Lyα-Lyman continuum connection in 3.5 ≤ z ≤ 4.3 star-forming galaxies from the VUDS survey[⋆]

F. Marchi[1], L. Pentericci[1], L. Guaita[13], D. Schaerer[3,4], A. Verhamme[3], M. Castellano[1], B. Ribeiro[2], B. Garilli[5], O. Le Fèvre[2], R. Amorin[11,12], S. Bardelli[8], P. Cassata[7], A. Durkalec[10], A. Grazian[1], N. P. Hathi[6], B. C. Lemaux[9], D. Maccagni[5], E. Vanzella[8], and E. Zucca[8]

[1] INAF–Osservatorio Astronomico di Roma, via Frascati 33, 00040 Monte Porzio Catone, Italy
e-mail: `francesca.marchi@oa-roma.inaf.it`
[2] Aix–Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille, UMR 7326, 13388 Marseille, France
[3] Geneva Observatory, University of Geneva, ch. des Maillettes 51, 1290 Versoix, Switzerland
[4] Institut de Recherche en Astrophysique et Planétologie - IRAP, CNRS, Université de Toulouse, UPS-OMP, 14 avenue É. Belin, 31400 Toulouse, France
[5] INAF–IASF Milano, via Bassini 15, 20133, Milano, Italy
[6] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[7] Instituto de Fisica y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Gran Bretaña 1111, Playa Ancha, Valparaíso, Chile
[8] INAF–Osservatorio Astronomico di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy
[9] Department of Physics, University of California, Davis, One Shields Ave., Davis, CA 95616, USA
[10] National Center for Nuclear Research, ul. Hoza 69, 00-681 Warszawa, Poland
[11] Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge, CB3 0HE, UK
[12] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[13] Nùcleo de Astronomìa, Facultad de Ingenierìa, Universidad Diego Portales, Av. Ejèrcito 441, Santiago, Chile

ABSTRACT

Context. To identify the galaxies responsible for the reionization of the Universe, we must rely on the investigation of the Lyman continuum (LyC) properties of $z \lesssim 5$ star-forming galaxies, where we can still directly observe their ionizing radiation.

Aims. The aim of this work is to explore the correlation between the LyC emission and some of the proposed indirect indicators of LyC radiation at $z \sim 4$ such as a bright Lyα emission and a compact UV continuum size.

Methods. We selected a sample of 201 star-forming galaxies from the Vimos Ultra Deep Survey (VUDS) at $3.5 \leq z \leq 4.3$ in the COSMOS, ECDFS, and VVDS-2h fields, including only those with reliable spectroscopic redshifts, a clean spectrum in the LyC range and clearly not contaminated by bright nearby sources in the same slit. For all galaxies we measured the Lyα EW, the Lyα velocity shift with respect to the systemic redshift, the Lyα spatial extension and the UV continuum effective radius. We then selected different sub-samples according to the properties predicted to be good LyC emission indicators: in particular we created sub-samples of galaxies with $EW(\mathrm{Ly}\alpha) \geq 70\,\text{Å}$, $\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$, $r_{UV} \leq 0.30\,\text{kpc}$ and $|\Delta v_{\mathrm{Ly}\alpha}| \leq 200\,\text{km s}^{-1}$. We stacked all the galaxies in each sub-sample and measured the flux density ratio ($f_\lambda(895)/f_\lambda(1470)$), that we considered to be a proxy for LyC emission. We then compared these ratios to those obtained for the complementary samples. Finally, to estimate the statistical contamination from lower redshift inter-lopers in our samples, we performed dedicated Monte Carlo simulations using an ultradeep U-band image of the ECDFS field.

Results. We find that the stacks of galaxies which are UV compact ($r_{UV} \leq 0.30\,\text{kpc}$) and have bright Lyα emission ($EW(\mathrm{Ly}\alpha) \geq 70\,\text{Å}$), have much higher LyC fluxes compared to the rest of the galaxy population. These parameters appear to be good indicators of LyC radiation in agreement with theoretical studies and previous observational works. In addition we find that galaxies with a low Lyα spatial extent ($\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$) have higher LyC flux compared to the rest of the population. Such a correlation had never been analysed before and seems even stronger than the correlation with high $EW(\mathrm{Ly}\alpha)$ and small r_{UV}. These results assume that the stacks from all sub-samples present the same statistical contamination from lower redshift interlopers. If we subtract a statistical contamination from low redshift interlopers obtained with the simulations from the flux density ratios ($f_\lambda(895)/f_\lambda(1470)$) of the significant sub-samples we find that these samples contain real LyC leaking flux with a very high probability, although the true average escape fractions are very uncertain.

Conclusions. Our work indicates that galaxies with very high $EW(\mathrm{Ly}\alpha)$, small $\mathrm{Ly}\alpha_{\text{ext}}$ and small r_{UV} are very likely the best candidates to show Lyman continuum radiation at $z \sim 4$ and could therefore be the galaxies that have contributed most to reionisation.

Key words. galaxies: high-redshift – Galaxy: evolution – surveys

1. Introduction

Understanding the processes that led to the reionization of the Universe is among the most challenging tasks of modern extra-galactic astronomy. The most likely objects responsible

for this phenomenon until now are star-forming galaxies and active galactic nuclei (AGN) that, at $z \sim 6$, completely ionized the intergalactic medium (IGM) thanks to the emission of the so-called Lyman continuum (LyC) radiation (Robertson et al. 2010, 2015; Shull et al. 2012; Becker & Bolton 2013; Giallongo et al. 2015), which is at $\lambda < 912$ Å. However, at redshifts higher than $z \sim 4.5$, the IGM becomes less transparent to LyC photons due to the increasing number of intervening absorption systems and can prevent the direct detection of Lyman continuum flux (Madau 1995; Madau et al. 1999; Dijkstra et al. 2004; Inoue & Iwata 2008; Prochaska et al. 2009; Laursen et al. 2011; Inoue et al. 2014; Worseck et al. 2014). It is therefore not possible to directly study the LyC emission of the sources responsible for the reionisation. What we can do, however, is to study the ionizing emission properties of lower redshift galaxies and later infer if these properties were more common during the reionization epoch.

At $z \gtrsim 2.5$, the LyC radiation is redshifted into the optical spectral region and therefore, for galaxies in the range $2.5 \lesssim z \lesssim 4.5$, it can be detected using ground-based observations. For galaxies at lower redshift, we must instead rely on space-based observations. The ionizing radiation is significantly attenuated by neutral gas and dust in the interstellar and circumgalactic medium of the sources (Leitherer et al. 1995; Deharveng et al. 2001). Therefore, the detection of LyC emission in individual galaxies is a rather difficult task. Furthermore, at high redshifts the search for LyC emitters (LCEs) is made much more complicated by the high probability of contamination by lower redshift interlopers, the faintness of the sources and the increase of the IGM opacity with redshift (Vanzella et al. 2010, 2012; Inoue et al. 2014). In particular, the line of sight (LoS) contamination is one of the main limitations of LyC studies when imaging and spectroscopic observations are taken from the ground (Vanzella et al. 2012). Indeed low-redshift galaxies can mimic the LyC emission from high-redshift sources if they are located very close to the target galaxies and the spatial resolution does not allow us to distinguish the two objects. These nearby contaminants can only be identified in high-resolution HST images because they appear to be blended in ground-based observations. In most cases, the putative LyC emission appears offset in HST images with respect to the main optical galaxy, indicating the presence of a possible lower-redshift contaminant (Nestor et al. 2011; Mostardi et al. 2013; Grazian et al. 2016).

Blind searches for LCEs have not, indeed, been very productive so far. Only three LyC emitters have been found with blind searches in the local Universe (Bergvall et al. 2006; Leitet et al. 2011, 2013; Leitherer et al. 2016) and only two detections have been reported at high redshifts (Shapley et al. 2016; Mostardi et al. 2015). To overcome this lack of detections, several pre-selection methods to find good LyC leaker candidates have been proposed, leading to the discovery of six further LCEs at low redshifts (Borthakur et al. 2014; Izotov et al. 2016) and one LCE at high redshift (Vanzella et al. 2016). It has been found indeed that some galaxy properties can be related to a high escape fraction of ionizing radiation. These features are the nebular emission line strengths (Zackrisson et al. 2013), high [OIII]λ5007/[OII]λ3727 ratios that could trace density-bounded HII regions (Jaskot & Oey 2013; Nakajima & Ouchi 2014) and the non-saturation of the metallic low-ionization absorption lines that trace a low covering fraction of the absorbing gas along the line of sight (Heckman et al. 2011; Alexandroff et al. 2015; Vasei et al. 2016). It is also believed that the faintest, low-mass star forming galaxies are responsible for the bulk of the ionizing radiation during the

reionization epoch: the observed UV luminosity function is very steep at high-z and these faint galaxies should be very numerous (Ouchi et al. 2009; Wise & Cen 2009b; Yajima et al. 2011; Bouwens et al. 2012; Mitra et al. 2013; Mason et al. 2015) hence providing the necessary ionizing budget. Since high redshift Lyman alpha emitters (LAEs) are, in general, low-mass galaxies (e.g. Finkelstein et al. 2007; Bouwens et al. 2007; Pentericci et al. 2010), characterized by a faint UV continuum, a bright Lyα emission line has also been proposed as a pre-selection tool to look for LyC emitters (Verhamme et al. 2015; Dijkstra et al. 2016). UV morphology can also be used as an indirect indicator of LyC emission; very compact star-forming regions can photoionize the ISM creating the so-called density-bounded regions (Nakajima & Ouchi 2014) or can shape low density channels through the ISM by mechanical feedback, facilitating the escape of LyC and Lyα radiation. A connection between the LyC, Lyα emissions and UV compactness is therefore expected at some level.

There have been several theoretical studies investigating the above correlations. For example Dijkstra et al. (2016) explore the correlation between LyC and Lyα radiation using a large suite of simplified models of the multi-phase ISM that span the wide range of astrophysical conditions encountered in observed galaxies. They model the source of LyC and Lyα radiation surrounded by a collection of spherical clumps of dust and neutral hydrogen gas which are opaque to LyC radiation and that are embedded within a hot inter-clump medium (as in Hansen & Oh 2006; Laursen et al. 2013; Gronke & Dijkstra 2014). In these models LyC photons can escape from the galaxy if their sight-line does not encounter any clump. They find that galaxies with a low escape fraction of Lyα photons ($f_{\rm esc}^{\rm Ly\alpha}$), also present a low escape fraction of LyC photons ($f_{\rm esc}^{\rm LyC}$), while galaxies with high values of $f_{\rm esc}^{\rm Ly\alpha}$ present a large spread in $f_{\rm esc}^{\rm LyC}$. Finally, they find galaxies that show LyC emission typically have narrower and more symmetric Lyα line profiles and a low velocity offset with respect to the systemic redshift. Verhamme et al. (2015) use a similar approach to study the LyC-Lyα connection. They use the classic shell model to picture the galaxy in two different configurations: 1) totally ionized ISM and 2) riddled (i.e., with ionized holes) ionization-bounded ISM. In the first case they find a Lyα spectrum characterized by a very narrow profile with a small shift with respect to the systemic redshift (as Dijkstra et al. 2016) whereas in the second case they find a Lyα peak well centered at the systemic z but with a flux redwards. They find also that, for galaxies with very low or null outflow velocity for which the Lyα profile is characterized by a double peak, a small separation between the two peaks is a strong indicator of LyC emission.

