimm, T., Blaizot, J., Garel, T., et al. 2019, *MNRAS*, **486**, 2215
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Kulkarni, G., Choudhury, T. R., Puchwein, E., & Haehnelt, M. G. 2017, *MNRAS*, **469**, 4283
- Lebouteiller, V., Kunth, D., Lequeux, J., et al. 2006, *A&A*, **459**, 161
- Lecavelier des Etangs, A., Désert, J. M., Kunth, D., et al. 2004, *A&A*, **413**, 131
- Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., & Zitrin, A. 2016, *ApJ*, **831**, 152
- Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, *A&A*, **553**, A106
- Leitherer, C., Ortiz Otálvaro, P. A., Bresolin, F., et al. 2010, *ApJS*, **189**, 309
- Leitherer, C., Schaerer, D., Goldader, J., et al. 2011a, *Astrophysics Source Code Library* [record ascl:1104.003]
- Leitherer, C., Tremonti, C. A., Heckman, T. M., & Calzetti, D. 2011b, *AJ*, **141**, 37
- Leitherer, C., Ekström, S., Meynet, G., et al. 2014, *ApJS*, **212**, 14
- Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, *ApJ*, **823**, 64
- Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, *MNRAS*, **449**, 4204
- Ma, X., Quataert, E., Wetzel, A., et al. 2020, *MNRAS*, **498**, 2001
- Madau, P., & Haardt, F. 2015, *ApJ*, **813**, L8
- Makan, K., Worseck, G., Davies, F. B., et al. 2021, *ApJ*, **912**, 38
- Malkan, M. A., & Malkan, B. K. 2021, *ApJ*, **909**, 92
- Marchi, F., Pentericci, L., Guaita, L., et al. 2017, *A&A*, **601**, A73
- Marchi, F., Pentericci, L., Guaita, L., et al. 2018, *A&A*, **614**, A11
- Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, *ApJ*, **869**, 150
- Mauerhofer, V., Verhamme, A., Blaizot, J., et al. 2021, *A&A*, **646**, A80
- McCandliss, S. R., & O’Meara, J. M. 2017, *ApJ*, **845**, 111
- McKinney, J. H., Jaskot, A. E., Oey, M. S., et al. 2019, *ApJ*, **874**, 52
- Meštrić, U., Ryan-Weber, E. V., Cooke, J., et al. 2021, *MNRAS*, **508**, 4443
- Meynet, G., Maeder, A., Schaller, G., Schaerer, D., & Charbonnel, C. 1994, *A&AS*, **103**, 97
- Micheva, G., Iwata, I., & Inoue, A. K. 2017a, *MNRAS*, **465**, 302
- Micheva, G., Iwata, I., Inoue, A. K., et al. 2017b, *MNRAS*, **465**, 316
- Morrissey, P., Conrow, T., Barlow, T. A., et al. 2007, *ApJS*, **173**, 682
- Mostardi, R. E., Shapley, A. E., Steidel, C. C., et al. 2015, *ApJ*, **810**, 107
- Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, *MNRAS*, **454**, 2691
- Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. 2018, *MNRAS*, **478**, 791
- Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, *ApJ*, **892**, 109
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, **442**, 900
- Nakajima, K., Ellis, R. S., Iwata, I., et al. 2016, *ApJ*, **831**, L9
- Nakajima, K., Schaerer, D., Le Fèvre, O., et al. 2018, *A&A*, **612**, A94
- Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, **889**, 161



- Newville, M., Stensitzki, T., Allen, D. B., et al. 2016, *Astrophysics Source Code Library* [record ascl:1606.014]
- Orlítóv, I., Verhamme, A., Henry, A., et al. 2018, *A&A*, **616**, A60
- stlin, G., Rivera-Thorsen, T. E., Menacho, V., et al. 2021, *ApJ*, **912**, 155
- Pahl, A. J., Shapley, A., Faisst, A. L., et al. 2020, *MNRAS*, **493**, 3194
- Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., & Reddy, N. A. 2021, *MNRAS*, **505**, 2447
- Pauldrach, A. W. A., Hoffmann, T. L., & Lennon, M. 2001, *A&A*, **375**, 161
- Planck Collaboration XLVII. 2016, *A&A*, **596**, A108
- Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B. 1984, *A&A*, **132**, 389
- Puschnig, J., Hayes, M., stlin, G., et al. 2017, *MNRAS*, **469**, 3252
- Ramambason, L., Schaerer, D., Stasinska, G., et al. 2020, *A&A*, **644**, A21
- Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, *ApJ*, **806**, 259
- Reddy, N. A., Steidel, C. C., Pettini, M., & Bogosavljevi, M. 2016a, *ApJ*, **828**, 107
- Reddy, N. A., Steidel, C. C., Pettini, M., Bogosavljevi, M., & Shapley, A. E. 2016b, *ApJ*, **828**, 108
- Reddy, N. A., Topping, M. W., Shapley, A. E., et al. 2022, *ApJ*, **926**, 31
- Rivera-Thorsen, T. E., Hayes, M., stlin, G., et al. 2015, *ApJ*, **805**, 14
- Rivera-Thorsen, T. E., Dahle, H., Gronke, M., et al. 2017, *A&A*, **608**, L4
- Rivera-Thorsen, T. E., Dahle, H., Chisholm, J., et al. 2019, *Science*, **366**, 738
- Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *ApJ*, **768**, 71
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, **802**, L19
- Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, *MNRAS*, **479**, 994
- Rutkowski, M. J., Scarlata, C., Henry, A., et al. 2017, *ApJ*, **841**, L27
- Saha, K., Tandon, S. N., Simmonds, C., et al. 2020, *Nat. Astron.*, **4**, 1185
- Salim, S., Boquien, M., & Lee, J. C. 2018, *ApJ*, **859**, 11
- Scarlata, C., & Panagia, N. 2015, *ApJ*, **801**, 43
- Scarlata, C., Colbert, J., Teplitz, H. I., et al. 2009, *ApJ*, **704**, L98
- Schaerer, D., de Barros, S., & Sklias, P. 2013, *A&A*, **549**, A4
- Schaerer, D., Izotov, Y. I., Verhamme, A., et al. 2016, *A&A*, **591**, L8
- Schaerer, D., Fragos, T., & Izotov, Y. I. 2019, *A&A*, **622**, L10
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, *ApJ*, **588**, 65
- Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, *ApJ*, **651**, 688
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJ*, **826**, L24
- Shivaei, I., Reddy, N., Rieke, G., et al. 2020, *ApJ*, **899**, 117
- Siana, B., Shapley, A. E., Kulas, K. R., et al. 2015, *ApJ*, **804**, 17
- Smith, A., Safrank-Shrader, C., Bromm, V., & Milosavljevi, M. 2015, *MNRAS*, **449**, 4336
- Smith, B. M., Windhorst, R. A., Cohen, S. H., et al. 2020, *ApJ*, **897**, 41
- Stasinska, G., Izotov, Y., Morisset, C., & Guseva, N. 2015, *A&A*, **576**, A83
- Steidel, C. C., Bogosavljevi, M., Shapley, A. E., et al. 2018, *ApJ*, **869**, 123
- Stevens, M. L., Shull, J. M., Danforth, C. W., & Tilton, E. M. 2014, *ApJ*, **794**, 75
- Sugahara, Y., Ouchi, M., Harikane, Y., et al. 2019, *ApJ*, **886**, 29
- Trainor, R. F., Steidel, C. C., Strom, A. L., & Rudie, G. C. 2015, *ApJ*, **809**, 89
- Trainor, R. F., Strom, A. L., Steidel, C. C., et al. 2019, *ApJ*, **887**, 85
- Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, *MNRAS*, **470**, 224
- Vanzella, E., Siana, B., Cristiani, S., & Nonino, M. 2010, *MNRAS*, **404**, 1672
- Vanzella, E., Guo, Y., Giavalisco, M., et al. 2012, *ApJ*, **751**, 70
- Vanzella, E., de Barros, S., Castellano, M., et al. 2015, *A&A*, **576**, A116
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, **825**, 41
- Vanzella, E., Nonino, M., Cupani, G., et al. 2018, *MNRAS*, **476**, L15
- Vanzella, E., Caminha, G. B., Calura, F., et al. 2020, *MNRAS*, **491**, 1093
- Vasei, K., Siana, B., Shapley, A. E., et al. 2016, *ApJ*, **831**, 38
- Verhamme, A., Orlitov, I., Schaerer, D., & Hayes, M. 2015, *A&A*, **578**, A7
- Verhamme, A., Orlitov, I., Schaerer, D., et al. 2017, *A&A*, **597**, A13
- Wang, B., Heckman, T. M., Leitherer, C., et al. 2019, *ApJ*, **885**, 57
- Wang, B., Heckman, T. M., Amorn, R., et al. 2021, *ApJ*, **916**, 3
- Xiao, L., Stanway, E. R., & Eldridge, J. J. 2018, *MNRAS*, **477**, 904
- Zackrisson, E., Inoue, A. K., & Jensen, H. 2013, *ApJ*, **777**, 39

<sup>1</sup> Department of Astronomy, University of Geneva, 51 Chemin Pegasi, 1290 Versoix, Switzerland

e-mail: alberto.saldanalopez@unige.ch

<sup>2</sup> CNRS, IRAP, 14 Avenue E. Belin, 31400 Toulouse, France

<sup>3</sup> Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA

<sup>4</sup> Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA

<sup>5</sup> Astronomy Department, Williams College, Williamstown, MA 01267, USA

<sup>6</sup> Institut fur Physik und Astronomie, Universitat Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany

