Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **465,** 3637–3655 (2017)
Advance Access publication 2016 November 17

doi:10.1093/mnras/stw2973

The production and escape of Lyman-Continuum radiation from star-forming galaxies at $z \sim 2$ and their redshift evolution

Jorryt Matthee,[1]★ David Sobral,[1,2] Philip Best,[3] Ali Ahmad Khostovan,[4] Iván Oteo,[3,5] Rychard Bouwens[1] and Huub Röttgering[1]

[1]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[2]*Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK*
[3]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[4]*University of California, Riverside, 900 University Ave, Riverside, CA 92521, USA*
[5]*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany*

Accepted 2016 November 15. Received 2016 November 15; in original form 2016 May 25

ABSTRACT

We study the production rate of ionizing photons of a sample of 588 Hα emitters (HAEs) and 160 Lyman-α emitters (LAEs) at $z = 2.2$ in the COSMOS field in order to assess the implied emissivity from galaxies, based on their ultraviolet (UV) luminosity. By exploring the rest-frame Lyman Continuum (LyC) with *GALEX*/*NUV* data, we find $f_{\rm esc} < 2.8\,(6.4)$ per cent through median (mean) stacking. By combining the Hα luminosity density with intergalactic medium emissivity measurements from absorption studies, we find a globally averaged $\langle f_{\rm esc}\rangle$ of $5.9^{+14.5}_{-4.2}$ per cent at $z = 2.2$ if we assume HAEs are the only source of ionizing photons. We find similarly low values of the global $\langle f_{\rm esc}\rangle$ at $z \approx 3$–5, also ruling out a high $\langle f_{\rm esc}\rangle$ at $z < 5$. These low escape fractions allow us to measure $\xi_{\rm ion}$, the number of produced ionizing photons per unit UV luminosity, and investigate how this depends on galaxy properties. We find a typical $\xi_{\rm ion} \approx 10^{24.77\pm0.04}$ Hz erg^{-1} for HAEs and $\xi_{\rm ion} \approx 10^{25.14\pm0.09}$ Hz erg^{-1} for LAEs. LAEs and low-mass HAEs at $z = 2.2$ show similar values of $\xi_{\rm ion}$ as typically assumed in the reionization era, while the typical HAE is three times less ionizing. Due to an increasing $\xi_{\rm ion}$ with increasing EW(Hα), $\xi_{\rm ion}$ likely increases with redshift. This evolution alone is fully in line with the observed evolution of $\xi_{\rm ion}$ between $z \approx 2$ and 5, indicating a typical value of $\xi_{\rm ion} \approx 10^{25.4}$ Hz erg^{-1} in the reionization era.

Key words: galaxies: evolution – galaxies: high-redshift – cosmology: observations – dark ages, reionization, first stars.

1 INTRODUCTION

One of the most important questions in galaxy formation is whether galaxies alone have been able to provide the ionizing photons which reionized the Universe. Optical depth measurements from the *Planck* satellite place the mean reionization redshift between $z \approx 7.8$ and 8.8 (Planck Collaboration XLVII et al. 2016). The end point of reionization has been marked by the Gun–Peterson trough in high-redshift quasars at $z \approx 5$–6, with a typical neutral fraction of $\sim 10^{-4}$ (e.g. Fan et al. 2006; McGreer, Mesinger & D'Odorico 2015). Moreover, recent observations indicate that there are large opacity fluctuations among various sightlines, indicating an inhomogeneous nature of reionization (Becker et al. 2015).

Assessing whether galaxies have been the main provider of ionizing photons at $z \gtrsim 5$ (alternatively to active galactic nuclei, AGNs; e.g. Madau & Haardt 2015; Giallongo et al. 2015; Weigel et al. 2015) crucially depends on (i) precise measurements of the number of galaxies at early cosmic times, (ii) the clumping factor of the intergalactic medium (IGM, e.g. Pawlik, Schaye & Dalla Vecchia 2015), (iii) the amount of ionizing photons that is produced (Lyman-Continuum photons, LyC, $\lambda < 912$ Å) and (iv) the fraction of ionizing photons that escapes into the IGM. All these numbers are currently uncertain, with the relative uncertainty greatly rising from (i) to (iv).