As mentioned above, these indirect indicators have been partly confirmed by observations. Izotov et al. (2016) showed that selecting for compact star-forming galaxies showing high [OIII]λ5007/[OII]λ3727 ratios (>5) appears to very efficiently pick up sources that have escaping Lyman continuum radiation at low redshift: they find LyC emission from all the five galaxies selected by these criteria, considerably increasing the number of known LCEs at low redshift. On the other hand, Rutkowski et al. (2017), selecting a sample of $z \sim 2$ star-forming galaxies with the same constraint on the [OIII]λ5007/[OII]λ3727 ratio as Izotov et al. (2016), did not find any individual detections. Verhamme et al. (2017) analysed the Lyα spectral properties of several known LyC emitters finding that all the LCEs in the local Universe are characterized by a double peak Lyα profile with a small peak separation in agreement with the theoretical

expectations. Furthermore, they find that several known LyC sources present a Lyα in emission with very high EW and a large Lyα escape fraction ($f_{esc}^{Ly\alpha} > 20\%$) as predicted by Dijkstra et al. (2016). Finally, they observe a correlation between the escape fraction of ionizing photons and the SFR surface density. This is evidence that the compactness of star-forming regions could play a significant role in the escape of ionizing radiation.

In our previous paper (Marchi et al. 2017), we also found an indication of a possible positive trend of the flux density ratio ($f_\nu(910)/f_\nu(1500)$) as a function of Lyα EW using a sample of 12 LAEs at $z \sim 3.8$ in the VIMOS Ultra Deep Survey (VUDS; Le Fèvre et al. 2015). A similar correlation has been also found by Micheva et al. (2017) with a sample of 18 LAEs at $z \geq 3.06$ in the SSA22 proto-cluster. These studies give support to the correlation between the escape of Lyα photons (correlated to the Lyα EW) and the escape of LyC radiation.

Finally we mention *Ion2*, a high redshift galaxy that was initially identified as a possible LyC emitter by Vanzella et al. (2015): subsequent observations confirmed the presence of LyC radiation from this source (Vanzella et al. 2016; de Barros et al. 2016). This galaxy also presents many physical properties typical of LCEs and discussed above as indirect indicators (e.g. small UV size and high Lyα EW). Finally, six more LyC emitting candidates have been recently proposed by Naidu et al. (2017) but their LyC emission has still to be verified.

These latest results seem to indicate that the use of indirect indicators to identify the best LyC emitters candidates could be very efficient. In this paper we extend this analysis to high redshift to test some of the proposed trends, using a large sample of galaxies at $z \sim 4$ with spectra from the VUDS survey. Thanks to this large dataset, we can select different sub-samples of galaxies according to the properties predicted to be good LyC emission indicators and then test if the population of galaxies in each sub-sample does present an excess in the LyC part of the spectrum.

This paper is organized as follows. In Sect. 2 we describe the selection criteria that we used to select our sample of high redshift star-forming galaxies from the VUDS dataset. In Sect. 3 we describe the methods to evaluate the different galaxies' properties, the criteria to select each sub-sample and the technique to evaluate the flux density ratio, $\frac{f_\lambda(895)}{f_\lambda(1470)}$, for each sub-sample. In Sect. 4 we present the obtained results. Finally, in Sect. 5 we discuss the effects of contamination from lower redshift interlopers on our samples. Throughout the paper we adopt the Λ cold dark matter (Λ-CDM) cosmological model ($H_0 = 70 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$). All magnitudes are in the AB system. All EWs presented in this paper are given in the rest-frame and positive values correspond to those lines measured in emission.

2. Sample selection

We selected a sample of star-forming galaxies from the VUDS database[1] (Le Fèvre et al. 2015; Tasca et al. 2017), which is the largest, up to date, spectroscopic survey of galaxies at $z > 2$. VUDS acquired approximately 7000 spectra of galaxies at $2 \leq z \leq 6$ in the COSMOS, ECDFS and VVDS-2h fields. We selected all galaxies with reliable redshifts in the range $3.5 \leq z \leq 4.3$. In order to measure a possible LyC signal in the wavelength range covered by the VUDS spectra (3800−9400 Å), we require sources at $z > 3.5$ to observe the LyC wavelength

domain redshifted into the spectral interval of the VUDS observations. The upper redshift limit is related to the almost totally opaque IGM at $z > 4.5$ (Madau 1995; Laursen et al. 2011; Inoue et al. 2014). We, however, use a more restricted cut at $z = 4.3$ to maximize the signal-to-noise ratio (S/N), given that at $z > 4.3$ there are only few good quality spectra.

The redshift determination of the VUDS targets is explained in details in Le Fèvre et al. (2015). Briefly, each spectrum is analysed by two different VUDS team members using the EZ tool (Garilli et al. 2010), a cross-correlation engine to compare spectra and a wide library of galaxy and star templates, along with a visual inspection of the spectra when only emission lines are present. A quality flag is assigned to each redshift, following the scheme that was tested in previous surveys (e.g. VVDS; Le Fèvre et al. 2014). In our analysis we included only galaxies with VUDS reliability flags 3, 4, 23, and 24 corresponding to a probability greater than 95% for the spectroscopic redshift to be correct (see Le Fèvre et al. 2015, for more details). The spectra are calibrated using spectrophotometric standard stars with a relative flux accuracy of better than ~5% over the wavelength range 3600 to 9300 Å. In addition, each spectrum is corrected for atmospheric extinction and for wavelength-dependent slit losses due to atmospheric refraction, taking into account the geometry of each source as projected into the slit (Thomas et al. 2017). The spectra are also corrected for the galactic extinction (Le Fèvre et al. 2015). With an integration of 14 net hours per target per grism, the VUDS spectra reach a S/N on the continuum at 8500 Å of $S/N = 5$ for $i_{AB} = 25$, and $S/N = 5$ for an emission line with a flux $f_\lambda = 1.5 \times 10^{-18} \text{erg s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$.

With these criteria we selected 246 galaxies. From this sample we excluded 37 galaxies due to spectral defects or strong residuals from skyline subtraction in the LyC region of the spectrum (870−910 Å rest-frame), and one galaxy with possible AGN features. We also discarded six further objects whose spectra were clearly contaminated by very bright neighbours, after visually checking the two-dimensional spectra and the available low resolution images. Finally, as even a few spurious excess in the LyC region of the spectrum can affect our measurements, we checked that the two-dimensional spectrum of each source was not contaminated by higher-order spectra coming from other slits in the same mask. We found only one source that could present this kind of contamination and we excluded it from the sample.

The final sample contains 201 galaxies. 106 of these are in the COSMOS field, 22 in the ECDFS field and 73 in the VVDS-2h field. The redshift and R magnitude distributions of the galaxies in the final sample, with the scatter plot between these two quantities, are shown in Fig. 1. The median redshift and R magnitude are $z = 3.81$ and $R = 24.91$, respectively.

In Marchi et al. (2017) and Guaita et al. (2016) a careful cleaning procedure based on multi-band high resolution HST imaging was applied, to exclude all the galaxies contaminated by any possible interloper. In these papers we found contamination fractions of 28% and 52%, respectively. This procedure cannot be applied on the present sample since the availability of multi-colour HST imaging would restrict the application of this procedure only to a very small subset (~35 galaxies) reducing the sample significantly. Instead, we simply excluded the clearly contaminated objects as previously explained in this section. We assumed for the rest of the analysis that all the sub-samples that we define in the next section contain a statistically similar contamination. We discuss the effects of the contamination from faint low-redshift interlopers on our samples in Sect. 5 and Appendix A.

---

[1] The first public release is available at `http://cesam.lam.fr/vuds/DR1/` but our sources are selected from the entire database
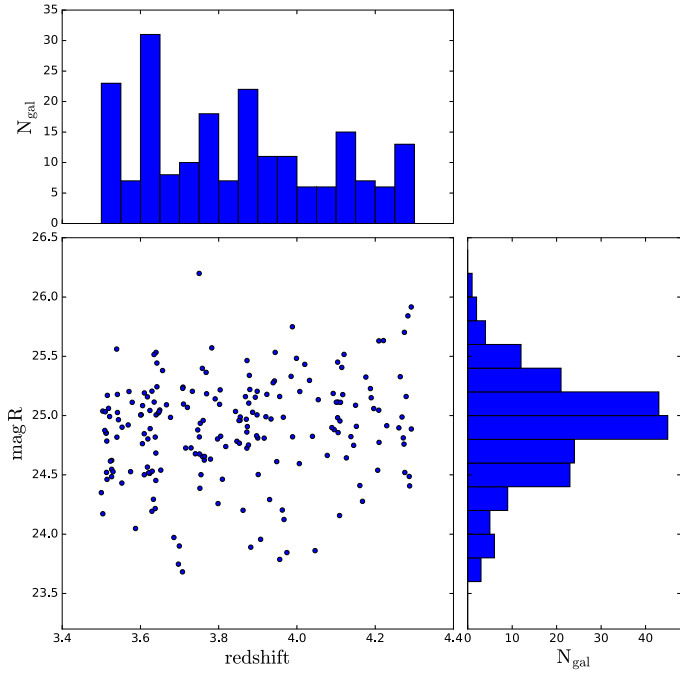
Fig. 1. R magnitude as a function of the redshift for the 201 star-forming galaxies in the final sample. The distributions of the two quantities are shown on the sides of the plot.

3. Method

3.1. Lyα and UV properties

The indirect indicators of LyC emission that we can in principle use exploiting the VUDS data, are: the Lyα EW and the Lyα FWHM (Dijkstra et al. 2016), the Lyα velocity offset with respect to the systemic redshift (Dijkstra et al. 2016; Verhamme et al. 2015), the Lyα extension (Verhamme et al. in prep) and the UV compactness (Wise & Cen 2009a; Izotov et al. 2016). In this section we describe the methods used to evaluate these quantities and their errors for the galaxies in our sample.

3.1.1. Lyα EW

We used the values of the Lyα EW from Cassata et al. (2015), that are evaluated using a continuum and line flux estimate from the IRAF splot tool. The EW(Lyα) distribution is shown in Fig. 2. Approximately 10% of the sample has $EW(\text{Ly}\alpha) \geq 55$ Å and ~25% has $EW(\text{Ly}\alpha) \geq 25$ Å in agreement with the statistics at this redshift (e.g. Shapley et al. 2003; Stark et al. 2010). We note that these fractions are slightly higher than those found by Cassata et al. (2015), who also used the VUDS data. This is due to the fact that we included in our sample only the spectra with quality flags 3 and 4 (probability greater than 95% for the spectroscopic redshift to be correct, see Le Fèvre et al. 2015) while Cassata et al. (2015) also used galaxies with a less secure spectroscopic redshift which usually do not show Lyα in emission. As already explained in the previous section however, we decided to include only the galaxies with secure spectroscopic redshifts in our sample to avoid the presence of spurious objects in our redshift range.