<sup>7</sup> Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

<sup>8</sup> Instituto de Investigacion Multidisciplinar en Ciencia y Tecnologa, Departamento de Fısica y Astronoma, Universidad de La Serena, Avda. Juan Cisternas 1200, La Serena, Chile

<sup>9</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>10</sup> INAF – Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio, 5, 35122 Padova, Italy

<sup>11</sup> Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA

<sup>12</sup> Department of Astronomy, Oskar Klein Centre; Stockholm University, 106 91 Stockholm, Sweden

<sup>13</sup> University of Michigan, Department of Astronomy, 323 West Hall, 1085 S. University Ave, Ann Arbor, MI 48109, USA

<sup>14</sup> INAF – Osservatorio Astronomico di Roma, Via Frascati 33, 00078, Monteporzio Catone, Italy

<sup>15</sup> Astronomy Department, University of Virginia, PO Box 400325, Charlottesville, VA 22904-4325, USA

<sup>16</sup> INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy

## Appendix A: Fits and absorption lines results

Summary tables with the main products from our UV–SED fits (+errors) and the absorption line measurements (+errors) are shown in this appendix for both the LzLCS and literature samples. Tables A.1 and A.3 include the  $E_{B-V}$  average stellar age and metallicities, inferred  $f_{esc}^{abs}$ , and intrinsic (dust-

free) and observed UV absolute magnitudes (in AB system). Tables A.2 and A.4 include the average HI and LIS equivalent widths, residual fluxes, and derived  $f_{esc}^{abs, HI}$  and  $f_{esc}^{abs, LIS}$ . Sources are identified and sorted by its coordinates, and the strong/weak/nonleaker categories are differentiated according to the LzLCS criteria.

**Table A.1.** UV–SED fits results for the LzLCS sample. Sources are sorted by Right-Ascension coordinate.

object	LyC type	$z$	$E_{B-V}$ (mag.)	Age (Myr)	$Z (Z_{\odot})$	$f_{esc}^{abs}$	$M_{1500}^{int.}$ (AB)	$M_{1500}^{obs.}$ (AB)
J003601+003307	non-leaker	0.3479	$0.013 \pm 0.033$	$3.61 \pm 1.40$	$0.14 \pm 0.08$	$\leq 0.029$	$-18.64 \pm 0.37$	$-18.53 \pm 0.10$
J004743+015440	weak	0.3535	$0.134 \pm 0.041$	$1.17 \pm 1.12$	$0.24 \pm 0.09$	$0.013_{-0.003}^{+0.021}$	$-21.90 \pm 0.45$	$-20.73 \pm 0.10$
J011309+000223	weak	0.3062	$0.136 \pm 0.049$	$3.97 \pm 1.30$	$0.50 \pm 0.19$	$0.022_{-0.012}^{+0.016}$	$-21.46 \pm 0.54$	$-20.27 \pm 0.12$
J012217+052044	weak	0.3656	$0.075 \pm 0.040$	$3.79 \pm 1.73$	$0.05 \pm 0.05$	$0.038_{-0.016}^{+0.046}$	$-20.79 \pm 0.44$	$-20.14 \pm 0.10$
J012910+145935	non-leaker	0.2800	$0.158 \pm 0.036$	$6.32 \pm 1.26$	$0.20 \pm 0.12$	$\leq 0.007$	$-21.82 \pm 0.40$	$-20.44 \pm 0.09$
J072326+414608	non-leaker	0.2969	$0.184 \pm 0.034$	$3.19 \pm 1.03$	$0.22 \pm 0.10$	$\leq 0.004$	$-21.51 \pm 0.37$	$-19.91 \pm 0.08$
J080425+472607	strong	0.3565	$0.062 \pm 0.057$	$6.85 \pm 1.93$	$0.10 \pm 0.11$	$0.584_{-0.370}^{+0.416}$	$-19.00 \pm 0.64$	$-18.46 \pm 0.15$
J081112+414146	weak	0.3331	$0.182 \pm 0.022$	$5.27 \pm 0.85$	$0.08 \pm 0.05$	$0.020_{-0.005}^{+0.012}$	$-21.11 \pm 0.23$	$-19.52 \pm 0.05$
J081409+211459	non-leaker	0.2272	$0.259 \pm 0.016$	$6.39 \pm 0.75$	$0.23 \pm 0.06$	$\leq 0.007$	$-23.19 \pm 0.17$	$-20.93 \pm 0.04$
J082652+182052	non-leaker	0.2972	$0.191 \pm 0.049$	$7.12 \pm 1.37$	$0.16 \pm 0.07$	$\leq 0.009$	$-20.79 \pm 0.54$	$-19.12 \pm 0.12$
J083440+480541	non-leaker	0.3426	$0.125 \pm 0.034$	$7.35 \pm 1.43$	$0.05 \pm 0.04$	$\leq 0.016$	$-21.03 \pm 0.37$	$-19.94 \pm 0.08$
J090918+392925	strong	0.2816	$0.037 \pm 0.045$	$6.39 \pm 1.88$	$0.05 \pm 0.07$	$0.491_{-0.230}^{+0.417}$	$-19.06 \pm 0.50$	$-18.73 \pm 0.11$
J091113+183108	weak	0.2622	$0.238 \pm 0.025$	$3.41 \pm 1.28$	$0.30 \pm 0.11$	$0.023_{-0.007}^{+0.018}$	$-22.63 \pm 0.25$	$-20.56 \pm 0.05$
J091207+523960	non-leaker	0.2470	$0.197 \pm 0.035$	$2.26 \pm 0.85$	$0.20 \pm 0.11$	$\leq 0.008$	$-20.68 \pm 0.37$	$-18.96 \pm 0.08$
J091208+505009	non-leaker	0.3276	$0.158 \pm 0.030$	$1.66 \pm 1.16$	$0.39 \pm 0.15$	$\leq 0.003$	$-21.37 \pm 0.32$	$-20.00 \pm 0.07$
J091703+315221	strong	0.3004	$0.082 \pm 0.025$	$4.81 \pm 1.07$	$0.39 \pm 0.10$	$0.161_{-0.055}^{+0.073}$	$-21.19 \pm 0.26$	$-20.52 \pm 0.06$
J092552+395714	non-leaker	0.3142	$0.176 \pm 0.072$	$2.33 \pm 2.14$	$0.19 \pm 0.11$	$\leq 0.004$	$-20.94 \pm 0.78$	$-19.40 \pm 0.17$
J093355+510925	strong	0.2913	$0.088 \pm 0.032$	$6.74 \pm 0.92$	$0.40 \pm 0.11$	$0.266_{-0.110}^{+0.106}$	$-19.63 \pm 0.35$	$-18.86 \pm 0.08$
J094001+593244	non-leaker	0.3716	$0.241 \pm 0.048$	$3.22 \pm 1.05$	$0.39 \pm 0.17$	$\leq 0.002$	$-22.44 \pm 0.54$	$-20.34 \pm 0.13$
J095236+405249	weak	0.3187	$0.040 \pm 0.021$	$5.45 \pm 0.72$	$0.21 \pm 0.09$	$0.042_{-0.013}^{+0.021}$	$-20.40 \pm 0.23$	$-20.06 \pm 0.05$
J095700+235709	non-leaker	0.2444	$0.373 \pm 0.023$	$2.91 \pm 0.67$	$0.18 \pm 0.06$	$\leq 0.001$	$-23.96 \pm 0.24$	$-20.70 \pm 0.05$
J095838+202508	weak	0.3017	$0.153 \pm 0.058$	$5.58 \pm 1.43$	$0.36 \pm 0.15$	$0.019_{-0.011}^{+0.028}$	$-20.58 \pm 0.65$	$-19.24 \pm 0.15$
J101401+523251	non-leaker	0.3549	$0.272 \pm 0.088$	$3.43 \pm 1.79$	$0.72 \pm 0.19$	$\leq 0.002$	$-21.99 \pm 1.00$	$-19.62 \pm 0.24$
J102615+633308	weak	0.3338	$0.139 \pm 0.036$	$4.76 \pm 1.23$	$0.49 \pm 0.15$	$0.015_{-0.008}^{+0.014}$	$-21.02 \pm 0.40$	$-19.81 \pm 0.09$
J103344+635317	strong	0.3467	$0.068 \pm 0.026$	$5.58 \pm 1.06$	$0.47 \pm 0.12$	$0.305_{-0.116}^{+0.154}$	$-21.59 \pm 0.28$	$-21.00 \pm 0.06$
J103816+452718	weak	0.3258	$0.206 \pm 0.019$	$5.24 \pm 0.84$	$0.20 \pm 0.06$	$0.007_{-0.002}^{+0.003}$	$-23.31 \pm 0.21$	$-21.51 \pm 0.04$
J105117+474357	non-leaker	0.3354	$0.451 \pm 0.080$	$4.56 \pm 1.69$	$0.19 \pm 0.17$	$\leq 0.000$	$-24.53 \pm 0.93$	$-20.59 \pm 0.24$
J105331+523753	weak	0.2526	$0.147 \pm 0.014$	$5.92 \pm 0.79$	$0.31 \pm 0.07$	$0.012_{-0.004}^{+0.006}$	$-22.62 \pm 0.15$	$-21.33 \pm 0.03$
J105525+485724	non-leaker	0.3801	$0.366 \pm 0.057$	$7.60 \pm 1.56$	$0.23 \pm 0.15$	$\leq 0.001$	$-23.55 \pm 0.66$	$-20.36 \pm 0.16$
J110452+475204	non-leaker	0.2801	$0.115 \pm 0.024$	$6.50 \pm 0.90$	$0.05 \pm 0.02$	$\leq 0.011$	$-20.74 \pm 0.27$	$-19.73 \pm 0.06$
J112224+412052	weak	0.3048	$0.206 \pm 0.087$	$4.14 \pm 1.43$	$0.45 \pm 0.22$	$0.026_{-0.018}^{+0.056}$	$-20.82 \pm 0.98$	$-19.03 \pm 0.23$
J112848+524509	weak	0.3644	$0.248 \pm 0.035$	$6.35 \pm 1.30$	$0.13 \pm 0.06$	$0.005_{-0.003}^{+0.005}$	$-22.61 \pm 0.39$	$-20.45 \pm 0.09$
J112933+493525	non-leaker	0.3451	$0.037 \pm 0.037$	$5.28 \pm 1.27$	$0.42 \pm 0.14$	$\leq 0.021$	$-19.89 \pm 0.42$	$-19.57 \pm 0.10$
J113304+651341	weak	0.2414	$0.137 \pm 0.023$	$3.83 \pm 0.98$	$0.05 \pm 0.05$	$0.022_{-0.009}^{+0.022}$	$-21.31 \pm 0.24$	$-20.12 \pm 0.05$
J115855+312559	strong	0.2430	$0.178 \pm 0.018$	$3.56 \pm 0.80$	$0.17 \pm 0.06$	$0.066_{-0.015}^{+0.030}$	$-22.42 \pm 0.18$	$-20.87 \pm 0.04$
J115959+382422	weak	0.2679	$0.061 \pm 0.035$	$3.24 \pm 1.76$	$0.12 \pm 0.10$	$0.043_{-0.016}^{+0.067}$	$-19.19 \pm 0.38$	$-18.65 \pm 0.09$
J120934+305326	non-leaker	0.2193	$0.169 \pm 0.011$	$4.22 \pm 0.78$	$0.36 \pm 0.06$	$\leq 0.013$	$-22.05 \pm 0.12$	$-20.57 \pm 0.03$
J121915+453930	weak	0.3038	$0.197 \pm 0.042$	$5.82 \pm 1.19$	$0.21 \pm 0.13$	$0.013_{-0.005}^{+0.016}$	$-21.33 \pm 0.47$	$-19.61 \pm 0.11$
J123519+063556	strong	0.3327	$0.080 \pm 0.020$	$6.16 \pm 0.81$	$0.27 \pm 0.08$	$0.052_{-0.014}^{+0.028}$	$-21.12 \pm 0.21$	$-20.42 \pm 0.05$
J124033+212716	non-leaker	0.2834	$0.106 \pm 0.017$	$6.65 \pm 0.92$	$0.12 \pm 0.05$	$\leq 0.011$	$-20.47 \pm 0.18$	$-19.54 \pm 0.04$
J124423+021540	non-leaker	0.2394	$0.141 \pm 0.030$	$5.45 \pm 1.32$	$0.24 \pm 0.10$	$\leq 0.015$	$-21.21 \pm 0.33$	$-19.98 \pm 0.07$
J124619+444902	weak	0.3222	$0.201 \pm 0.056$	$4.09 \pm 1.74$	$0.06 \pm 0.06$	$0.004_{-0.002}^{+0.009}$	$-22.09 \pm 0.62$	$-20.33 \pm 0.14$
J124835+123403	weak	0.2634	$0.099 \pm 0.026$	$7.40 \pm 1.14$	$0.28 \pm 0.11$	$0.047_{-0.026}^{+0.043}$	$-21.21 \pm 0.28$	$-20.34 \pm 0.06$
J124911+464535	non-leaker	0.3725	$0.272 \pm 0.130$	$2.42 \pm 1.92$	$0.20 \pm 0.21$	$\leq 0.003$	$-21.85 \pm 1.51$	$-19.47 \pm 0.38$
J125503+255609	non-leaker	0.3119	$0.080 \pm 0.026$	$6.41 \pm 1.10$	$0.11 \pm 0.07$	$\leq 0.009$	$-20.83 \pm 0.29$	$-20.14 \pm 0.06$
J125718+410221	non-leaker	0.3131	$0.095 \pm 0.041$	$6.57 \pm 1.41$	$0.24 \pm 0.11$	$\leq 0.014$	$-20.28 \pm 0.46$	$-19.45 \pm 0.11$