Many studies so far have focused on counting the number of galaxies as a function of their ultraviolet (UV) luminosity (luminosity functions) at $z > 7$ (e.g. McLure et al. 2013; Bowler et al. 2014; Atek et al. 2015; Bouwens et al. 2015a; Finkelstein et al. 2015; Ishigaki et al. 2015; McLeod et al. 2015; Castellano et al. 2016; Livermore, Finkelstein & Lotz 2016). These studies typically infer

★ E-mail: matthee@strw.leidenuniv.nl

luminosity functions with steep faint-end slopes, and a steepening of the faint-end slope with increasing redshift (see for example, the recent review from Finkelstein 2015), leading to a high number of faint galaxies. Assuming 'standard' values for the other parameters such as the escape fraction, simplistic models indicate that galaxies may indeed have provided the ionizing photons to reionize the Universe (e.g. Madau, Haardt & Rees 1999; Robertson et al. 2015), and that the ionizing background at $z \sim 5$ is consistent with the derived emissivity from galaxies (Choudhury et al. 2015; Bouwens et al. 2015b). However, without validation of input assumptions regarding the production and escape of ionizing photons (for example, these simplistic models assume that the escape fraction does not depend on UV luminosity), the usability of these models remains to be evaluated.

The most commonly adopted escape fraction of ionizing photons, $f_{\rm esc}$, is 10–20 per cent, independent of mass or luminosity (e.g. Mitra, Choudhury & Ferrara 2015; Robertson et al. 2015). However, hydrodynamical simulations indicate that $f_{\rm esc}$ is likely very anisotropic and time dependent (Cen & Kimm 2015; Ma et al. 2015). An escape fraction which depends on galaxy properties (for example, a higher $f_{\rm esc}$ for lower mass galaxies, e.g. Paardekooper, Khochfar & Dalla Vecchia 2015) would influence the way reionization happened (e.g. Sharma et al. 2016). Most importantly, it is impossible to measure $f_{\rm esc}$ directly at high redshift ($z > 6$) because of the high opacity of the IGM for ionizing photons (e.g. Inoue et al. 2014). Furthermore, to estimate $f_{\rm esc}$ it is required that the intrinsic amount of ionizing photons is measured accurately, which requires accurate understanding of the stellar populations, star formation rate (SFR) and dust attenuation (i.e. De Barros et al. 2016).

Nevertheless, several attempts have been made to measure $f_{\rm esc}$, both in the local Universe (e.g. Leitherer et al. 1995; Deharveng et al. 2001; Leitet et al. 2013; Alexandroff et al. 2015) and at intermediate redshift, $z \sim 3$, where it is possible to observe redshifted LyC radiation with optical CCDs (e.g. Inoue, Iwata & Deharveng 2006; Boutsia et al. 2011; Vanzella et al. 2012; Bergvall et al. 2013; Mostardi et al. 2015). However, the number of reliable direct detections is limited to a handful, both in the local Universe and at intermediate redshift (e.g. Borthakur et al. 2014; Izotov et al. 2016a,b; De Barros et al. 2016; Leitherer et al. 2016), and strong limits of $f_{\rm esc} \lesssim 5$–10 per cent exist for the majority (e.g. Grazian et al. 2016; Guaita et al. 2016; Rutkowski et al. 2016). An important reason is that contamination from sources in the foreground may mimic escaping LyC, and high-resolution UV imaging is thus required (e.g. Mostardi et al. 2015; Siana et al. 2015). Even for sources with established LyC leakage, estimating $f_{\rm esc}$ reliably depends on the ability to accurately estimate the intrinsically produced amount of LyC photons and precisely model the transmission of the IGM (e.g. Vanzella et al. 2016).

The amount of ionizing photons that are produced per unit UV (rest frame ≈1500 Å) luminosity ($\xi_{\rm ion}$) is generally calculated using Spectral Energy Distribution (SED) modelling (e.g. Madau et al. 1999; Bouwens et al. 2012; Kuhlen & Faucher-Giguère 2012) or (in a related method) estimated from the observed values of the UV slopes of high-redshift galaxies (e.g. Robertson et al. 2013; Duncan & Conselice 2015). Most of these studies find values around $\xi_{\rm ion} \approx 10^{25.2-25.3}$ Hz erg^{-1} at $z \sim 8$. More recently, Bouwens et al. (2016) estimated the number of ionizing photons in a sample of Lyman-break galaxies (LBGs) at $z \sim 4$ to be $\xi_{\rm ion} \approx 10^{25.3}$ Hz erg^{-1} by estimating Hα luminosities with *Spitzer*/IRAC photometry.