The errors on the EW(Lyα) were evaluated following Eq. (7) in Vollmann & Eversberg (2006) and are on average ~17% of the measured values.

3.1.2. Lyα velocity offset ($\Delta v_{\text{Ly}\alpha}$)

To estimate the Lyα velocity offset with respect to the systemic redshift, a good knowledge of the systemic redshift is needed: this can be obtained either from photospheric stellar absorption lines (such as OIV(1343 Å) and SiIII(1417 Å)) which are too weak in our individual spectra, or from nebular emission lines. Only 12 galaxies in our sample show CIII(1907.07Å) in emission: in these cases the systemic redshift was estimated from the centroid of the line. We note that, since the VUDS resolution does not allow us to distinguish the CIII(1907.07Å) doublet, we used a single Gaussian fit to estimate the centroid. For 33 further sources, characterized by strong SiII(1260.42Å), CII(1334.53Å) and/or SiII(1526.71Å) in absorption, we applied the relation found by Steidel et al. (2010) between the redshift of the inter-stellar absorption lines (z_{IS}) and the systemic redshift:

$$z_{\text{sys}} = z_{IS} + 0.00299 - 0.00291(2.7 - z_{IS}), \tag{1}$$

that was derived analysing a sample of 86 galaxies at $z \sim 2.3$ for which the systemic redshift was known from the Hα emission line. We assume here that this relation is also valid at our sample's redshifts (though see the discussion in the next paragraph). We then measured the center of the Lyα line and estimated $\Delta v_{\text{Ly}\alpha}$ with respect to the systemic redshift.

The distribution of the Lyα velocity offset is shown in Fig. 2. It presents an extended tail to negative velocities, which are not commonly observed in star-forming galaxies, both at low and high redshifts (e.g. Guaita et al. 2013; McLinden et al. 2014; Song et al. 2014). We could not test the Steidel et al. (2010) relation at our redshifts, so it might be that these negative velocities come from a bad evaluation of the systemic redshifts from the IS absorption lines. However, as shown in Fig. 2, this is likely not the case, since there are galaxies with negative velocities where the systemic redshift was evaluated from the CIII(1907.07 Å) emission. We also note that Gronke (2017) recently found that a small fraction of Lyα emitters from the "MUSE-Wide" survey shows negative velocities[2] up to ~ -300 km s^{-1}, although their systemic redshifts were derived in a different way, namely from a shell model fit of the Lyα line.

The errors on the velocity shifts were evaluated summing quadratically the uncertainties on the determination of the Lyα centroid and the centroid of the CIII(1907.07 Å) emission or IS absorption lines. The individual centroid errors depend on the S/N at the peak of the line, the resolution of the spectrum ($\sim$1300 km s^{-1}) and on a constant that depends on the kind of noise, that we assume 1/1.46 as in Lenz & Ayres (1992). For example, for an average S/N of approximately five both at the peak of the Lyα and at the peak of the CIII(1907.07 Å) or at the bottom of the IS absorption lines, we obtained individual errors of $\sim$80 km s^{-1}, and a final error of $\sim$110 km s^{-1}. We note that for very bright lines, the S/N is higher than five and therefore the uncertainty on the Lyα velocity shift is slightly lower.

3.1.3. Lyα spatial extent (Lyα$_{\text{ext}}$)

The Lyα spatial extent was evaluated directly from the two-dimensional spectrum of the sources. We collapsed the two-dimensional spectrum in the wavelength range of the Lyα and then applied a Gaussian fit along the y-axis perpendicular to the

---

[2] We note that Gronke (2017) defines the velocity offset as $\Delta v_{\text{Ly}\alpha} = c\frac{z - z_{\text{Ly}\alpha}}{z + z_{\text{Ly}\alpha}}$, so their positive offsets correspond to negative values according to our definition.
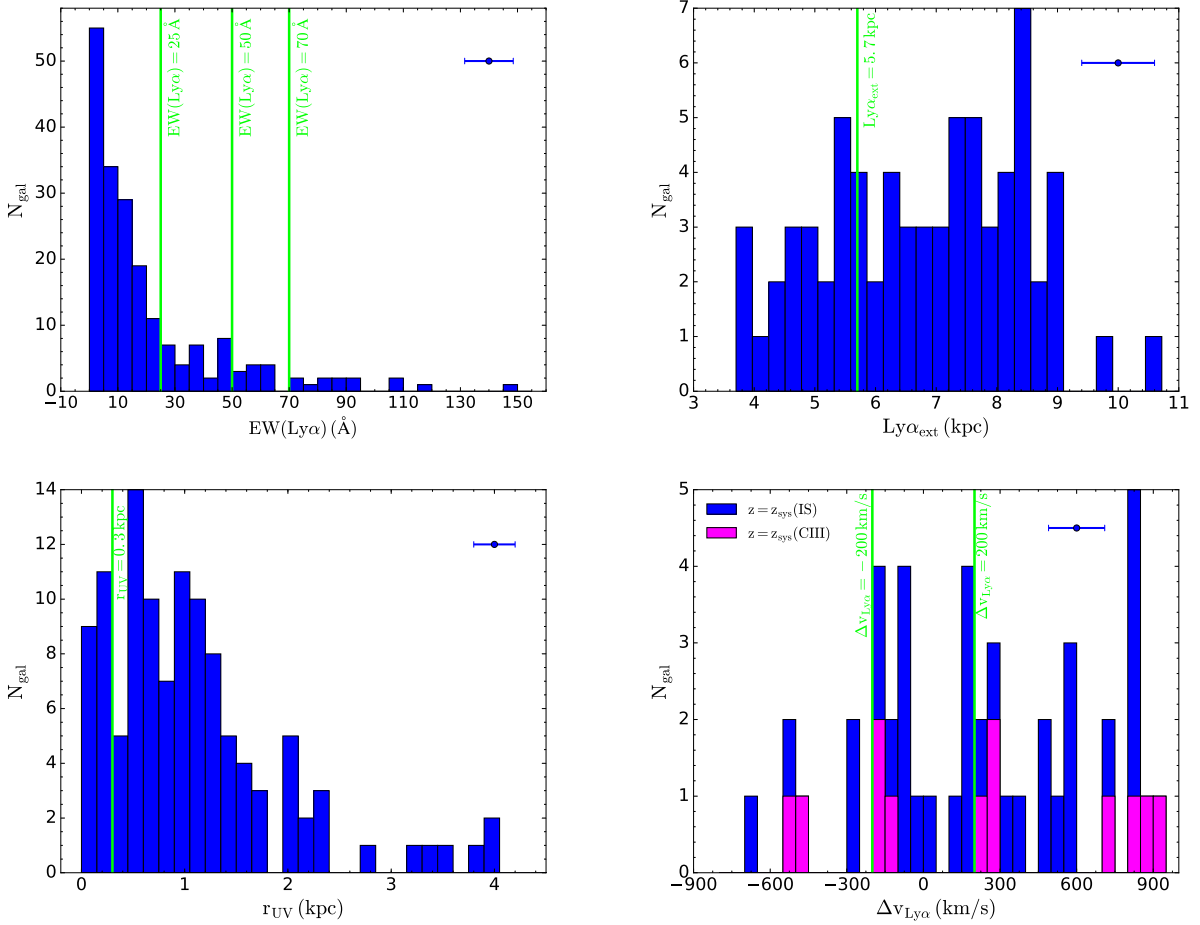
Fig. 2. Distributions of the parameters that we analysed in this work (see Sect. 3.1). The green vertical lines correspond to the cuts that we applied to select the sub-samples (see Sect. 3.3). We note that we left out one source with $EW(Ly\alpha) = 269$ Å from the $EW(Ly\alpha)$ histogram for clarity. We also colour coded the distribution of the Lyα velocity offset according to the line used to measure z_{sys}: magenta indicates z_{sys} from the CIII(1907.07 Å) line and blue indicates the z_{sys} from the interstellar absorption lines. In each panel we show the average error on the relevant parameter in the upper right. We note that for the $EW(Ly\alpha)$ we show the error for $EW(Ly\alpha) = 50$ Å; for the Lyα_{ext} we plot for reference an error of 0.6 kpc, which corresponds to a 10% error of the value used to divide the samples.

spectral dispersion. The Lyα spatial extent was then evaluated as the FWHM of the best fit. We could measure this parameter for the 70 galaxies with $EW(Ly\alpha) > 10$ Å that did not present any skylines close to the position of the Lyα in the spectrum. For the objects with lower EW(Lyα) the S/N was not sufficient to derive the measurement. We note that our estimates of the Lyα_{ext} cannot be taken as absolute estimates of this quantity since we did not deconvolve it by the point spread function (PSF) of the observations. However, this is not an issue for our analysis since we are only interested in the relative Lyα extensions and, since the observations were taken under very similar conditions, we can assume that the PSF affects all the sources in the same way. The distribution of the Lyα spatial extent is shown in Fig. 2.

Given that the Lyα_{ext} is measured only on objects with bright Lyα, the main uncertainty involved in its estimate is the choice of the background in the two-dimensional spectrum, before fitting the Gaussian profile. The reason is that this background is estimated on relatively few pixels, due to the limited length of the slits in the spatial direction. We therefore allowed the background to vary within the 2σ error of its mean value for 1000 times, and then fitted again a Gaussian profile to the line. The errors are finally obtained from the dispersion of the FWHM distribution and vary from 0.1 kpc to

a maximum of 1 kpc, with an average value of $\sim$10% of the measured Lyα_{ext}.

3.1.4. UV rest-frame morphologies (r_{UV})

We used the effective radii obtained by Ribeiro et al. (2016) with GALFIT for the rest-frame UV continuum sizes. These measurements were available for the 115 objects, covered by deep I band imaging (HST, F814W) in the COSMOS and ECDFS fields (U-band rest-frame). The distribution of these values is show in Fig. 2. We note that we took only the objects for which GALFIT converged. The errors on r_{UV} were also evaluated with GALFIT and are on average $\sim$0.2 kpc.

In principle also the FWHM of the Lyα line would be measurable from the spectra. However, models predict that the LyC emitters are those with very narrow emission ($\sim$200 km s^{-1} according to Dijkstra et al. 2016), which are not measurable in the low resolution VUDS spectra ($R \sim 230$).

3.2. The escape fraction of LyC photons

The escape fraction of LyC photons is the fraction of ionizing radiation that is able to escape from the galaxy into the IGM without being absorbed, relative to the total number of photons produced (Wyithe & Cen 2007; Wise & Cen 2009b),

and references therein). Determining this quantity, known as the absolute escape fraction, f_{esc}, requires knowledge of the intrinsic number of ionizing photons produced by the galaxy itself. However, the intrinsic spectral energy distribution (SED) of a galaxy is not known a priori, especially in the rest-frame far UV where dust reddening could be severe.