Table A.1. continued.

object	LyC type	$z$	$E_{B-V}$ (mag.)	Age (Myr)	$Z (Z_{\odot})$	$f_{\text{esc}}^{\text{abs}}$	$M_{1500}^{\text{int.}}$ (AB)	$M_{1500}^{\text{obs.}}$ (AB)
J130059+451057	non-leaker	0.3309	$0.439 \pm 0.141$	$2.46 \pm 2.17$	$0.41 \pm 0.24$	$\leq 0.001$	$-22.40 \pm 1.66$	$-18.58 \pm 0.44$
J130128+510451	weak	0.3478	$0.097 \pm 0.031$	$3.55 \pm 1.04$	$0.21 \pm 0.10$	$0.027^{+0.013}_{-0.014}$	$-21.88 \pm 0.33$	$-21.04 \pm 0.07$
J130559+422638	strong	0.3157	$0.014 \pm 0.019$	$4.39 \pm 1.12$	$0.33 \pm 0.13$	$0.178^{+0.078}_{-0.058}$	$-18.41 \pm 0.21$	$-18.29 \pm 0.06$
J131037+214817	weak	0.2831	$0.185 \pm 0.038$	$2.73 \pm 1.26$	$0.29 \pm 0.12$	$0.016^{+0.020}_{-0.006}$	$-21.71 \pm 0.41$	$-20.10 \pm 0.08$
J131419+104739	non-leaker	0.2961	$0.260 \pm 0.029$	$4.55 \pm 1.04$	$0.33 \pm 0.12$	$\leq 0.001$	$-23.23 \pm 0.31$	$-20.97 \pm 0.06$
J131904+510309	non-leaker	0.3176	$0.395 \pm 0.071$	$6.97 \pm 1.59$	$0.75 \pm 0.19$	$\leq 0.002$	$-22.53 \pm 0.80$	$-19.08 \pm 0.19$
J132633+421824	weak	0.3177	$0.055 \pm 0.080$	$5.00 \pm 1.43$	$0.05 \pm 0.12$	$0.118^{+0.137}_{-0.084}$	$-20.27 \pm 0.93$	$-19.79 \pm 0.23$
J132937+573315	non-leaker	0.3091	$0.194 \pm 0.040$	$1.87 \pm 1.81$	$0.21 \pm 0.16$	$\leq 0.001$	$-22.22 \pm 0.44$	$-20.53 \pm 0.10$
J134559+112848	non-leaker	0.2373	$0.302 \pm 0.022$	$3.36 \pm 0.41$	$0.22 \pm 0.09$	$\leq 0.002$	$-23.33 \pm 0.24$	$-20.69 \pm 0.05$
J135008+662438	non-leaker	0.3839	$0.402 \pm 0.073$	$3.08 \pm 1.48$	$0.16 \pm 0.09$	$\leq 0.001$	$-23.60 \pm 0.85$	$-20.10 \pm 0.22$
J140333+612115	weak	0.2816	$0.139 \pm 0.019$	$7.12 \pm 1.01$	$0.26 \pm 0.09$	$0.031^{+0.019}_{-0.014}$	$-20.22 \pm 0.21$	$-19.01 \pm 0.04$
J141013+434435	strong	0.3558	$0.072 \pm 0.056$	$3.33 \pm 1.24$	$0.12 \pm 0.11$	$0.105^{+0.071}_{-0.070}$	$-19.83 \pm 0.61$	$-19.19 \pm 0.14$
J144010+461937	weak	0.3008	$0.183 \pm 0.017$	$3.08 \pm 0.59$	$0.21 \pm 0.04$	$0.005^{+0.002}_{-0.002}$	$-22.77 \pm 0.19$	$-21.18 \pm 0.04$
J151707+370512	weak	0.3533	$0.228 \pm 0.088$	$3.44 \pm 1.52$	$0.37 \pm 0.23$	$0.041^{+0.080}_{-0.023}$	$-22.45 \pm 1.03$	$-20.46 \pm 0.27$
J154050+572442	non-leaker	0.2944	$0.218 \pm 0.020$	$5.62 \pm 0.85$	$0.16 \pm 0.07$	$\leq 0.001$	$-23.18 \pm 0.22$	$-21.28 \pm 0.04$
J155945+403325	non-leaker	0.2268	$0.135 \pm 0.016$	$5.72 \pm 0.47$	$0.33 \pm 0.09$	$\leq 0.025$	$-21.28 \pm 0.17$	$-20.11 \pm 0.03$
J160437+081959	non-leaker	0.3123	$0.246 \pm 0.066$	$7.18 \pm 1.54$	$0.25 \pm 0.13$	$\leq 0.007$	$-21.54 \pm 0.74$	$-19.39 \pm 0.17$
J164607+313054	weak	0.2906	$0.109 \pm 0.024$	$5.58 \pm 1.02$	$0.38 \pm 0.10$	$0.023^{+0.010}_{-0.010}$	$-20.98 \pm 0.26$	$-20.02 \pm 0.06$
J164849+495751	weak	0.3819	$0.166 \pm 0.075$	$4.16 \pm 1.28$	$0.10 \pm 0.13$	$0.026^{+0.051}_{-0.014}$	$-21.25 \pm 0.86$	$-19.80 \pm 0.21$
J172010+542133	weak	0.2938	$0.069 \pm 0.028$	$5.22 \pm 1.11$	$0.21 \pm 0.08$	$0.031^{+0.026}_{-0.014}$	$-20.12 \pm 0.31$	$-19.52 \pm 0.07$

**Notes.** Column 1: object name. Column 2: object type (strong, weak or nonleaker). Column 3: redshift. Column 4: UV dust-attenuation parameter ( $E_{B-V}$ , in mag.). Column 5: light-weighted stellar age (in Myr). Column 6: light-weighted stellar metallicity (in solar units,  $Z_{\odot}$ ). Column 6: LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ). Column 7: intrinsic (dust corrected,  $M_{1500}^{\text{int.}}$ ). Column 8: observed UV absolute magnitudes ( $M_{1500}^{\text{obs.}}$ ).