Progress in the understanding of $f_{\rm esc}$ and $\xi_{\rm ion}$ can be made by expanding the searched parameter space to lower redshifts, where rest-frame optical emission lines (e.g. Hα) can provide valuable information on the production rate of LyC photons and where it is possible to obtain a complete selection of star-forming galaxies.

In this paper, we use a large sample of Hα emitters (HAEs) and Lyα emitters (LAEs) at $z = 2.2$ to constrain $f_{\rm esc}$ and measure $\xi_{\rm ion}$ and how this may depend on galaxy properties. Our measurements of $\xi_{\rm ion}$ rely on the assumption that $f_{\rm esc}$ is negligible (<10 per cent), which we validate by constraining $f_{\rm esc}$ with archival *GALEX NUV* imaging and by comparing the estimated emissivity of HAEs with IGM emissivity measurements from quasar absorption lines (e.g. Becker & Bolton 2013). Combined with rest-frame UV photometry, accurate measurements of $\xi_{\rm ion}$ are possible on a source by source basis for HAEs, allowing us to explore correlations with galaxy properties. Since only a handful of LAEs are detected in Hα (see Matthee et al. 2016), we measure the median $\xi_{\rm ion}$ from stacks of LAEs from Sobral et al. (2016a).

We describe the galaxy sample and definitions of galaxy properties in Section 2. Section 3 presents the *GALEX* imaging. We present upper limits on $f_{\rm esc}$ in Section 4. We indirectly estimate $f_{\rm esc}$ from the Hα luminosity function and the IGM emissivity in Section 5 and measure the ionizing properties of galaxies and its redshift evolution in Section 6. Section 7 discusses the implications for reionization. Finally, our results are summarized in Section 8. We adopt a ΛCDM cosmology with $H_0 = 70$ km s^{-1}Mpc^{-1}, $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$. Magnitudes are in the AB system. At $z = 2.2$, 1 arcsec corresponds to a physical scale of 8.2 kpc.

2 GALAXY SAMPLE

We use a sample of Hα selected star-forming galaxies from the High-z Emission Line Survey (HiZELS; Geach et al. 2008; Sobral et al. 2009) at $z = 2.2$ in the COSMOS field. These galaxies were selected using narrow-band (NB) imaging in the *K* band with the United Kingdom InfraRed Telescope. HAEs were identified among the line emitters using *BzK* and *BRU* colours and photometric redshifts, as described in Sobral et al. (2013), and thus have a photometric redshift of $z = 2.22 \pm 0.02$ where the error is due to the width of the NB filter. In total, there are 588 HAEs at $z = 2.2$ in COSMOS.[1]

HAEs are selected to have $\mathrm{EW}_{0,\mathrm{H}\alpha+[\mathrm{NII}]} > 25$ Å. Since the COSMOS field has been covered by multiple NB filters, a fraction of $z = 2.2$ sources are detected with multiple major emission lines in addition to Hα: [OIII], [OII] (e.g. Sobral et al. 2012; Nakajima et al. 2012; Sobral et al. 2013) or Lyα (e.g. Oteo et al. 2015; Matthee et al. 2016). Multi-wavelength photometry from the observed UV to mid-infrared (IR) is widely available in COSMOS. In this paper, we make explicit use of *V* and *R* band in order to measure the UV luminosity and UV slope β (see Section 2.1.3), but all bands have been used for photometric redshifts (see Sobral et al. 2013, and e.g. Ilbert et al. 2009) and SED fitting (Sobral et al. 2014; Oteo et al. 2015; Khostovan et al. 2016).