A related quantity, more used in observational studies, is the relative escape fraction, which is defined as (Steidel et al. 2001; Siana et al. 2007):

$$f_{esc}^{rel}(\text{LyC}) = \frac{L_\nu(1470)/L_\nu(895) \cdot f_\nu(895)/f_\nu(1470)}{e^{-\tau_{IGM,z}}}, \qquad (2)$$

where $L_\nu(1470)/L_\nu(895)$ is the ratio of the intrinsic luminosities at 1470 and 895 Å rest frame and $f_\nu(895)/f_\nu(1470)$ is the ratio of the observed flux densities at the same wavelengths. The transmissivity, $e^{-\tau_{IGM,z}}$, takes into account the photoelectric absorption of photons with $\lambda \leq 912$ Å by the IGM and that then depends on the redshift.

The intrinsic luminosity ratio, $L_\nu(1470)/L_\nu(895)$, depends on the physical properties of the galaxies, such as the mean stellar ages, metallicities, stellar initial mass functions (IMFs), and star formation histories (SFHs). Under reasonable assumptions, it is usually taken as between three and five (see Guaita et al. 2016, and the discussion in Sect. 4). The transmissivity can be estimated by simulating several absorbers in different lines of sight and evaluating the mean intergalactic attenuation curve (Inoue et al. 2014; Worseck et al. 2014). We highlight here that there is a large scatter in the IGM transmission around the mean at each given redshift, as shown, for example, in Fig. 2 of Vanzella et al. (2015) and as computed directly from spectral fitting (Thomas et al. 2017).

The quantity that we can measure directly from the VUDS spectra is the flux density ratio $\left(\frac{f_\nu(895)}{f_\nu(1470)}\right)$ that we use in this paper as an indicator of LyC emission. We explore the caveats of this assumption in Sect. 4.

Since the single spectra have low S/N at LyC wavelengths and cannot give precise values of the LyC signal, to increase the sensitivity of our measures, we grouped the galaxies in different sub-samples according to the properties derived in the previous section, and then, for each sub-sample, we produced stacked spectra where we measured directly the ratio between the LyC and the UV continuum density fluxes, $\frac{f_\lambda(895)}{f_\lambda(1470)}$, as explained in Sect. 3.4. We note that in our previous paper (Marchi et al. 2017), we tried to evaluate the flux density ratio from the individual spectra and obtained a tentative trend between the flux density ratio and the Lyα EW for the galaxies in our sample. However, those values were only upper limits and had large errors. For this reason, we do not attempt here to carry out the same procedure, but instead rely on the stacks of sub-samples selected according to the observed properties to have better estimates of this quantity.

3.3. Definition of the sub-samples

In this section we explain the criteria that we used to select the different sub-samples. Given that the properties we want to explore were measured in different subsets (e.g., $\Delta v_{Ly\alpha}$ in 45 galaxies, r_{UV} in 115 objects and so on), the division in sub-samples according to the galaxies' properties is also different.

EW(Lyα) sub-samples. Theoretical predictions and observational studies in the local Universe (e.g. Dijkstra et al. 2016; Verhamme et al. 2017) suggest that the escape fractions of Lyα and LyC photons are correlated. Since high Lyα escape fractions

generally imply high Lyα EWs, we can use this quantity to test the correlation between Lyα and LyC radiation. We measured the EW(Lyα) for all the 201 galaxies in the final sample. For this reason, we could study the dependence of the flux density ratio with the EW(Lyα) imposing different cuts on this quantity and see how the LyC emission changes increasing the EW(Lyα) cut. We start with a cut at $EW(Ly\alpha) = 25$ Å, to include in the sub-sample all the LAEs, and then we increase it to $EW(Ly\alpha) = 50$ Å and finally to $EW(Ly\alpha) = 70$ Å. We therefore selected and stacked all the galaxies with $EW(Ly\alpha) \geq 25$ Å, $EW(Ly\alpha) \geq 50$ Å and $EW(Ly\alpha) \geq 70$ Å and finally compared the values of $\frac{f_\lambda(895)}{f_\lambda(1470)}$ obtained for these samples with those of their complementaries.

$\Delta v_{Ly\alpha}$ sub-samples. According to theoretical predictions (Dijkstra et al. 2016; Verhamme et al. 2015), the galaxies with a higher probability of leaking LyC emission, have a Lyα emission line that emerges very close to the systemic redshift (for example, $\Delta v_{Ly\alpha} < 150$ km s^{-1} according to Verhamme et al. 2015). We therefore selected and stacked all the sources with Lyα velocity offset around zero, $-200 \leq \Delta v_{Ly\alpha} \leq 200$ km s^{-1}, and those with $\Delta v_{Ly\alpha} > 200$ km s^{-1} and $\Delta v_{Ly\alpha} < -200$ km s^{-1}. We had to select galaxies in a larger range of $\Delta v_{Ly\alpha}$ with respect to the predicted values because we would have had too few galaxies in the interval $|\Delta v_{Ly\alpha}| < 150$ km s^{-1} to perform a reliable stack.

Lyα_{ext} sub-samples. We did not find in the literature any study focused on the relation between the Lyα emission size and the presence of LyC radiation, that could drive our sample division. The distribution of the observed sizes is also more or less flat between 4 and 9 kpc (see Fig. 2). Recently Yang et al. (2017) found an anti-correlation between the Lyα extension and the Lyα escape fraction. Since we expect a correlation between Lyα and LyC escape fractions (Dijkstra et al. 2016), we suppose that the Lyα compactness could favour the escape of LyC radiation. To investigate this scenario, we selected the 20 most compact sources in Lyα to see if this sample has a higher LyC emission compared to the galaxies with a larger Lyα extension. We selected a sub-sample with Ly$\alpha_{ext} \leq 5.7$ kpc and a complementary sample with Ly$\alpha_{ext} > 5.7$ kpc. We chose 20 objects because it is the minimum number required to produce a stack not dominated by the errors.

r_{UV} sub-samples. Izotov et al. (2016) demonstrated that a selection for UV compact objects, along with a selection for high [OIII]λ5007/[OII]λ3727 emitters, appears to pick up very efficiently LCEs in the local Universe. We want to test this correlation at higher redshifts. The only indication that we have at $z \sim 3$ is the size of the well known LCE *Ion2* (Vanzella et al. 2016; de Barros et al. 2016), with $r_{UV} = 0.3$ kpc. Since we have a sufficient number of objects with $r_{UV} \leq 0.3$ kpc (20 sources), we decided to define this as a sub-sample. To investigate the dependence between the UV radius and the LyC signal, we compared its flux density ratio with that of the complementary sample ($r_{UV} > 0.3$ kpc).

We summarize in Table 1 the division in different sub-samples and complementary samples, showing also the number of galaxies in each sample. In Fig. 2 we also indicate the sub-sample division.

3.4. Stacking procedure and $f_\lambda(895)/f_\lambda(1470)$ evaluation

We followed the same stacking procedure described in Marchi et al. (2017). We first shifted each one-dimensional spectrum to its rest-frame and normalized it using its mean value in the wavelength range 1420−1520 Å where no particular features are

Table 1. Number of sources in each sub-sample considered.

Sub-sample	Number of sources	Complementary sample	Number of sources
$EW(\mathrm{Ly}\alpha) \geq 25\,\text{Å}$	52	$EW(\mathrm{Ly}\alpha) < 25\,\text{Å}$	149
$EW(\mathrm{Ly}\alpha) \geq 50\,\text{Å}$	25	$EW(\mathrm{Ly}\alpha) < 50\,\text{Å}$	176
$EW(\mathrm{Ly}\alpha) \geq 70\,\text{Å}$	14	$EW(\mathrm{Ly}\alpha) < 70\,\text{Å}$	187
$\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$	20	$\mathrm{Ly}\alpha_{\text{ext}} > 5.7\,\text{kpc}$	50
$|\Delta v_{\mathrm{Ly}\alpha}| \leq 200\,\text{km s}^{-1}$	17	$|\Delta v_{\mathrm{Ly}\alpha}| > 200\,\text{km s}^{-1}$	28
$r_{UV} \leq 0.3\,\text{kpc}$	20	$r > 0.3\,\text{kpc}$	95

Notes. (1) Selection criterion for the sub-sample; (2) number of sources in the sub-sample; (3) selection criterion for the complementary sub-sample; (4) number of sources in the complementary sub-sample.

present. To bring the spectra rest-frame, we used the systemic redshifts, for the sources that presented inter-stellar absorption lines or the CIII(1907.07 Å) line (45 galaxies, see Sect. 3.1), and the VUDS official redshifts for the other sources (see Le Fèvre et al. 2015, for details on the redshift evaluation). To take into account the noise of each spectrum, we computed the stack as a weighted average of the normalized spectra in each sub-sample, using the statistical errors of the individual flux density ratios $\frac{f_\lambda(895)}{f_\lambda(1470)}$ as weights[3]. To make the average, we resampled each spectrum to the same grid that goes from 870 Å to 1700 Å with a step of $5.355/(1 + z_{\text{median}})$. 5.355 is the nominal VIMOS Low Resolution Å/pixel scale of the observed-frame spectra, and z_{median} is the median redshift of the sub-sample.

The flux density ratio $f_\lambda(895)/f_\lambda(1470)$ can be measured directly from the stack, averaging the signal in the interval 870−910 Å. However, given the very small size of some of our sub-samples, taking this value simply from the stack would not give a reliable estimate of the flux density ratio and its error. For this reason, we used the bootstrapping technique to estimate the uncertainty on this quantity. Using this approach does not require making any assumptions on the distribution of our data. For each sub-sample of N galaxies, we therefore created 5000 realizations of it, randomly extracting N galaxies with replacement. We then stacked each realization and computed the flux density ratio. We finally evaluated the mean and 68% confidence level of each $f_\lambda(895)/f_\lambda(1470)$ distribution.

4. Results

The flux density ratios representative of each sub-sample and of the total sample of 201 galaxies, measured as explained in Sect. 3.4 and converted from f_λ to f_ν, are shown in Fig. 3. We expect that the average signal of LyC is almost washed away in the total sample, given that this is probably dominated by sources with no LyC leakage, and we assume, for the moment, that the signal $\frac{f_\nu(895)}{f_\nu(1470)}$ observed in the total sample basically comes from contamination from low redshift inter-lopers. This will be further discussed in Sect. 5.

We can see from Fig. 3 that most of the sub-samples (blue dots) show a $\frac{f_\nu(895)}{f_\nu(1470)}$ much larger than the 1σ interval obtained for the entire sample (lavender band) and that their complementary samples (magenta dots) are instead in agreement with it. The only exceptions are the $\Delta v_{\mathrm{Ly}\alpha}$ samples that we better discuss later.

The parameters which appear to (anti)correlate the strongest with the flux density ratio are $EW(\mathrm{Ly}\alpha)$, $\mathrm{Ly}\alpha_{\text{ext}}$ and r_{UV}. The higher is the Lyα equivalent width, the higher is the flux density ratio, and the smaller are the Lyα spatial extent and the UV effective radius, the higher is the LyC radiation. In particular, the sub-samples with $\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$ and $r_{UV} \leq 0.30\,\text{kpc}$ have values for the flux density ratio that are more than 2σ higher than their complementaries.

We note that the uncertainties in the flux density ratio of some sub-samples are particularly large. This is because some of these sub-samples contain very few sources (in particular the sub-sample with $EW(\mathrm{Ly}\alpha) \geq 70\,\text{Å}$ contains only 14 galaxies).