Table A.2. Absorption lines results for the LzLCS sample. Sources are sorted by Right-Ascension coordinate.

object	LyC type	$z$	$W_{\text{HI}}$ (Å)	$W_{\text{LIS}}$ (Å)	$R_{\text{HI}}$	$R_{\text{LIS}}$	$f_{\text{esc}}^{\text{abs, HI}}$	$f_{\text{esc}}^{\text{abs, LIS}}$
J003601+003307	non-leaker	0.3479	$2.51 \pm 0.23$	$0.97 \pm 0.88$	$0.17 \pm 0.07$	$0.28 \pm 0.29$	$0.153 \pm 0.099$	$0.044 \pm 0.151$
J004743+015440	weak	0.3535	$3.33 \pm 0.17$	$2.16 \pm 0.80$	$0.01 \pm 0.03$	$0.26 \pm 0.26$	$0.001 \pm 0.008$	$0.008 \pm 0.032$
J011309+000223	weak	0.3062	$2.42 \pm 0.34$	$0.58 \pm 0.70$	$0.19 \pm 0.11$	$0.66 \pm 0.24$	$0.055 \pm 0.050$	$0.048 \pm 0.055$
J012217+052044	weak	0.3656	$2.97 \pm 0.16$	$-0.10 \pm 0.76$	$0.23 \pm 0.05$	$0.54 \pm 0.24$	$0.093 \pm 0.059$	$0.078 \pm 0.098$
J012910+145935	non-leaker	0.2800	$2.40 \pm 0.30$	$1.87 \pm 0.49$	$0.14 \pm 0.09$	$0.36 \pm 0.15$	$0.015 \pm 0.016$	$0.001 \pm 0.014$
J072326+414608	non-leaker	0.2969	$1.83 \pm 0.27$	$0.82 \pm 0.48$	$0.13 \pm 0.06$	$0.56 \pm 0.16$	$0.014 \pm 0.010$	$0.021 \pm 0.017$
J080425+472607	strong	0.3565	$0.87 \pm 0.28$	$0.38 \pm 1.46$	$0.30 \pm 0.10$	$0.72 \pm 0.51$	$0.135 \pm 0.186$	$0.136 \pm 0.298$
J081112+414146	weak	0.3331	$1.77 \pm 0.16$	$0.65 \pm 0.38$	$0.27 \pm 0.05$	$0.77 \pm 0.15$	$0.033 \pm 0.011$	$0.039 \pm 0.017$
J081409+211459	non-leaker	0.2272	$4.20 \pm 0.11$	$1.28 \pm 0.24$	$0.21 \pm 0.04$	$0.76 \pm 0.10$	$0.010 \pm 0.003$	$0.015 \pm 0.004$
J082652+182052	non-leaker	0.2972	$3.09 \pm 0.41$	$2.71 \pm 0.61$	$0.00 \pm 0.07$	$0.27 \pm 0.16$	$0.003 \pm 0.010$	–
J083440+480541	non-leaker	0.3426	$2.79 \pm 0.21$	$1.79 \pm 0.58$	$0.09 \pm 0.05$	$0.38 \pm 0.19$	$0.019 \pm 0.015$	$0.017 \pm 0.031$
J090918+392925	strong	0.2816	$0.70 \pm 0.41$	$0.37 \pm 0.82$	$0.24 \pm 0.09$	$0.57 \pm 0.33$	$0.193 \pm 0.148$	$0.249 \pm 0.218$
J091113+183108	weak	0.2622	$1.75 \pm 0.24$	$0.97 \pm 0.40$	$0.22 \pm 0.07$	$0.55 \pm 0.13$	$0.013 \pm 0.006$	$0.014 \pm 0.006$
J091207+523960	non-leaker	0.2470	$2.78 \pm 0.25$	$1.65 \pm 0.45$	$0.08 \pm 0.08$	$0.55 \pm 0.16$	$0.007 \pm 0.010$	$0.017 \pm 0.016$
J091208+505009	non-leaker	0.3276	$3.09 \pm 0.17$	$0.42 \pm 0.53$	$0.13 \pm 0.04$	$0.78 \pm 0.19$	$0.019 \pm 0.011$	$0.053 \pm 0.033$
J091703+315221	strong	0.3004	$1.98 \pm 0.28$	$1.29 \pm 0.48$	$0.04 \pm 0.05$	$0.63 \pm 0.17$	$0.042 \pm 0.025$	$0.099 \pm 0.061$
J092552+395714	non-leaker	0.3142	$3.03 \pm 0.33$	$1.30 \pm 1.01$	$0.01 \pm 0.08$	$0.38 \pm 0.24$	$0.002 \pm 0.020$	$0.013 \pm 0.043$
J093355+510925	strong	0.2913	$2.08 \pm 0.29$	$-0.44 \pm 0.55$	$0.28 \pm 0.11$	$0.80 \pm 0.19$	$0.125 \pm 0.066$	$0.108 \pm 0.065$
J094001+593244	non-leaker	0.3716	$3.05 \pm 0.26$	$2.66 \pm 0.62$	$0.11 \pm 0.08$	$0.16 \pm 0.20$	$0.006 \pm 0.009$	–
J095236+405249	weak	0.3187	$3.30 \pm 0.13$	$0.42 \pm 0.42$	$0.17 \pm 0.04$	$0.67 \pm 0.14$	$0.122 \pm 0.042$	$0.168 \pm 0.076$
J095700+235709	non-leaker	0.2444	$3.82 \pm 0.18$	$3.13 \pm 0.30$	$0.14 \pm 0.07$	$0.28 \pm 0.08$	$0.002 \pm 0.001$	$0.000 \pm 0.001$
J095838+202508	weak	0.3017	$1.42 \pm 0.41$	$1.62 \pm 0.84$	$0.17 \pm 0.12$	$0.19 \pm 0.21$	$0.025 \pm 0.039$	$0.006 \pm 0.033$
J101401+523251	non-leaker	0.3549	$2.96 \pm 0.34$	$1.09 \pm 1.30$	–	$0.51 \pm 0.36$	–	$0.005 \pm 0.028$
J102615+633308	weak	0.3338	$1.94 \pm 0.26$	$1.25 \pm 0.54$	$0.30 \pm 0.08$	$0.63 \pm 0.21$	$0.074 \pm 0.042$	$0.044 \pm 0.042$
J103344+635317	strong	0.3467	$2.41 \pm 0.16$	$0.21 \pm 0.51$	$0.37 \pm 0.06$	$0.70 \pm 0.17$	$0.162 \pm 0.058$	$0.111 \pm 0.060$



Table A.2. continued.