We also include 160 LAEs at $z = 2.2$ from the CAlibrating LYMan-α with Hα survey (CALYMHA; Matthee et al. 2016; Sobral et al. 2016a). For completeness at bright luminosities, LAEs were selected with $\mathrm{EW}_{0,\mathrm{Ly}\alpha} > 5$ Å, while LAEs are typically selected with a higher EW_0 cut of 25 Å (see e.g. Matthee et al. 2015 and references therein). Only 15 per cent of our LAEs have $\mathrm{EW}_{0,\mathrm{Ly}\alpha} <$ 25 Å and these are typically AGN, see Sobral et al. (2016a), but they represent some of the brightest. We note that 40 per cent of LAEs

[1] The sample of HAEs from Sobral et al. (2013) is publicly available through e.g. VizieR, http://vizier.cfa.harvard.edu.

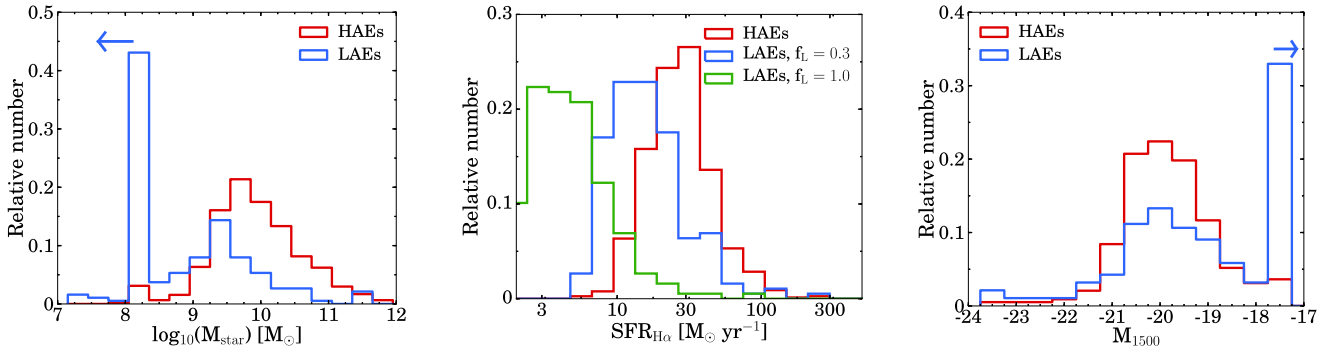

Figure 1. Histogram of the properties of HAEs and LAEs. Stellar mass is obtained through SED fitting (see Section 2.1.1). For HAEs, SFR(Hα) is obtained from dust-corrected Hα (see Section 2.1.2). LAEs that are undetected in broad-bands (and thus without SED fits) are assigned $M_{\rm star} = 10^8$ M$_\odot$ and $M_{1500} = -17$, corresponding to a V-band magnitude of 27 and we assumed those galaxies have no dust in computing SFR(Hα). For LAEs, we use the observed Lyα luminosity and convert this to Hα for two different Lyα escape fractions (f_L, the typical escape fraction for LAEs (30 per cent) and the maximum of 100 per cent, see Sobral et al. 2016a). M_{1500} is obtained by converting the observed V magnitude to absolute magnitude. In general, LAEs trace a galaxy population with lower stellar masses and SFR and fainter UV magnitudes.

are too faint to be detected in broad-bands, and we thus have only upper limits on their stellar mass and UV magnitude (see Fig. 1). By design, CALYMHA observes both Lyα and Hα for Hα selected galaxies. As presented in Matthee et al. (2016), 17 HAEs are also detected in Lyα with the current depth of Lyα NB imaging. These are considered as HAEs in the remainder of the paper.

We show the general properties of our sample of galaxies in Fig. 1. It can be seen that compared to HAEs, LAEs are typically somewhat fainter in the UV, have a lower mass and lower SFR, although they are also some of the brightest UV objects.

Our sample of HAEs and LAEs was chosen for the following reasons: (i) all are at the same redshift slice where the LyC can be observed with the *GALEX NUV* filter and Hα with the NB$_K$ filter, (ii) the sample spans a large range in mass, SFR and environments (Fig. 1 and Geach et al. 2012; Sobral et al. 2014) and (iii) as discussed in Oteo et al. (2015), Hα selected galaxies span the entire range of star-forming galaxies, from dust-free to relatively dust-rich (unlike e.g. LBGs).