In Fig. 4 we show the normalized spectral stacks of the sub-samples with $\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$ (cyan line) and with $\mathrm{Ly}\alpha_{\text{ext}} > 5.7\,\text{kpc}$ (purple line) for reference. The excess in the LyC range is clearly visible even from the simple stack in the sub-sample with small $\mathrm{Ly}\alpha_{\text{ext}}$, while a much lower signal is present in the stack of the sub-sample with higher $\mathrm{Ly}\alpha_{\text{ext}}$ in this range. We emphasize that we were looking for a differential LyC signal between the sub-samples, because at this stage, we were not taking into account the contamination from lower-redshift interlopers and we are assuming the same statistical contamination in all the sub-samples.

The spectral stacks for the three most significant parameters ($EW(\mathrm{Ly}\alpha) \geq 70\,\text{Å}$, $\mathrm{Ly}\alpha_{\text{ext}} \leq 5.7\,\text{kpc}$ and $r_{UV} \leq 0.30\,\text{kpc}$) are not independent, since many objects belong to more than one of these sub-samples. We show in Fig. 5 the Venn diagram with the number of sources contained in each sub-sample and the sources in common. We could not measure all the parameters for all the galaxies in the final sample, so the sources in common between the three sub-samples could be more than the present numbers. We are not able to study the LyC properties of the sources in common between the sub-samples because the number of galaxies that satisfy all the three conditions is too low (see Fig. 5) and the stack is dominated by noise.

To test the correlation between the different parameters, we evaluated the Spearman's rank correlation co-efficients between the sub-samples for which we measured EW(Lyα), $\mathrm{Ly}\alpha_{\text{ext}}$ and r_{UV}. We found that the most correlated parameters are, as expected, the Lyα spatial extension and the UV effective radius, with a coefficient of 0.32, that implies that the smaller is r_{UV}, the smaller is $\mathrm{Ly}\alpha_{\text{ext}}$. We also found a relatively weak anti-correlation between the Lyα spatial extension and the EW(Lyα), in the sense that the galaxies with lower EW(Lyα), present a more extended Lyα profile. In principle this weak trend could be due to an improper extraction of the one-dimensional spectra for objects with large Lyα spatial extension because the extraction aperture is determined from the UV portion of the spectra. Therefore, for extended Lyα emission, part of the Lyα

---

[3] The errors have been evaluated with the classic errors propagation from the standard deviations around the mean values in the LyC and UV ranges (870−910 Å and 1420−1520 Å, respectively).
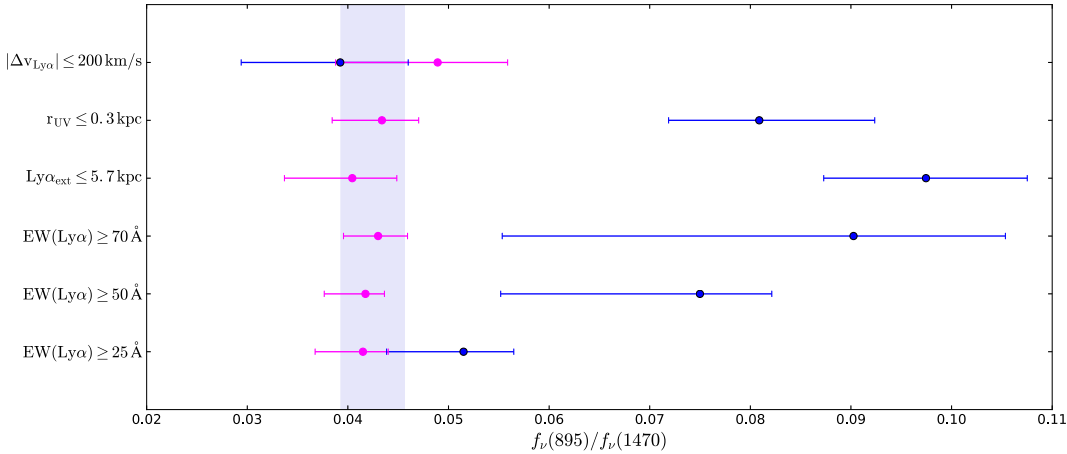
Fig. 3. Flux density ratios evaluated from the stacks of the samples in the y-axis (blue dots) and from the complementary samples (magenta dots) as indicated in Table 1. The lavender vertical band is the 1σ confidence interval evaluated for the total sample of 201 galaxies.
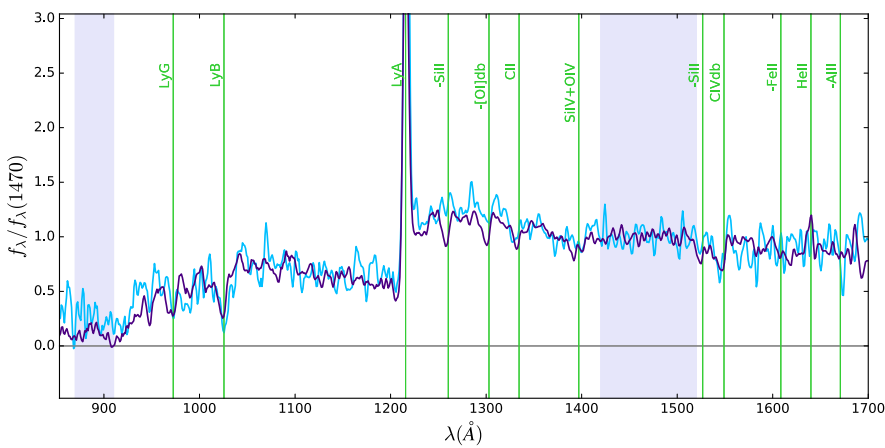

Fig. 4. Comparison of the spectral stacks of the sub-samples with $Ly\alpha_{ext} \leq 5.7$ kpc (cyan line) and with $Ly\alpha_{ext} > 5.7$ kpc (purple line). The vertical lavender bands indicate the LyC range (870−910 Å) and the UV range (1420−1520 Å) where we have normalized each spectrum. The signal in the LyC range in the stack of the sub-sample with $Ly\alpha_{ext} \leq 5.7$ kpc is about 2.5 times higher than that of the sub-sample with $Ly\alpha_{ext} > 5.7$ kpc. The spectra in the figure have been smoothed by three times the step of the stacked spectrum, which is 1.06 Å for the stack of the sub-sample with $Ly\alpha_{ext} \leq 5.7$ kpc and 1.12 Å for the sub-sample with $Ly\alpha_{ext} > 5.7$ kpc, to emphasize the difference in the LyC region between the two sub-samples.

flux could be lost and lower EW could be measured from the one-dimensional spectrum. However, we checked that this is not the case for our galaxies by re-extracting the spectra directly from the original two-dimensional frames for a sub-sample with large $Ly\alpha$ extension. In addition, a similar result was also found independently by Momose et al. (2016). Finally, we also observed a weak anti-correlation between $Ly\alpha$ EW and the UV size in agreement with Law et al. (2012). All the three parameters, therefore, seem to show some level of correlation.

So far we have compared the flux density ratio of the different sub-samples and not their relative escape fraction of LyC photons, that depends also on the intrinsic luminosity ratio and on the mean trasmissity of the sub-sample considered (see Eq. (2)). We could reasonably conclude that galaxies in the significant sub-samples likely emit more ionizing radiation, only if the intrinsic luminosity ratio and the transmissivity do not intervene in changing our results. The transmissivity depends on the redshift of the sources (see Sect. 3.2), and since our sub-samples have very similar median redshifts, the transmissivity is approximately the same (within 10%) for all the sub-samples. The intrinsic luminosity ratio depends instead on different galaxies' properties, in particular the age of the stellar populations (see for example Table 3 in Guaita et al. 2016). As evaluated in Grazian et al. (2016), for typical star-forming galaxies at $z \sim 3$, it varies between 1.7 and 7.1 for ages between 1 Myr and 0.2 Gyr, adopting the Bruzual & Charlot (2003) library. There is no evidence to date that galaxies that are UV compact or have a $Ly\alpha$ with a small spatial extent, have younger ages, so we expect that our sub-samples selected according to these two quantities have approximately the same age. For these sub-samples

we can therefore interpret the differences in the measured density flux ratios in terms of relative escape fraction. On the other hand, galaxies with a very high $Ly\alpha$ EW are generally believed to be younger than the rest of the star-forming galaxies population (e.g. Malhotra & Rhoads 2002; Gawiser et al. 2007) and have therefore lower intrinsic luminosity ratios. For this reason, the higher flux density ratio observed in the sub-sample with $EW(Ly\alpha) \geq 70$ Å with respect to its complementary, does not necessarily have to be due entirely to a higher relative escape fraction. Since we have shown that the three parameters are marginally correlated this could be also partially true for the other sub-samples: however, the correlation between $Ly\alpha$ EW and the spatial extent (both UV and $Ly\alpha$) is quite weak. We note that the differences between the flux density ratios obtained for the interesting sub-samples and those obtained for the complementaries are too large to be entirely erased by a change in the intrinsic luminosity ratio between the sub-samples.

Finally, we tested that the errors on the individual measurements ($EW(Ly\alpha)$, $Ly\alpha_{ext}$ and r_{UV}) do not influence the results of the stacks, by performing Monte Carlo simulations. For each parameter, we created 100 different versions of the original subset by varying each real measured value within a Gaussian distribution with the error on the measure as sigma. For each of these new subsets, we then re-selected the interesting sub-samples (e.g. $EW(Ly\alpha) \geq 70$Å), bootstrapped them and evaluated the mean flux density ratio. For the $Ly\alpha$ EW and the UV effective radius, we obtained flux density ratio distributions that were completely inside the intervals shown in Fig. 3. For the $Ly\alpha$ spatial extent, we obtained a distribution that is slightly larger than the flux density ratio interval

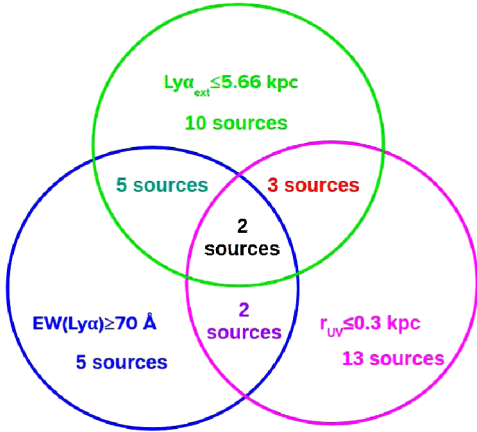**Fig. 5.** Venn diagram showing the number of sources contained in each significant sub-sample ($EW(\text{Ly}\alpha) \geq 70$ Å, $\text{Ly}\alpha_{\text{ext}} \leq 5.7$ kpc and $r_{UV} \leq 0.30$ kpc) and the sources in common.

shown in Fig. 3, even though it is still much higher than the complementary sample. This test proves the validity of our sub-sample division, even considering the uncertainties on the parameters.