object	LyC type	$z$	$W_{\text{HI}} (\text{\AA})$	$W_{\text{LIS}} (\text{\AA})$	$R_{\text{HI}}$	$R_{\text{LIS}}$	$f_{\text{esc}}^{\text{abs, HI}}$	$f_{\text{esc}}^{\text{abs, LIS}}$
J103816+452718	weak	0.3258	$2.87 \pm 0.14$	$1.20 \pm 0.34$	$0.23 \pm 0.05$	$0.57 \pm 0.11$	$0.020 \pm 0.006$	$0.018 \pm 0.008$
J105117+474357	non-leaker	0.3354	$3.20 \pm 0.42$	$3.84 \pm 0.82$	–	–	–	–
J105331+523753	weak	0.2526	$3.47 \pm 0.09$	$1.11 \pm 0.21$	$0.17 \pm 0.03$	$0.66 \pm 0.07$	$0.029 \pm 0.008$	$0.045 \pm 0.011$
J105525+485724	non-leaker	0.3801	$3.60 \pm 0.27$	$3.61 \pm 0.91$	$0.02 \pm 0.08$	$0.13 \pm 0.32$	$0.000 \pm 0.002$	–
J110452+475204	non-leaker	0.2801	$2.64 \pm 0.20$	$1.06 \pm 0.42$	$0.17 \pm 0.06$	$0.68 \pm 0.13$	$0.047 \pm 0.023$	$0.072 \pm 0.034$
J112224+412052	weak	0.3048	$0.95 \pm 0.58$	$0.07 \pm 1.33$	$0.17 \pm 0.17$	$0.50 \pm 0.35$	$0.010 \pm 0.048$	$0.007 \pm 0.058$
J112848+524509	weak	0.3644	$3.05 \pm 0.19$	$0.63 \pm 0.65$	$0.11 \pm 0.06$	$0.33 \pm 0.18$	$0.008 \pm 0.005$	$0.002 \pm 0.006$
J112933+493525	non-leaker	0.3451	$2.73 \pm 0.23$	$-0.37 \pm 0.86$	$0.12 \pm 0.06$	$0.49 \pm 0.22$	$0.070 \pm 0.060$	$0.082 \pm 0.124$
J113304+651341	weak	0.2414	$2.74 \pm 0.18$	$0.15 \pm 0.35$	$0.29 \pm 0.06$	$0.76 \pm 0.14$	$0.063 \pm 0.021$	$0.064 \pm 0.027$
J115855+312559	strong	0.2430	$1.96 \pm 0.16$	$0.88 \pm 0.26$	$0.40 \pm 0.06$	$0.79 \pm 0.11$	$0.057 \pm 0.014$	$0.037 \pm 0.011$
J115959+382422	weak	0.2679	$1.48 \pm 0.30$	$0.05 \pm 0.60$	$0.09 \pm 0.06$	$0.70 \pm 0.19$	$0.041 \pm 0.040$	$0.172 \pm 0.103$
J120934+305326	non-leaker	0.2193	$2.57 \pm 0.08$	$1.20 \pm 0.14$	$0.23 \pm 0.03$	$0.64 \pm 0.05$	$0.032 \pm 0.006$	$0.034 \pm 0.006$
J121915+453930	weak	0.3038	$3.06 \pm 0.28$	$1.89 \pm 0.58$	$0.21 \pm 0.11$	$0.42 \pm 0.20$	$0.018 \pm 0.018$	$0.009 \pm 0.017$
J123519+063556	strong	0.3327	$2.57 \pm 0.12$	$0.33 \pm 0.43$	$0.27 \pm 0.05$	$0.64 \pm 0.16$	$0.107 \pm 0.034$	$0.108 \pm 0.046$
J124033+212716	non-leaker	0.2834	$2.68 \pm 0.16$	$0.19 \pm 0.36$	$0.14 \pm 0.05$	$0.78 \pm 0.14$	$0.052 \pm 0.017$	$0.061 \pm 0.023$
J124423+021540	non-leaker	0.2394	$2.90 \pm 0.23$	$1.50 \pm 0.37$	$0.22 \pm 0.08$	$0.43 \pm 0.11$	$0.040 \pm 0.025$	$0.018 \pm 0.018$
J124619+444902	weak	0.3222	$2.33 \pm 0.24$	$1.12 \pm 0.81$	$0.02 \pm 0.04$	$0.55 \pm 0.23$	$0.001 \pm 0.005$	$0.016 \pm 0.025$
J124835+123403	weak	0.2634	$1.96 \pm 0.23$	$1.12 \pm 0.48$	$0.03 \pm 0.04$	$0.40 \pm 0.14$	$0.018 \pm 0.015$	$0.026 \pm 0.032$
J124911+464535	non-leaker	0.3725	$2.72 \pm 0.55$	$1.03 \pm 3.23$	–	–	–	–
J125503+255609	non-leaker	0.3119	$3.22 \pm 0.17$	$0.98 \pm 0.39$	$0.16 \pm 0.05$	$0.49 \pm 0.12$	$0.085 \pm 0.035$	$0.075 \pm 0.040$
J125718+410221	non-leaker	0.3131	$2.62 \pm 0.23$	$1.42 \pm 0.64$	$0.08 \pm 0.05$	$0.48 \pm 0.22$	$0.038 \pm 0.033$	$0.046 \pm 0.058$
J130059+451057	non-leaker	0.3309	$3.35 \pm 0.65$	$2.17 \pm 2.12$	–	$0.36 \pm 0.61$	–	$0.000 \pm 0.031$
J130128+510451	weak	0.3478	$3.06 \pm 0.17$	$0.10 \pm 0.58$	$0.15 \pm 0.05$	$0.50 \pm 0.17$	$0.047 \pm 0.028$	$0.051 \pm 0.044$
J130559+422638	strong	0.3157	$1.56 \pm 0.21$	$0.10 \pm 0.57$	$0.30 \pm 0.07$	$0.66 \pm 0.20$	$0.242 \pm 0.083$	$0.265 \pm 0.124$
J131037+214817	weak	0.2831	$2.24 \pm 0.29$	$1.02 \pm 0.61$	$0.32 \pm 0.10$	$0.44 \pm 0.18$	$0.034 \pm 0.022$	$0.011 \pm 0.018$
J131419+104739	non-leaker	0.2961	$2.88 \pm 0.25$	$2.85 \pm 0.38$	$0.16 \pm 0.06$	$0.35 \pm 0.11$	$0.007 \pm 0.004$	$0.002 \pm 0.004$
J131904+510309	non-leaker	0.3176	$4.64 \pm 0.40$	$3.12 \pm 0.75$	–	$0.09 \pm 0.22$	–	–
J132633+421824	weak	0.3177	$1.66 \pm 0.40$	$0.70 \pm 1.55$	$0.06 \pm 0.07$	$0.41 \pm 0.51$	$0.042 \pm 0.123$	$0.098 \pm 0.327$
J132937+573315	non-leaker	0.3091	$4.28 \pm 0.26$	$0.95 \pm 0.46$	$0.07 \pm 0.05$	$0.49 \pm 0.15$	$0.007 \pm 0.008$	$0.013 \pm 0.015$
J134559+112848	non-leaker	0.2373	$4.30 \pm 0.12$	$3.00 \pm 0.25$	$0.02 \pm 0.03$	$0.32 \pm 0.07$	$0.001 \pm 0.001$	$0.001 \pm 0.001$
J135008+662438	non-leaker	0.3839	$3.17 \pm 0.39$	$1.86 \pm 1.50$	$0.05 \pm 0.10$	$0.39 \pm 0.41$	$0.001 \pm 0.002$	$0.000 \pm 0.005$
J140333+612115	weak	0.2816	$1.48 \pm 0.22$	$0.43 \pm 0.37$	$0.40 \pm 0.09$	$0.57 \pm 0.11$	$0.083 \pm 0.025$	$0.038 \pm 0.018$
J141013+434435	strong	0.3558	$2.29 \pm 0.24$	$1.29 \pm 0.83$	$0.09 \pm 0.06$	$0.22 \pm 0.23$	$0.035 \pm 0.054$	–
J144010+461937	weak	0.3008	$4.76 \pm 0.13$	$1.10 \pm 0.39$	$0.16 \pm 0.04$	$0.49 \pm 0.11$	$0.023 \pm 0.007$	$0.016 \pm 0.010$
J151707+370512	weak	0.3533	$1.85 \pm 0.40$	$1.90 \pm 1.37$	$0.19 \pm 0.11$	$0.12 \pm 0.30$	$0.011 \pm 0.035$	–
J154050+572442	non-leaker	0.2944	$3.52 \pm 0.12$	$2.00 \pm 0.29$	$0.16 \pm 0.04$	$0.53 \pm 0.09$	$0.012 \pm 0.005$	$0.013 \pm 0.006$
J155945+403325	non-leaker	0.2268	$2.98 \pm 0.11$	$0.70 \pm 0.24$	$0.23 \pm 0.03$	$0.72 \pm 0.07$	$0.046 \pm 0.011$	$0.061 \pm 0.015$
J160437+081959	non-leaker	0.3123	$2.71 \pm 0.39$	$1.37 \pm 0.73$	$0.04 \pm 0.13$	$0.48 \pm 0.22$	$0.006 \pm 0.016$	$0.006 \pm 0.018$
J164607+313054	weak	0.2906	$2.19 \pm 0.19$	$-0.10 \pm 0.39$	$0.29 \pm 0.07$	$0.80 \pm 0.14$	$0.108 \pm 0.038$	$0.102 \pm 0.037$
J164849+495751	weak	0.3819	$2.66 \pm 0.31$	$1.62 \pm 1.33$	$0.25 \pm 0.09$	$0.29 \pm 0.48$	$0.039 \pm 0.077$	$0.025 \pm 0.072$
J172010+542133	weak	0.2938	$3.30 \pm 0.21$	$0.12 \pm 0.50$	$0.16 \pm 0.06$	$0.69 \pm 0.15$	$0.096 \pm 0.044$	$0.117 \pm 0.067$

**Notes.** Column 1: object name. Column 2: object type (strong, weak or non-leaker). Column 3: redshift. Columns 4 and 5: weighted-average HI and LIS equivalent widths ( $W_{\text{HI}}$ ,  $W_{\text{LIS}}$ ). Columns 6 and 7: weighted-average HI and LIS residual fluxes ( $C_f(\text{HI})$ ,  $C_f(\text{LIS})$ ). Columns 8 and 9: Lyman series and LIS derived absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs, HI}}$ ,  $f_{\text{esc}}^{\text{abs, LIS}}$ ).