2.1 Definition of galaxy properties

We define the galaxy properties that are used in the analysis in this subsection. These properties are either obtained from: (1) SED fitting of the multi-wavelength photometry, (2) observed Hα flux, or (3) observed rest-frame UV photometry.

2.1.1 SED fitting

For HAEs, stellar masses ($M_{\rm star}$) and stellar dust attenuations ($E(B-V)$) are taken from Sobral et al. (2014). In this study, synthetic galaxy SEDs are simulated with Bruzual & Charlot (2003) stellar templates with metallicities ranging from $Z = 0.0001$ to 0.05, following a Chabrier (2003) initial mass function (IMF) and with exponentially declining star formation histories (with e-folding times ranging from 0.1 to 10 Gyr). The dust attenuation is described by a Calzetti et al. (2000) law. The observed UV to IR photometry is then fitted to these synthetic SEDs. The values of $M_{\rm star}$ and $E(B-V)$ that we use are the median values of all synthetic models which have a χ^2 within 1σ of the best-fitting model. The 1σ uncertainties are typically 0.1–0.2 dex for $M_{\rm star}$ and 0.05–0.1 dex for $E(B-V)$. The smallest errors are found at high masses and high extinctions. The same SED fitting method is applied to the photometry of LAEs.

We note that the SED fitting from Sobral et al. (2014) uses SED models which do not take contribution from nebular emission lines into account. This means that some stellar masses could be overestimated. However, the SED fits have been performed on over >20 different filters, such that even if a few filters are contaminated by emission lines, the χ^2 values are not strongly affected. Importantly, the *Spitzer*/IRAC bands (included in SED fitting and most important for measuring stellar mass at $z = 2.2$) are unaffected by strong nebular emission lines at $z = 2.2$.

We still investigate the importance of emission lines further by comparing the SED results with those from Oteo et al. (2015), who performed SED fits for a subsample (≈60 per cent) of the HAEs and LAEs, including emission lines. We find that the stellar masses and dust attenuations correlate very well, although stellar masses from Oteo et al. (2015) are on average lower by 0.15 dex. We look at the galaxies with the strongest lines (highest observed EWs) and find that the difference in the stellar mass is actually smaller than for galaxies with low Hα EW. This indicates that the different mass estimates are not due to the inclusion of emission lines, but rather due to the details of the SED fitting implementation, such as the age-grid ages and allowed range of metallicities. We therefore use the stellar masses from Sobral et al. (2014). Our sample spans galaxies with masses $M_{\rm star} \approx 10^{7.5-12}$ M$_\odot$, see Fig. 1.

2.1.2 Intrinsic Hα luminosity

The intrinsic Hα luminosity is used to compute instantaneous SFRs and the number of produced ionizing photons. To measure the intrinsic Hα luminosity, we first correct the observed line flux in the NB$_K$ filter for the contribution of the adjacent [NII] emission-line doublet. We also correct the observed line flux for attenuation due to dust.

We correct for the contribution from [NII] using the relation between [NII]/Hα and EW$_{0,[\rm NII]+\rm H\alpha}$ from Sobral et al. (2012). This relation is confirmed to hold up to at least $z \sim 1$ (Sobral et al. 2015) and the median ratio of [NII]/(Hα+ [NII]) = 0.19 ± 0.06 is consistent with spectroscopic follow-up at $z \approx 2$ (e.g. Swinbank et al. 2012; Sanders et al. 2015), such that we do not expect that metallicity evolution between $z = 1$ and 2 has a strong effect on the applied correction. For 1 out of the 588 HAEs, we do not detect the continuum in the K band, such that we use the 1σ detection limit in K to estimate the EW and the contribution from [NII]. We apply the same

correction to HAEs that are detected as X-ray AGN (see Matthee et al. 2016 for details on the AGN identification).

If we alternatively use the relation between stellar mass and [NII]/Hα from Erb et al. (2006) at $z \sim 2$, we find [NII]/(Hα+ [NII]) = 0.10 ± 0.03. This different [NII] estimate is likely caused by the lower metallicity of the Erb et al. (2006) sample, which may be a selection effect (UV selected galaxies typically have less dust than Hα selected galaxies, and are thus also expected to be more metal poor, i.e. Oteo et al. 2015). The difference in [NII] contributions estimated either from the EW or mass is smaller for higher mass HAEs, which have a higher metallicity. Due to the uncertainties in the [NII] correction we add 50 per cent of the correction to the uncertainty in the Hα luminosity in quadrature.