The results obtained for the velocity shift of the Lyα with respect to the systemic redshift, are not in agreement with theoretical expectations (Dijkstra et al. 2016; Verhamme et al. 2015). This could be for several reasons: First, our evaluation of the systemic redshift could be wrong, since it relies in most cases on an average relation that was tested only at lower redshifts and the uncertainties on the measurements are relatively high. Second, we know that in ~30% of the galaxies the Lyα emission is actually double peaked (e.g. Yamada et al. 2012; Kulas et al. 2012), but the peak separation is smaller than the VUDS resolution and we might be underestimating the velocity of the main (red) peak. Therefore the division in sub-samples according to the velocity offset could be incorrect. Last, we cannot entirely discard the possibility that there is no real relation between the leakage of LyC emission and the velocity offset of the Lyα line, contrary to model predictions.

5. Estimating $f_{\text{esc}}^{\text{rel}}$ after correction from contamination

As explained in Sect. 2, our sample is most probably contaminated, in the sense that some of the objects selected could have nearby lower redshift faint galaxies (interlopers) that contribute to the flux in the LyC range of the extracted spectrum. For this reason, it would be incorrect to simply transform the values of the flux density ratio derived above into a LyC escape fraction. Since most of the galaxies in our sample (~165 galaxies) do not have multi-band high resolution HST imaging, the careful cleaning procedure performed in Marchi et al. (2017) and Guaita et al. (2016) is not possible. In principle, we do not expect any correlation between contamination and the physical properties analysed in this work. If anything, we would expect a higher probability of contamination in objects with a more extended profile in the UV rest-frame than in compact objects, while the observations of the flux in the 895 Å region indicate the contrary if this was completely a contamination effect. In any case, to transform the results obtained in the previous section into relative escape fractions of LyC photons, we need to estimate the amount of

flux that could come from the lower redshift interlopers in each sub-sample.

One way to do this is to perform simulations, following the method outlined in Vanzella et al. (2010). The procedure is fully explained in Appendix A: briefly, we used a very deep U-band image of the CDFS field (Nonino et al. 2009), which corresponds to the LyC flux rest-frame for sources at redshift of approximately four, to determine the expected average integrated contribution of the foreground blue sources, by placing different rectangular apertures (corresponding to our spectroscopic slits) and performing Monte Carlo simulations. We find that the median value of the simulated flux, which corresponds to the contamination, is very similar for all of our sub-samples, depending only slightly on the sample-size, in the sense that larger samples have slightly larger median contamination, thus validating our previous assumptions. On the other hand, the estimated contamination depends strongly on the maximum magnitude of possible contaminants that we set in the simulations (U_{max}, see Sect. 2).

Adopting a conservative approach and setting $U_{\text{max}} = 25$, we proceed to estimate the relative escape fraction of LyC photons of the significant sub-samples, i.e., the samples where we believe there is real LyC signal in the 895 Å region. We therefore ran our simulations for the sub-samples selected as $EW(\text{Ly}\alpha) \geq 70$ Å, $\text{Ly}\alpha_{\text{ext}} \leq 5.7$ kpc and $r_{UV} \leq 0.30$ kpc to obtain the contamination. As an example, we show in Fig. 6 the distribution of the simulated fluxes for the sub-sample with $\text{Ly}\alpha_{\text{ext}} \leq 5.7$ kpc. We show both the distribution obtained assuming $U_{\text{max}} = 25$ (magenta histogram) and that assuming $U_{\text{max}} = 26$ (blue histogram). In both cases the observed flux, which is the symbol in the figure, is much higher than the simulated value, implying that there is probably real LyC escaping the galaxies in the sub-sample.

To evaluate the average relative escape fraction, we first evaluated the presumed LyC flux of our sub-samples, $f_\nu(\text{LyC})$, subtracting the 2σ percentile of the distribution obtained with the simulations from the observed value, $f_\nu^{\text{obs}}(895)$, and then we estimated the resulting flux density ratio, $\frac{f_\nu(\text{LyC})}{f_\nu(1470)}$. To convert it to an escape fraction, we evaluated the transmissivity averaging the individual galaxy's transmissivities of each sub-sample (0.25 for the $EW(\text{Ly}\alpha) \geq 70$ Å sub-sample, 0.22 for the $\text{Ly}\alpha_{\text{ext}} \leq 5.7$ kpc sub-sample and 0.28 for the $r_{UV} \leq 0.30$ kpc sub-sample). The individual values have been evaluated using the analytical prescription given by Inoue et al. (2014). For easier comparison with earlier studies (e.g., Steidel et al. 2001; Grazian et al. 2016, 2017; Marchi et al. 2017), we adopted a value of $L_\nu(1470)/L_\nu(895) = 3$, which corresponds to young star-forming galaxies with age ~10 Myr, assuming a constant Bruzual & Charlot SFH and a Chabrier IMF (Chabrier 2003). We point out that this is simply a multiplicative factor in the evaluation of the relative escape fraction, and it is therefore possible to re-scale $f_{\text{esc}}^{\text{rel}}$ with other values of $L_\nu(1470)/L_\nu(895)$ if needed.

We obtain upper limits of $f_{\text{esc}}^{\text{rel}} \sim 33\%$ for the sub-sample with $EW(\text{Ly}\alpha) \geq 70$ Å (67% if we subtract the 1σ percentile of the simulated distribution), $f_{\text{esc}}^{\text{rel}} \sim 48\%$ for the sub-sample with $\text{Ly}\alpha_{\text{ext}} \leq 5.7$ kpc (85% subtracting 1σ) and $f_{\text{esc}}^{\text{rel}} \sim 23\%$ for the sub-sample with $r_{UV} \leq 0.30$ kpc (50% subtracting 1σ). Clearly these values are just indicative, since they crucially depend on the real contamination that is present in each sub-sample. Given the small sizes of these significant sub-samples (14, 20 and 20 objects respectively) and the fact that there are several objects in common, even the presence of a single contaminated source could sensibly change the results. In addition, as evident from
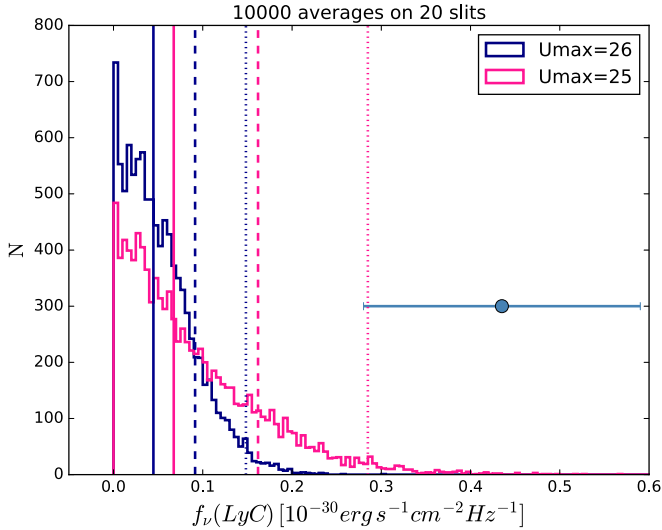
Fig. 6. Distribution of the simulated observed-frame $f_\nu^{\rm sim}$ coming from foreground sources for the sub-sample with $Ly\alpha_{\rm ext} \leq 5.7$ kpc imposing $U_{\rm max} = 25$ (magenta histogram) and $U_{\rm max} = 26$ (blue histogram). The continuum vertical lines are the median of the distributions and the dashed and dotted lines are the 1σ and 2σ confidence intervals, respectively. The symbol in the figure is the average observed-frame $f_\nu^{\rm obs}(895)$ obtained for the sub-sample.

Fig. A.1, with this choice of $U_{\rm max}$ and slits size in the simulations, we are probably underestimating the contamination since the flux obtained from the simulations does not account for the flux in the full sample stack. We note however, that if instead of using the simulations to estimate the contamination, we simply assumed that the full sample of 201 galaxies is dominated by contamination and subtract its flux density ratio from the significant sub-samples, we obtain similar high values for $f_{\rm esc}^{\rm rel}$ for the three samples above.

The values we find for $f_{\rm esc}^{\rm rel}$ are quite high and could seem at odds with observational results finding lower average values or very stringent upper limits (Boutsia et al. 2011; Grazian et al. 2016; Guaita et al. 2016): however the significant sub-samples represent only less than 10% of the total galaxy population. If the galaxies with high LyC emission are indeed only a very small fraction of the entire star forming population, we do not expect to detect significant emission when stacking large samples of objects without any pre-selection.

6. Summary and conclusions

In this work we have analysed some of the proposed correlations between LyC emission and other galaxies properties at high redshifts, exploiting the high quality and large number of spectra coming from the VUDS survey. We initially selected star-forming galaxies from the VUDS dataset at $3.5 \leq z \leq 4.3$, paying particular attention in retaining only galaxies with a clean spectrum in the LyC region and galaxies with no clear contamination from bright neighbours in the same slit. For each of the 201 selected galaxies, we then evaluated (if possible) the $Ly\alpha$ EW, the $Ly\alpha$ velocity offset with respect to the systemic redshift, the $Ly\alpha$ spatial extension and the UV effective radius. Unfortunately, we could not estimate all these parameters for all the galaxies in the sample, because for example in some cases an evaluation of the systemic redshift was not possible, or the lack of high resolution HST data did not allow us to estimate the

UV radius. To analyse the correlations between these parameters and LyC emission, we defined different sub-samples according to the properties predicted to be good LyC emission indicators. We therefore selected the most compact galaxies in $Ly\alpha$ ($Ly\alpha_{\rm ext} \leq 5.7$ kpc) and UV ($r_{UV} \leq 0.30$ kpc), the galaxies with the highest $EW(Ly\alpha)$ ($EW(Ly\alpha) \geq 70$ Å) and those with the $Ly\alpha$ peak closest to the systemic redshift ($-200 \leq \Delta v_{Ly\alpha} \leq 200$ km s^{-1}). Since we cannot reliably measure the flux in the LyC range for individual objects, we created spectral stacks of all the galaxies in each sub-sample and estimated the average flux density ratios $\left(\frac{f_\nu(895)}{f_\nu(1470)}\right)$ from the stacks. Our main results are:

- We find that galaxies which are UV compact ($r_{UV} \leq 0.30$ kpc) and have a high $Ly\alpha$ emission ($EW(Ly\alpha) \geq 70$ Å) are likely to have higher LyC emission, since these stack show a significant excess of flux in the LyC range compared to the complementary samples. This is in agreement with theoretical studies (Dijkstra et al. 2016; Verhamme et al. 2015; Wise & Cen 2009a) and with previous observational studies at low redshift (Izotov et al. 2016; Verhamme et al. 2017). An indication that a high $Ly\alpha$ emission was related to the presence of LyC emission was also found in our recent work (Marchi et al. 2017).

- We find that galaxies with a small $Ly\alpha$ spatial extent ($Ly\alpha_{\rm ext} \leq 5.7$ kpc) have much larger fluxes in the LyC range, compared to the complementary samples. A possible relation between the spatial extent of $Ly\alpha$ and the presence of LyC emission has never been studied before. According to our data, this parameter might even be more correlated to the presence of LyC emission than the other parameters, since the stack of the spatially small emitters shows a flux in the LyC range that is the highest of all the sub-samples.