**Table A.3.** UV–SED fits results for the literature sample<sup>1</sup>. Columns are the same as in Table A.1. Sources are sorted by year of publication and Right-Ascension coordinate.

object	LyC type	$z$	$E_{B-V}$ (mag.)	Age (Myr)	$Z (Z_{\odot})$	$f_{\text{esc}}^{\text{abs}}$	$M_{1500}^{\text{int.}}$ (AB)	$M_{1500}^{\text{obs.}}$ (AB)
J0925+1403	strong	0.3013	0.106 ± 0.024	4.32 ± 0.94	0.43 ± 0.09	0.092 <sup>+0.019</sup> <sub>-0.034</sub>	-20.74 ± 0.26	-19.81 ± 0.05
J1152+3400	strong	0.3419	0.095 ± 0.029	6.85 ± 1.28	0.11 ± 0.08	0.177 <sup>+0.195</sup> <sub>-0.063</sub>	-21.51 ± 0.32	-20.68 ± 0.07
J1333+6246	strong	0.3181	0.097 ± 0.044	7.00 ± 1.17	0.10 ± 0.04	0.128 <sup>+0.126</sup> <sub>-0.062</sub>	-20.60 ± 0.49	-19.75 ± 0.11
J1442-0209	strong	0.2937	0.091 ± 0.028	5.79 ± 1.08	0.16 ± 0.04	0.120 <sup>+0.064</sup> <sub>-0.050</sub>	-20.94 ± 0.30	-20.14 ± 0.06
J1503+3644	weak	0.3557	0.165 ± 0.026	1.69 ± 0.53	0.49 ± 0.18	0.033 <sup>+0.019</sup> <sub>-0.006</sub>	-21.52 ± 0.28	-20.08 ± 0.06
J1154+2443	strong	0.3690	0.046 ± 0.029	3.69 ± 1.28	0.13 ± 0.08	0.625 <sup>+0.375</sup> <sub>-0.241</sub>	-19.62 ± 0.32	-19.22 ± 0.08
J0901+2119	weak	0.2993	0.208 ± 0.073	6.12 ± 1.85	0.36 ± 0.17	0.026 <sup>+0.063</sup> <sub>-0.017</sub>	-20.30 ± 0.82	-18.48 ± 0.18
J1011+1947	strong	0.3322	0.131 ± 0.058	2.00 ± 1.74	0.07 ± 0.11	0.090 <sup>+0.232</sup> <sub>-0.030</sub>	-19.99 ± 0.63	-18.85 ± 0.14
J1243+4646	strong	0.4317	0.027 ± 0.019	2.34 ± 1.13	0.22 ± 0.04	0.889 <sup>+0.111</sup> <sub>-0.263</sub>	-20.29 ± 0.21	-20.05 ± 0.05
J1248+4259	non-leaker	0.3629	0.214 ± 0.085	7.34 ± 2.06	0.05 ± 0.07	≤ 0.013	-20.95 ± 0.97	-19.09 ± 0.24
J1256+4509	strong	0.3530	0.062 ± 0.032	1.00 ± 0.67	0.10 ± 0.07	0.385 <sup>+0.249</sup> <sub>-0.093</sub>	-19.63 ± 0.36	-19.08 ± 0.08
J0232-0426	non-leaker	0.4524	0.020 ± 0.042	6.47 ± 1.79	0.10 ± 0.10	≤ 0.047	-19.03 ± 0.49	-18.86 ± 0.14
J0919+4906	weak	0.4051	0.181 ± 0.081	4.12 ± 2.45	0.07 ± 0.11	0.049 <sup>+0.183</sup> <sub>-0.027</sub>	-20.79 ± 0.94	-19.21 ± 0.24
J1046+5827	non-leaker	0.3968	0.036 ± 0.027	2.58 ± 1.13	0.16 ± 0.07	≤ 0.010	-20.07 ± 0.29	-19.75 ± 0.07
J1121+3806	strong	0.3179	0.032 ± 0.033	4.99 ± 1.91	0.08 ± 0.10	0.433 <sup>+0.510</sup> <sub>-0.195</sub>	-18.81 ± 0.36	-18.53 ± 0.09
J1127+4610	non-leaker	0.3223	0.036 ± 0.029	8.53 ± 1.34	0.15 ± 0.10	≤ 0.059	-18.72 ± 0.33	-18.41 ± 0.08
J1233+4959	strong	0.4219	0.037 ± 0.045	6.38 ± 1.47	0.21 ± 0.10	0.192 <sup>+0.063</sup> <sub>-0.124</sub>	-19.76 ± 0.52	-19.44 ± 0.13
J1349+5631	non-leaker	0.3637	0.156 ± 0.072	2.57 ± 0.93	0.69 ± 0.23	≤ 0.009	-20.02 ± 0.83	-18.65 ± 0.20
J1355+1457	non-leaker	0.3651	0.048 ± 0.042	7.81 ± 0.89	0.06 ± 0.06	≤ 0.053	-19.31 ± 0.48	-18.90 ± 0.12
J1455+6107	non-leaker	0.3679	0.075 ± 0.042	5.38 ± 0.93	0.23 ± 0.08	≤ 0.019	-19.80 ± 0.48	-19.15 ± 0.12
J091021+610550	weak	0.2720	0.244 ± 0.010	2.43 ± 0.30	0.45 ± 0.06	0.028 <sup>+0.004</sup> <sub>-0.004</sub>	-23.10 ± 0.11	-20.97 ± 0.02
J143256+274249	strong	0.2660	0.207 ± 0.009	3.73 ± 0.19	0.61 ± 0.13	0.060 <sup>+0.009</sup> <sub>-0.007</sub>	-23.00 ± 0.10	-21.19 ± 0.02
J124206+011537	non-leaker	0.2710	0.313 ± 0.009	2.63 ± 0.28	0.81 ± 0.16	≤ 0.000	-23.44 ± 0.09	-20.70 ± 0.02

**Notes.** <sup>1</sup> References: Izotov et al. (2016a,b, 2018a,b, 2021), Wang et al. (2019).

**Table A.4.** Absorption lines results for the literature sample<sup>1</sup>. Columns are the same as in Table A.2. Sources are sorted by year of publication and Right-Ascension coordinate.

object	LyC type	$z$	$W_{\text{HI}}$ (Å)	$W_{\text{LIS}}$ (Å)	$R_{\text{HI}}$	$R_{\text{LIS}}$	$f_{\text{esc}}^{\text{abs, HI}}$	$f_{\text{esc}}^{\text{abs, LIS}}$
J0925+1403	strong	0.3013	0.41 ± 0.48	0.82 ± 0.35	0.33 ± 0.19	0.70 ± 0.11	0.092 ± 0.062	0.083 ± 0.032
J1152+3400	strong	0.3419	2.21 ± 0.30	0.58 ± 0.51	0.17 ± 0.10	0.66 ± 0.17	0.052 ± 0.043	0.090 ± 0.054
J1333+6246	strong	0.3181	2.47 ± 0.26	0.92 ± 0.66	0.09 ± 0.06	0.54 ± 0.21	0.044 ± 0.039	0.075 ± 0.073
J1442-0209	strong	0.2937	1.61 ± 0.29	0.10 ± 0.44	0.26 ± 0.10	0.70 ± 0.16	0.121 ± 0.053	0.105 ± 0.052
J1503+3644	weak	0.3557	1.53 ± 0.21	0.73 ± 0.50	0.30 ± 0.06	0.52 ± 0.14	0.043 ± 0.017	0.021 ± 0.014
J1154+2443	strong	0.3690	1.33 ± 0.20	0.86 ± 0.90	0.47 ± 0.08	0.53 ± 0.22	0.272 ± 0.115	0.104 ± 0.090
J0901+2119	weak	0.2993	0.97 ± 0.56	1.78 ± 0.79	0.14 ± 0.15	0.42 ± 0.22	0.019 ± 0.038	0.010 ± 0.025
J1011+1947	strong	0.3322	2.18 ± 0.31	0.01 ± 0.88	0.25 ± 0.09	0.59 ± 0.26	0.055 ± 0.060	0.046 ± 0.064
J1243+4646	strong	0.4317	0.77 ± 0.16	-0.28 ± 0.55	0.62 ± 0.06	0.71 ± 0.16	0.470 ± 0.124	0.231 ± 0.087
J1248+4259	non-leaker	0.3629	1.28 ± 0.43	0.88 ± 1.39	0.03 ± 0.13	0.57 ± 0.36	0.003 ± 0.021	0.013 ± 0.049
J1256+4509	strong	0.3530	0.80 ± 0.21	-0.60 ± 0.67	0.37 ± 0.07	0.70 ± 0.17	0.175 ± 0.081	0.156 ± 0.086
J0232-0426	non-leaker	0.4524	2.96 ± 0.24	-0.89 ± 3.64	0.04 ± 0.04	0.47 ± 0.84	0.030 ± 0.050	0.169 ± 0.297
J0919+4906	weak	0.4051	1.42 ± 0.34	0.87 ± 1.23	0.26 ± 0.11	0.50 ± 0.34	0.028 ± 0.069	0.012 ± 0.065
J1046+5827	non-leaker	0.3968	2.98 ± 0.18	1.74 ± 0.59	0.11 ± 0.05	0.40 ± 0.15	0.092 ± 0.043	0.048 ± 0.061
J1121+3806	strong	0.3179	0.85 ± 0.27	-0.03 ± 0.98	0.32 ± 0.07	0.65 ± 0.34	0.219 ± 0.111	0.213 ± 0.172
J1127+4610	non-leaker	0.3223	1.98 ± 0.19	0.34 ± 0.68	0.10 ± 0.04	0.53 ± 0.19	0.064 ± 0.038	0.128 ± 0.101
J1233+4959	strong	0.4219	1.68 ± 0.22	0.53 ± 1.05	0.20 ± 0.06	0.61 ± 0.31	0.119 ± 0.094	0.159 ± 0.192
J1349+5631	non-leaker	0.3637	2.39 ± 0.35	0.34 ± 1.11	0.04 ± 0.08	0.54 ± 0.24	0.009 ± 0.026	0.024 ± 0.068
J1355+1457	non-leaker	0.3651	2.85 ± 0.21	0.72 ± 0.85	0.08 ± 0.06	0.48 ± 0.21	0.059 ± 0.053	0.093 ± 0.114
J1455+6107	non-leaker	0.3679	2.67 ± 0.21	1.73 ± 0.58	0.10 ± 0.05	0.62 ± 0.19	0.037 ± 0.035	0.107 ± 0.087
J091021+610550	weak	0.2720	1.48 ± 0.16	1.32 ± 0.16	0.58 ± 0.07	0.68 ± 0.06	0.019 ± 0.003	0.017 ± 0.003
J143256+274249	strong	0.2660	2.60 ± 0.09	0.97 ± 0.15	0.41 ± 0.04	0.71 ± 0.06	0.035 ± 0.005	0.027 ± 0.004
J124206+011537	non-leaker	0.2710	4.06 ± 0.07	3.09 ± 0.10	0.11 ± 0.02	0.30 ± 0.03	0.004 ± 0.001	0.001 ± 0.001