Attenuation due to dust is estimated with a Calzetti et al. (2000) attenuation curve and by assuming that the nebular attenuation equals the stellar attenuation, $E(B-V)_{\rm gas} = E(B-V)_{\rm stars}$. This is in agreement with the average results from the Hα sample from MOSDEF (Shivaei et al. 2015), although we note that there are indications that the nebular attenuation is stronger for galaxies with higher SFRs and masses (e.g. Reddy et al. 2015; Puglisi et al. 2016) and other studies indicate slightly higher nebular attenuations (e.g. Förster Schreiber et al. 2009; Wuyts et al. 2011; Kashino et al. 2013). We note that we vary the method to correct for dust in the relevant sections (e.g. Section 6.3) in two ways: either based on the UV slope (Meurer, Heckman & Calzetti 1999), or from the local relation between dust attenuation and stellar mass (Garn & Best 2010).

SFRs are obtained from dust-corrected L(Hα) and using a Chabrier (2003) IMF: SFR $= 4.4 \times 10^{-42}$ L(Hα) (e.g. Kennicutt 1998), where the SFR is in $M_\odot$ yr^{-1} and L(Hα) in erg s^{-1}. The SFRs of galaxies in our sample range from 3 to 300 $M_\odot$ yr^{-1}, with a typical SFR of ≈30 $M_\odot$ yr^{-1}, see Fig. 1.

2.1.3 *Rest-frame UV photometry and UV slopes*

For our galaxy sample at $z = 2.2$, the rest-frame UV (∼1500 Å) is traced by the V band, which is not contaminated by (possibly) strong Lyα emission. Our full sample of galaxies is imaged in the optical V and R filters with Subaru Suprime-Cam as part of the COSMOS survey (Taniguchi et al. 2007). The 5σ depths of V and R are 26.2–26.4 AB magnitude (see e.g. Muzzin et al. 2013) and have a full width at half-maximum (FWHM) of ∼0.″8. The typical HAE in our sample has a V-band magnitude of ≈25 and is thus significantly detected. 5–7 per cent of the HAEs in our sample are not detected in either the V or R band.

We correct the UV luminosities from the V band for dust with the Calzetti et al. (2000) attenuation curve and the fitted $E(B-V)$ values. The absolute magnitude, M_{1500}, is obtained by subtracting a distance modulus of $\mu = 44.97$ (obtained from the luminosity distance and corrected for bandwidth stretching with $2.5\log_{10}(1+z)$, $z = 2.23$) from the observed V-band magnitudes. The UV slope β is measured with observed V and R magnitudes following:

$$\beta = -\frac{V-R}{2.5\log_{10}(\lambda_V/\lambda_R)} - 2 \quad (1)$$

Here, $\lambda_V = 5477.83$ Å, the effective wavelength of the V filter and $\lambda_R = 6288.71$ Å, the effective wavelength of the R filter. With this combination of filters, β is measured around a rest-frame wavelength of ∼1800 Å.

3 *GALEX* UV DATA

For galaxies observed at $z = 2.2$, rest-frame LyC photons can be observed with the *NUV* filter on the *GALEX* space telescope. In COSMOS, there is deep *GALEX* data (3σ AB magnitude limit ∼25.2, see e.g. Martin et al. 2005; Muzzin et al. 2013) available from the public Deep Imaging Survey. We stress that the FWHM of the point spread function (PSF) of the *NUV* imaging is 5.″4 (Martin et al. 2003) and that the pixel scale is 1.″5 pix^{-1}. We have acquired *NUV* images in COSMOS from the Mikulski Archive at the Space Telescope Science Institute (MAST).[2] All HAEs and LAEs in COSMOS are covered by *GALEX* observations, due to the large circular field of view with 1.°25 diameter. Five pointings in the COSMOS field overlap in the centre, which results in a total median exposure time of 91.4 ks and a maximum exposure time of 236.8 ks.