- Although the above three parameters are correlated at some level, there are only very few objects in which all three conditions ($r_{UV} \leq 0.30$ kpc, $EW(Ly\alpha) \geq 70$ Å and $Ly\alpha_{\rm ext} \leq 5.7$ kpc) are met. Part of the reason for this is that we could not measure all parameters for all galaxies in the initial sample. These few objects are the best candidates for being real LyC emitters. However, since they are so few, their spectral stack is dominated by noise in the LyC range.

- We could not apply any cleaning procedure to exclude contamination from lower-redshift interlopers in our samples because multi-wavelength HST imaging, needed for identifying the contaminants, is not available for most of the sources. We believe that the contamination should be approximately the same in all the sub-samples, given that it should not correlate with any of the galaxy properties. We attempted to estimate (and subtract) a statistical contamination to the LyC flux using Monte Carlo simulations performed on a very deep U band image of the ECDFS field, which covers the observed wavelength of the LyC emission at $z \sim 4$. We find that it is difficult to give an accurate estimate of the real contamination with the simulations because of the uncertainties involved, especially in the choice of the parameters to reproduce the observations ($U_{\rm max}$, see Appendix A.1). Also, the small number of galaxies in the significant sub-samples means that the simulated contamination flux has a very large distribution. For a reasonable choice of parameters in the simulation, we find that comparing the observed $f_\nu(895)f_\nu(1470)$ in the three significant sub-samples to the simulated ones, there is a very high probability that a significant fraction of the $f_\nu(895)$ flux comes from real LyC leakage, resulting in large escape fractions for the

galaxies with very high $EW(\mathrm{Ly}\alpha)$, small r_{UV} and small $\mathrm{Ly}\alpha_{\mathrm{ext}}$.

The physical picture that emerges from our results is that the conditions that regulate the escape of LyC and Lyα photons, must be closely related. Galaxies with compact UV morphologies and compact and strong Lyα emission are the sources that most likely show LyC emission. This is in agreement with the scenario proposed by Nakajima & Ouchi (2014) where compact star-forming regions can photoionize the ISM, creating density-bounded regions or optically thin paths through the ISM. These paths could be seen as "holes" in the ISM caused for example by supernovae driven winds (Dove et al. 2000; Sharma et al. 2017), and would allow the simultaneous escape of LyC and Lyα radiation. The natural consequence would be the observed strong correlation between Lyα EW and LyC emission. In addition, small HI column densities cause less scattering of the Lyα photons through the ISM, resulting in more compact Lyα spatial profiles (Yang et al. 2017). This would produce the observed anti-correlation between Lyα size and LyC emission.

References

Alexandroff, R. M., Heckman, T. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 810, 104
Becker, G. D., & Bolton, J. S. 2013, MNRAS, 436, 1023
Bergvall, N., Zackrisson, E., Andersson, B.-G., et al. 2006, A&A, 448, 513
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. 2014, Science, 346, 216
Boutsia, K., Grazian, A., Giallongo, E., et al. 2011, ApJ, 736, 41
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, ApJ, 754, 83
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Cassata, P., Tasca, L. A. M., Le Fèvre, O., et al. 2015, A&A, 573, A24
Chabrier, G. 2003, PASP, 115, 763
de Barros, S., Vanzella, E., Amorín, R., et al. 2016, A&A, 585, A51
Deharveng, J.-M., Buat, V., Le Brun, V., et al. 2001, A&A, 375, 805
Dijkstra, M., Haiman, Z., Rees, M. J., & Weinberg, D. H. 2004, ApJ, 601, 666
Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, ApJ, 828, 71
Dove, J. B., Shull, J. M., & Ferrara, A. 2000, ApJ, 531, 846
Finkelstein, S. L., Rhoads, J. E., Malhotra, S., Pirzkal, N., & Wang, J. 2007, ApJ, 660, 1023
Garilli, B., Fumana, M., Franzetti, P., et al. 2010, PASP, 122, 827
Gawiser, E., Francke, H., Lai, K., et al. 2007, ApJ, 671, 278
Giallongo, E., Grazian, A., Fiore, F., et al. 2015, A&A, 578, A83
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
Gronke, M. 2017, A&A, 608, A139
Gronke, M., & Dijkstra, M. 2014, MNRAS, 444, 1095
Guaita, L., Francke, H., Gawiser, E., et al. 2013, A&A, 551, A93
Guaita, L., Pentericci, L., Grazian, A., et al. 2016, A&A, 587, A133
Hansen, M., & Oh, S. P. 2006, MNRAS, 367, 979
Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, ApJ, 730, 5
Inoue, A. K., & Iwata, I. 2008, MNRAS, 387, 1681

Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, MNRAS, 442, 1805
Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016, MNRAS, 461, 3683
Jaskot, A. E., & Oey, M. S. 2013, ApJ, 766, 91
Kulas, K. R., Shapley, A. E., Kollmeier, J. A., et al. 2012, ApJ, 745, 33
Laursen, P., Sommer-Larsen, J., & Razoumov, A. O. 2011, ApJ, 728, 52
Laursen, P., Duval, F., & Östlin, G. 2013, ApJ, 766, 124
Law, D. R., Steidel, C. C., Shapley, A. E., et al. 2012, ApJ, 759, 29
Le Fèvre, O., Adami, C., Arnouts, S., et al. 2014, The Messenger, 155, 33
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79
Leitet, E., Bergvall, N., Piskunov, N., & Andersson, B.-G. 2011, A&A, 532, A107
Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, A&A, 553, A106
Leitherer, C., Ferguson, H. C., Heckman, T. M., & Lowenthal, J. D. 1995, ApJ, 454, L19
Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, ApJ, 823, 64
Lenz, D. D., & Ayres, T. R. 1992, PASP, 104, 1104
Madau, P. 1995, ApJ, 441, 18
Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648
Malhotra, S., & Rhoads, J. E. 2002, ApJ, 565, L71
Marchi, F., Pentericci, L., Guaita, L., et al. 2017, A&A, 601, A73
Mason, C. A., Trenti, M., & Treu, T. 2015, ApJ, 813, 21
McLinden, E. M., Rhoads, J. E., Malhotra, S., et al. 2014, MNRAS, 439, 446
Micheva, G., Iwata, I., Inoue, A. K., et al. 2017, MNRAS, 465, 316
Mitra, S., Ferrara, A., & Choudhury, T. R. 2013, MNRAS, 428, L1
Momose, R., Ouchi, M., Nakajima, K., et al. 2016, MNRAS, 457, 2318
Mostardi, R. E., Shapley, A. E., Nestor, D. B., et al. 2013, ApJ, 779, 65
Mostardi, R. E., Shapley, A. E., Steidel, C. C., et al. 2015, ApJ, 810, 107
Naidu, R. P., Oesch, P. A., Reddy, N., et al. 2017, ApJ, 847, 12
Nakajima, K., & Ouchi, M. 2014, MNRAS, 442, 900
Nestor, D. B., Shapley, A. E., Steidel, C. C., & Siana, B. 2011, ApJ, 736, 18
Nonino, M., Dickinson, M., Rosati, P., et al. 2009, ApJS, 183, 244
Ouchi, M., Mobasher, B., Shimasaku, K., et al. 2009, ApJ, 706, 1136
Pentericci, L., Grazian, A., Scarlata, C., et al. 2010, A&A, 514, A64
Prochaska, J. X., Worseck, G., & O'Meara, J. M. 2009, ApJ, 705, L113
Ribeiro, B., Le Fèvre, O., Tasca, L. A. M., et al. 2016, A&A, 593, A22
Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J., & Stark, D. P. 2010, Nature, 468, 49
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJ, 802, L19
Rutkowski, M. J., Scarlata, C., Henry, A., et al. 2017, ApJ, 841, L27
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, ApJ, 826, L24
Sharma, M., Theuns, T., Frenk, C., et al. 2017, MNRAS, 468, 2176
Shull, J. M., Harness, A., Trenti, M., & Smith, B. D. 2012, ApJ, 747, 100
Siana, B., Teplitz, H. I., Colbert, J., et al. 2007, ApJ, 668, 62
Song, M., Finkelstein, S. L., Gebhardt, K., et al. 2014, ApJ, 791, 3
Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408, 1628
Steidel, C. C., Pettini, M., & Adelberger, K. L. 2001, ApJ, 546, 665
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Tasca, L. A. M., Le Fèvre, O., Ribeiro, B., et al. 2017, A&A, 600, A110
Thomas, R., Le Fèvre, O., Le Brun, V., et al. 2017, A&A, 597, A88
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
Vasei, K., Siana, B., Shapley, A. E., et al. 2016, ApJ, 831, 38
Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, A&A, 578, A7
Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, A&A, 597, A13
Vollmann, K., & Eversberg, T. 2006, Astron. Nachr., 327, 862
Wise, J., & Cen, R. 2009a, in AAS Meeting Abstracts #213, Bull. Am. Astron. Soc., 41, 484
Wise, J. H., & Cen, R. 2009b, ApJ, 693, 984
Worseck, G., Prochaska, J. X., O'Meara, J. M., et al. 2014, MNRAS, 445, 1745
Wyithe, J. S. B., & Cen, R. 2007, ApJ, 659, 890
Yajima, H., Choi, J.-H., & Nagamine, K. 2011, MNRAS, 412, 411
Yamada, T., Matsuda, Y., Kousai, K., et al. 2012, ApJ, 751, 29
Yang, H., Malhotra, S., Gronke, M., et al. 2017, ApJ, 844, 171
Zackrisson, E., Inoue, A. K., & Jensen, H. 2013, ApJ, 777, 39

Appendix A: Statistical contamination from lower-redshift interlopers

A.1. Method

To simulate the contribution of low redshift interlopers to the flux in the LyC range we followed the method outlined in Vanzella et al. (2010). We started from a very deep *U*-band image of the ECDFS field (Nonino et al. 2009), which corresponds to the LyC flux rest-frame for sources at a redshift of approximately four as those in our sample, and we calculated the expected average integrated contribution of the foreground blue sources placing different rectangular apertures (corresponding to our spectroscopic slits) and performing Monte Carlo simulations. Using the SExtractor algorithm (Bertin & Arnouts 1996), we initially detected all sources down to a magnitude limit of about 30 and then generated a new *U*-band image composed only by the detected sources separated by null pixels requiring the SExtractor checkimage "OBJECT". We did this to avoid the contamination from background fluctuations and therefore consider only the contamination from lower redshift interlopers. Since in the real case, relatively bright neighbours are easily recognized as contaminants in the two-dimensional spectra (that we have carefully checked), we have also excluded from the image, the sources brighter than a given U-mag (U_{max}) and replaced them with null pixels. In general it is also true that when planning spectroscopic observations of high redshift (faint) sources, slits would not be placed on targets with very bright neighbours.