**Notes.** <sup>1</sup> References: Izotov et al. (2016a,b, 2018a,b, 2021), Wang et al. (2019).

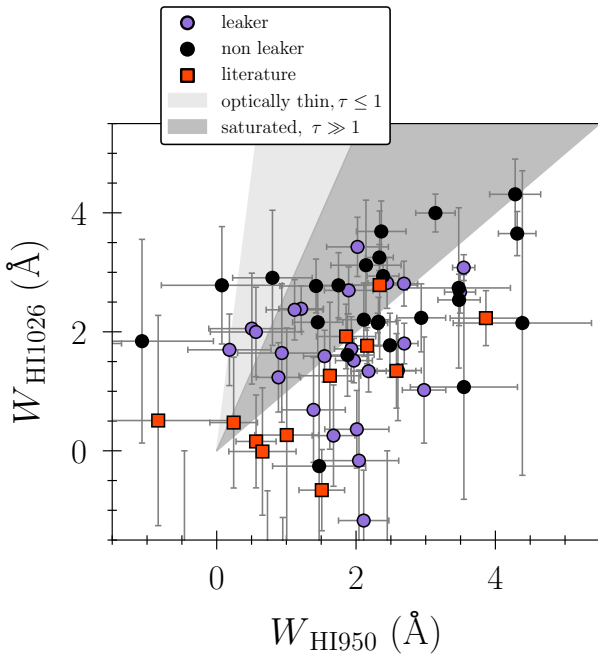
## Appendix B: Absorption lines – saturation

The ratio of two absorption lines of the same element and ionization state can indicate whether the species in question is saturated or not. Let  $f_i, \lambda_i$  be the oscillator strength and central rest-frame wavelength of the line in question, where the 2(1) subscript represents the stronger (weaker) transition. If both transitions are saturated, the equivalent width ratio  $W_2/W_1$  can be approximated by [Draine \(2011\)](#)

$$\frac{W_2}{W_1} \approx \left[ 1 + \frac{\ln(f_2 \lambda_2 / f_1 \lambda_1)}{\ln(\tau_1 / \ln 2)} \right]^{(1/2)}, \quad (\text{B.1})$$

where  $\tau_1$  is the optical depth of the stronger transition, being  $\tau_1 \gg 1$  the optically thick limit. If, on the opposite, both lines behave under the optically thin regime ( $\tau_1 \ll 1$ ), the equivalent width doublet-ratio is simply:

$$\frac{W_2}{W_1} = \frac{f_2 \lambda_2}{f_1 \lambda_1}. \quad (\text{B.2})$$



**Fig. B.1.** Absorption line H I equivalent width comparisons for studying the condition of saturation. The optically thin (light-gray) and optically thick (dark-gray) regions are delimited using the approximations given in [Draine \(2011\)](#). As in the entire document, purple symbols represent the leakers while the black symbols show the nonleakers. Red squares correspond to the [Izotov et al. \(2016a,b, 2018a,b, 2021\)](#) and [Wang et al. \(2019\)](#) galaxies.

In practice, we assume  $\tau_1 > 1$  for saturation, and a smooth transition between the two regimes. To test the saturation of H I in our galaxy spectra, we compare the equivalent width of H I  $\lambda 1026$  to H I  $\lambda 950$ . The use of lines pairs which are close in wavelength ensures that both lines are in the same optical regime simultaneously.

The observations are shown in Figure B.1, together with the regions delimiting the optically thick (saturation) and optically thin (linear) regimes, respectively, following the equations above. As stressed in the text, most of the sources are compatible with saturation for H I within the  $1\sigma$  errors, and only a low number of them fall in the optically thin limit. Several sources

are found to show a higher equivalent width in H I  $\lambda 950$  than H I  $\lambda 1026$ , which is not expected from simple curve-of-growth analysis. The exact reasons for this behavior are not known but could be due to multiple effects, such as systematic errors in the continuum placement in the blue part of the spectrum, blends with other lines such as O I, and other causes. In any case, in most of the cases, they are compatible with the saturated region at  $1\sigma$ . Spectra with higher resolution and S/N, such as the subset of galaxies analyzed by [Gazagnes et al. \(2018\)](#), would be required to better examine the question of saturation.

## Appendix C: Absorption lines – systematic errors

Here, we carefully account for the impact of the resolution and S/N in the measured covering fraction. The spectrograph resolution tends to make the absorption lines wider but less deep. Together with the noise and sensitivity –and other possible instrumental effects– this can lead to a systematically overestimated measurement of the residual flux.

We use absorption line simulations to study how the resolution ( $R$ ) and S/N both impact our measurements of the residual flux. We first simulate Ly $\delta$ , Ly $\beta$ , Si II  $\lambda 1260$ , and C II  $\lambda 1334$  line intensities ( $I_\lambda$ ) in a foreground dust-screen picket-fence model:

$$I_\lambda = C_f(\lambda) \times \exp(-\sigma_\lambda N_\lambda) + 1 - C_f(\lambda), \quad (\text{C.1})$$

where  $C_f(\lambda)$  represents the covering fraction of the line in question, i.e., the fraction of sight lines for which the transition is optically thick around the host galaxy.  $C_f(\lambda)$  can be inferred from the line depth as  $C_f(\lambda) = 1 - R(\lambda)$  for the dust-screen picket-fence model. The  $\tau_\lambda = \sigma_\lambda N_\lambda$  product is usually known as the optical depth of the line, and the  $\lambda$  subscripts belong to the transition in question. The line cross-section,  $\sigma_\lambda$ , which shapes the absorption profile, is given by a usual Voigt function so that:

$$\sigma_\lambda = \frac{\sqrt{\pi} e^2 f_\lambda \lambda}{m_e c} \frac{\lambda}{b} \text{Voigt}(\lambda, A_\lambda, b). \quad (\text{C.2})$$

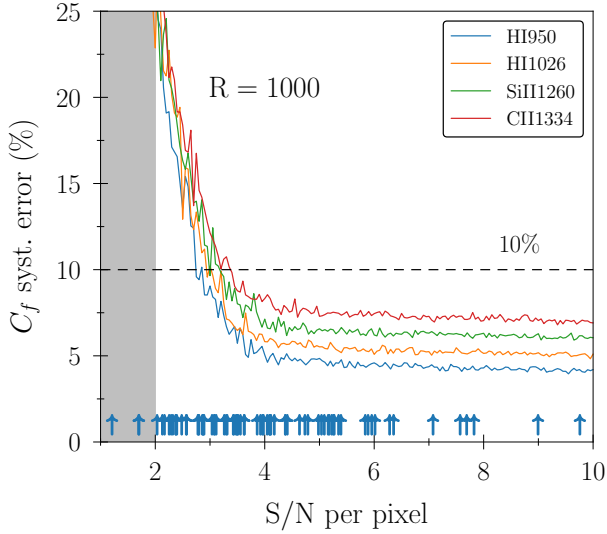
For the Voigt( $\lambda, A_\lambda, b$ ) function, we use the numerical approximation described in [Smith et al. \(2015\)](#). The Doppler broadening parameter is input through  $b$  (km s $^{-1}$ ). Finally, oscillator strengths  $f_\lambda$  and Einstein coefficients  $A_\lambda$  are taken from the NIST database<sup>12</sup>. We do not consider any velocity gradient effect (outflows) in the mock realizations.

For our purposes, we assume a Doppler parameter of  $b = 125$  km/s, and typical column densities in SF galaxies of  $N_\lambda = 10^{19}, 10^{19}, 10^{17}, 10^{17}$  cm $^{-2}$  for Ly $\delta$ , Ly $\beta$ , Si II  $\lambda 1260$ , and C II  $\lambda 1334$  lines, respectively. We also explore a wide range of ( $N, b$ ) parameters and conclude that results do not change drastically as soon as the lines are in the saturation region of the curve of growth. The covering fraction is fixed to  $C_f(\lambda) = 0.85$  throughout the processes.

The mock lines are convolved with a Gaussian kernel to match the COS/G140L nominal resolution ( $R = 1000$ ) and interpolated into the rest-frame COS/G140L wavelength spacing (0.5621/(1+z) Å, where  $z = 0.3$ ). A set of Gaussian S/N are then randomly added to every pixel. Finally, the residual flux is measured in the same way as in the original real spectra, and related to the covering fraction by Eq. (4) (uniform dust-screen). This process is repeated 500 times for every S/N value, and the median covering fraction is compared to the input value.

<sup>12</sup> NIST stands for National Institute of Standards and Technology, whose website is: [https://physics.nist.gov/PhysRefData/ASD/lines\\_form.html](https://physics.nist.gov/PhysRefData/ASD/lines_form.html)





**Fig. C.1.** Estimation of the systematic error on the observed covering fraction as a function of the spectral S/N per pixel, for different simulated lines (see legend). A resolution of  $R = 1000$  is assumed. The upward arrows at the bottom correspond to the measured S/N for the LzLCS sample around the HI950 line, for comparison (the S/N has been measured at the continuum level close to the line in question). In general, the observed continuum S/N increases at shorter wavelengths, i.e., it is larger for Ly $\delta$  and Ly $\beta$ , but systematically lower for Si II  $\lambda$ 1260 and C II  $\lambda$ 1334.

This method is summarized in Fig. C.1, which shows the covering fraction relative difference (in percentage) as a function of the input S/N per pixel in the simulations. The four colored lines corresponds to the different lines considered. The upward arrows at the bottom correspond to the measured S/N for the LzLCS sample around the HI950 line, for comparison (the S/N has been measured at the continuum level close to the line in question). In general, the observed continuum S/N increases at shorter wavelengths, i.e., it is larger for Ly $\delta$  and Ly $\beta$ , but systematically lower for Si II  $\lambda$ 1260 and C II  $\lambda$ 1334.

In all cases, and as indicated in the text, the systematic error on the covering fraction spans between 5% and 20% for these four representative lines (with a typical value of 10%). These values are within the original error bars we are reporting for single covering fraction measurements. We therefore leave our initial lines measurements and errors unaltered, and we prefer not to correct the  $C_f(\lambda)$  values by this systematic offset.

## Appendix D: The LINMIX package

The LINMIX estimator (Kelly 2007) is a Python-based Bayesian fitting code that allows the user to model a two-dimensional data set  $(x, y)$  with a linear regression, accounting for errors on both variables and intrinsic random scatter, and with the capability of including censored (upper or lower limits) data.

LINMIX assumes that measured data points  $(x, y)$  can be sampled from a two-dimensional Gaussian distribution  $P_1$  with mean  $(\xi, \eta)$ , and covariance matrix determined from measurement uncertainties. It also assumes that the measurement errors

in the  $x$ - and  $y$ -directions are normally distributed and uncorrelated.

The “true” value of the  $\eta$  variable is drawn from another Gaussian distribution  $P_2$ , with mean  $\alpha + \beta\eta$ , and variance  $\sigma_y^2$ , where  $\beta$  describes the slope,  $\alpha$  the intercept of the straight line, and  $\sigma_y^2$  is the intrinsic  $y$ -dispersion (also normally distributed). Finally, the “true” value of the independent variable  $\xi$  is sampled from a weighted sum of  $K$  Gaussian distributions  $P_3$  (we choose  $K = 2$ ). The distributions  $P_1, P_2, P_3$  are convolved hierarchically to compute the full likelihood of obtaining the observed data  $(x, y)$  given the parameters of the fit:  $\alpha, \beta, \sigma_y^2$ .

Assuming uniform prior distributions for the three parameters:  $\alpha, \beta \in (-\infty, \infty)$ ,  $\sigma_y^2 \in [0, \infty)$ ; a Markov Chain Monte-Carlo (MCMC) algorithm is used to sample from the posterior distributions until the convergence of the chains ( $\times 4$  chains, 5000 to 10000 iterations each). For a practical example, see e.g., Kennicutt & De Los Reyes (2021). A PYTHON version of the LINMIX code can be found in <https://linmix.readthedocs.io/en/latest/index.html>.

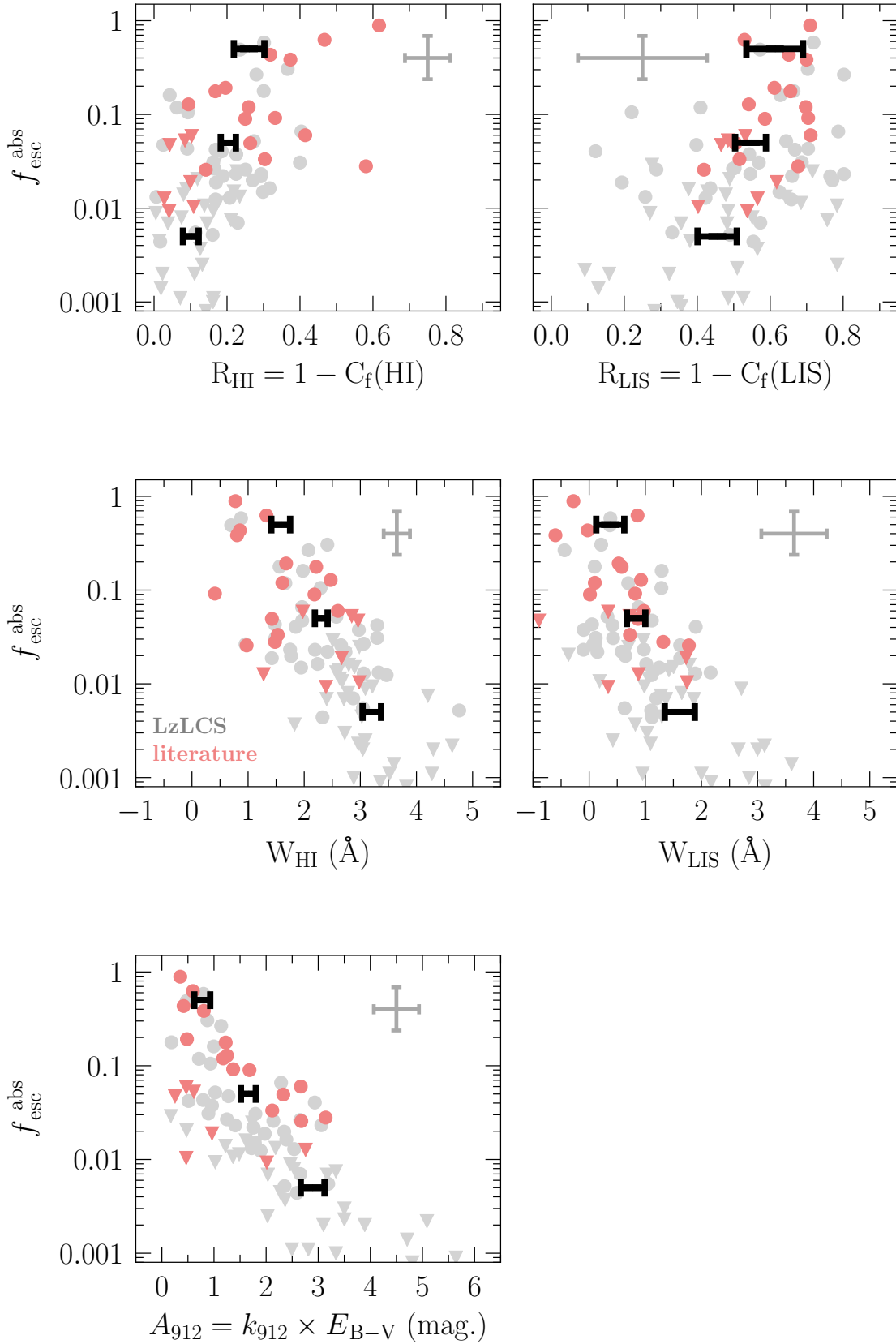
## Appendix E: Main results at a glance

This appendix is dedicated to illustrating the global behavior of the absolute LyC photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) with the main UV spectral properties measured in this work, namely: R(HI), R(LIS),  $W_{\text{HI}}$ ,  $W_{\text{LIS}}$  and  $E_{\text{B-V}}$ . These results are presented in Figs. 11–13, but here we focus on the median values of the working sample as a take-home message. To do this, we compute the median value of R(HI), R(LIS),  $W_{\text{HI}}$ ,  $W_{\text{LIS}}$  and  $E_{\text{B-V}}$  in three regimes of  $f_{\text{esc}}^{\text{abs}}$ , specifically at  $f_{\text{esc}}^{\text{abs}} \in [0\%–5\%), [5\%–10\%), [10\%–100\%)$ . We additionally apply a bootstrap + Monte Carlo algorithm to the number of sources at each  $f_{\text{esc}}^{\text{abs}}$  interval allowing us to compute the 0.16 and 0.84 quantiles at each bin. The results are summarized in Table E.1.

**Table E.1.** Running median results for R(HI), R(LIS),  $W_{\text{HI}}$ ,  $W_{\text{LIS}}$  (in Å) and  $E_{\text{B-V}}$  (in mag.) in three  $f_{\text{esc}}^{\text{abs}}$  intervals (including 0.16 and 0.84 quantiles).

	LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ )		
	0 – 5 (%)	5 – 10 (%)	10 – 100 (%)
R(HI)	0.10 (0.08, 0.12)	0.20 (0.18, 0.22)	0.26 (0.21, 0.30)
R(LIS)	0.46 (0.40, 0.51)	0.54 (0.50, 0.58)	0.61 (0.54, 0.69)
$W_{\text{HI}}$	3.23 (3.06, 3.39)	2.30 (2.19, 2.41)	1.58 (1.41, 1.75)
$W_{\text{LIS}}$	1.61 (1.35, 1.87)	0.85 (0.69, 1.01)	0.37 (0.11, 0.62)
$A_{\text{UV}}$	2.88 (2.66, 3.11)	1.67 (1.53, 1.81)	0.77 (0.63, 0.93)

In Fig. E.1, the running medians are plotted as a function of  $f_{\text{esc}}^{\text{abs}}$  (in black), together with the individual LzLCS+literature measurements (gray and red dots in the background). The reader may find a detailed discussion about these results in Sect. 6. A proper analysis of the median results of the sample will be investigated in future LzLCS publications, by the use of stacked spectra.



**Fig. E.1.** Summary plot with the correlations between the LyC absolute photon escape fraction ( $f_{\text{esc}}^{\text{abs}}$ ) and the main UV spectral properties derived in this work:  $R(\text{HI})$ ,  $R(\text{LIS})$ ,  $W_{\text{HI}}$ ,  $W_{\text{LIS}}$  and  $E_{\text{B-V}}$  (from top to bottom and left to right, respectively). Black thick error bars comprise the running median values and 0.16 and 0.84 quantiles of the parameters as a function of  $f_{\text{esc}}^{\text{abs}}$ . The individual measurements are shown in the background through gray (LzLCS) and red (literature) symbols. Error bars at the top of the panels represent the median  $1\sigma$  uncertainty of each parameter, based on individual errors.