We then placed *N* random rectangular slits (with *N* being the size of the sub-sample that we want to simulate) on the final image, and measured the mean flux coming from these slits. The rectangular apertures that we put on the image have dimensions 17×33 pix^2 which correspond to 1×2 arcsec^2. 1 arcsec mimics the slit width used to acquire the VUDS spectra and 2 arcsec takes into account the fact that we cannot discriminate foreground sources closer than the seeing ($2 \times$ seeing). We have assumed a mean seeing of 1 arcsec, which was the nominal condition of the VUDS survey, although observations were carried out in service mode and we do not have a precise statistics of the weather conditions during the acquisition of the spectra.

Finally, we estimated the expected average integrated contribution of the foreground sources for each sub-sample, performing the above procedure 10 000 times and evaluating the median, 1σ and 2σ confidence levels of the mean f_ν^{sim} distribution. The median and dispersion of the final distribution depend on the number *N* of slits and on the choice of U_{max}.

A.2. Dependence on U_{max}

The most delicate parameter to choose is U_{max}. Indeed, modifying this value, significantly changes the results of our simulations.

In Fig. A.1 we show the simulated observed-frame f_ν^{sim} coming from foreground sources for the total sample of 201 galaxies using two different cuts, $U_{max} = 25$ and $U_{max} = 26$ following Vanzella et al. (2010). From the stack we cannot directly measure the average flux in the LyC range, because we normalized each spectrum at its value in the UV range during the stacking procedure. For this reason, we get $f_\nu^{obs}(895)$ multiplying the flux density ratio with the mean UV flux of the sample. To evaluate the latter, we averaged, for each galaxy in the sample, the signal in the UV range of the spectrum (1420−1520 Å) and then evaluated the mean and the 1σ error using the bootstrapping technique.
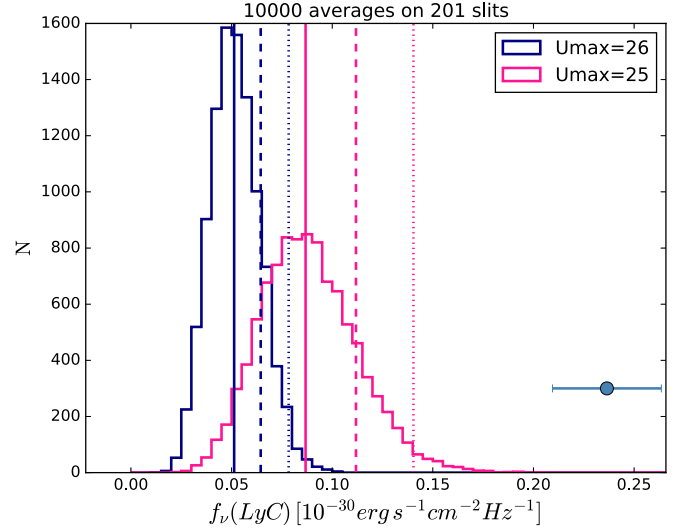

Fig. A.1. Distributions of the simulated observed-frame f_ν^{sim} coming from foreground sources for the total sample of 201 galaxies imposing $U_{max} = 26$ (blue histogram) and $U_{max} = 25$ (magenta histogram). The continuum vertical lines are the medians of the two distributions and the dashed and dotted lines correspond to the 1σ and 2σ confidence intervals. The symbol in the figure is the average observed-frame $f_\nu^{obs}(895)$ obtained for the total sample.

This simple exercise outlines the effect of the choice of U_{max} on the results of the simulations. Indeed, if we assume $U_{max} = 25$, the 2σ value of the distribution of the simulated flux is about twice the value of the distribution obtained assuming $U_{max} = 26$, implying a much higher contamination. In both cases the observed 895 Å flux is much higher than the simulated values. This could imply that the total sample contains not only contaminated sources but also some real LyC emitting galaxies, although this possibility is quite unlikely since we expect that the LyC emission is almost wiped out in the stack. Most probably our simulations underestimate the real contamination, at least for this particular choice of parameters. One further possibility is that our choice of slits width is too optimistic and the spectra contain scattered light from objects that are located just outside the slit. We explore this scenario in Appendix A.4.

A.3. Dependence on the sample size

The shape of the distribution of the simulated flux, as well as the median and the 1σ and 2σ values, depend also on the number of slits putted on the image in each simulation. We show in Fig. A.2 the two distributions obtained for the total sample of 201 galaxies (blue histogram) and for the sub-sample with Ly$\alpha_{ext} \leq 5.7$ kpc (magenta histogram), which is formed by 20 galaxies, using the same U_{max} (we chose the value of $U_{max} = 25$). It is now possible to see the effect of using a different number of slits on the distribution of the simulated contamination flux. Indeed, when a lower number of slits is used, the distribution is much more extended (magenta histogram). Furthermore, while the median value of the distribution (indicated in the figure with a vertical line) does not change much, only few percents as expected, the values of the 1σ (dashed lines) and 2σ (dotted lines) confidence intervals significantly change.

To validate the results obtained in Sect. 4, we need to verify the assumption that we made in Sect. 2, that is that all our sub-samples have approximately the same level of contamination. With this purpose, we performed a set of simulations for each
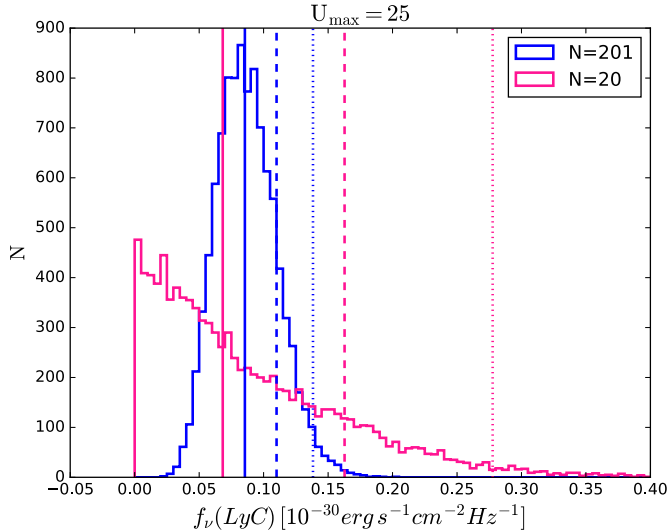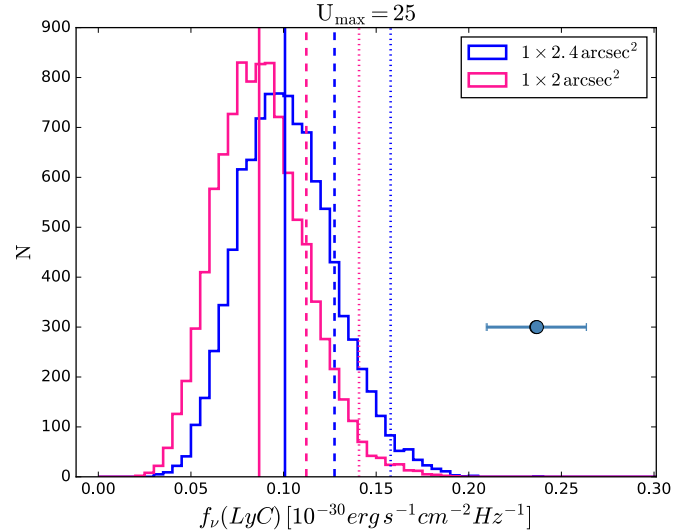
Fig. A.2. Comparison of the two distributions of the simulated observed-frame f_ν^{sim} coming from foreground sources for the total sample (blue histogram, 201 slits per simulation) and for the sub-sample with $Ly\alpha_{ext} \leq 5.7$ kpc (magenta histogram, 20 slits per simulation) using $U_{max} = 25$. The continuum vertical lines are the medians of the two distributions, the dashed lines correspond to the 1σ confidence intervals while the dotted lines to the 2σ confidence intervals.
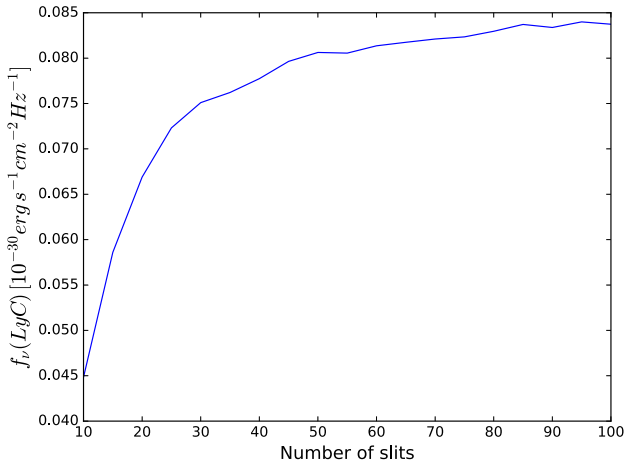

Fig. A.3. Simulated contamination flux as a function of the number of slits used in the simulations. The flux in the figure is the median value of the distribution and it is expressed in units of 10^{-30} erg s^{-1}cm^{-2} Hz^{-1}.

value of N in a range between 10 and 100 with a step of 5, fixing $U_{max} = 25$. We show in Fig. A.3 the median values of each distribution of simulated fluxes as a function of the number of slits used in the simulations. We obtain that the median contamination does not change much varying the number of slits used. Actually, Fig. A.3 shows that if we put in the simulations a smaller number of slits, the contamination is lower. Our most significant sub-samples are formed by very few galaxies compared to their complementaries (see Table 1),

Fig. A.4. Comparison of the two distributions of the simulated observed-frame f_ν^{sim} coming from foreground sources for the total sample (201 slits per simulation) using two different slits dimensions (1×2 arcsec^2, magenta histogram, and 1×2.4 arcsec^2, blue histogram) and fixing $U_{max} = 25$. The continuum vertical lines are the medians of the two distributions, the dashed lines correspond to the 1σ confidence intervals while the dotted lines to the 2σ confidence intervals. The symbol in the figure is the average observed-frame $f_\nu^{obs}(895)$ obtained for the total sample.

therefore the excess in the flux density ratio observed in these sub-samples should not be interpreted as contaminated flux but as true emission.

A.4. Dependence on the slits dimension

As described in Appendix A.1, so far we have used slits of 1×2 arcsec^2, assuming an average seeing of 1 arcsec, which was the nominal condition of the VUDS survey. However, the observations were carried out by ESO in service mode so it is possible that the seeing was not always strictly within this limit. In addition light could scatter into the slits from sources that are placed immediately outside the slit. For this reason we performed again our simulations using a slightly larger slit to test the dependence of the results. As an example, we use slits of 1×2.4 arcsec^2 to simulate an average seeing of 1.2 arcsec (the same effect would be obtained using larger slits to take into account scatter light). In Fig. A.4 we show the comparison of the two distributions of the simulated observed-frame f_ν^{sim} for the total sample (201 slits per simulation) using the different slit dimensions (1×2 arcsec^2, magenta histogram and 1×2.4 arcsec^2, blue histogram) and fixing $U_{max} = 25$. As expected, the distribution obtained using wider slits is shifted to higher values of contamination, although the effect is lower than the one produced by the change in U_{max} described in Appendix A.2. Also in this case the observed 895 Å flux value from the entire sample of 201 galaxies (which is the symbol in Fig. A.4) is larger than the simulated values for both slit sizes.