3.1 Removing foreground/neighbouring contamination

The large PSF-FWHM of *GALEX NUV* imaging leads to a major limitation in measuring escaping LyC photons from galaxies at $z = 2.2$. This is because the observed flux in the *NUV* filter could (partly) be coming from a neighbouring foreground source at lower redshift. In order to overcome this limitation, we use available high-resolution deep optical *Hubble Space Telescope (HST)*/ACS *F814W* (rest frame ≈2500 Å, Koekemoer et al. 2007) imaging to identify sources for which the *NUV* flux might be confused due to possible foreground or neighbouring sources and remove these sources from the sample. In addition, we use visual inspections of deep ground-based U-band imaging as a cross-check for the bluest sources which may be missed with the *HST* imaging. These data are available through the COSMOS archive.[3]

Neighbours are identified using the photometric catalogue from Ilbert et al. (2009), which is selected on deep *HST*/ACS *F814W* data. We find that 195 out of the 588 HAEs in COSMOS have no neighbour inside a radius of 2.″7. We refer to this subsample as our CLEAN sample of galaxies in the remainder of the text. The average properties (dust attenuation, UV magnitude mass and SFR) of this sample is similar to the full sample of Star forming galaxys (SFGs).

3.2 Transmission redward of 912 Å

For sources at $z = 2.22$, the *NUV* filter has non-negligible transmission from $\lambda_0 = 912$ to 933 Å of ≈1.5 per cent. This limits the search for escaping LyC radiation. The fraction of the observed flux in the *NUV* filter that originates from $\lambda_0 > 912$ Å depends on the galaxy's SED, the IGM transmission and the filter transmission. In order to estimate this contribution, we first use a set of STARBURST99 (Leitherer et al. 1999) SED models to estimate the shape of the galaxy's SED in the far-UV. We assume a single burst of star formation with a Salpeter IMF with upper mass limit 100 $M_\odot$, Geneva stellar templates without rotation (Mowlavi et al. 2012) and metallicity $Z = 0.02$. Then, we convolve this SED with the filter and IGM transmission curves, to obtain the fraction of the flux in the *NUV* filter that is non-ionizing at $z = 2.2$ (compared to the flux in the *NUV* filter that is ionizing). By using the SED models with Hα EWs within our measured range, we find that 2.6 ± 0.4 per cent of the flux observed in the *NUV* filter is non-ionizing. This means that upper limits from non-detections are slightly overestimated. For

[2] https://mast.stsci.edu/
[3] http://irsa.ipac.caltech.edu/data/COSMOS/

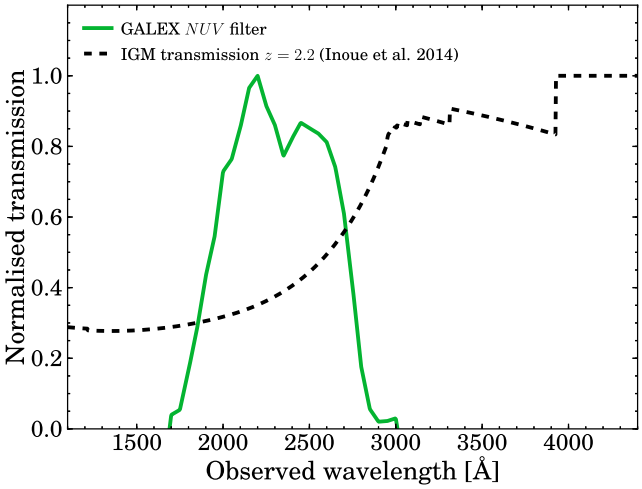

Figure 2. Filter transmission of the *GALEX NUV* filter (green line) and mean IGM transmission versus observed wavelength (dashed black line). We compute the IGM transmission at $z = 2.2$ using the models from Inoue et al. (2014). The bandpass-averaged IGM transmission is 40.4 per cent. As highlighted by a simulation from Vasei et al. (2016), the mean value of $T_{\rm IGM}$ is not the most common value. The distribution is bimodal, with a narrow peak at $T_{\rm IGM} \approx 0.0$ and a broad peak around $T_{\rm IGM} = 0.7$.

individually detected sources, it is theoretically possible that the *NUV* detection is completely due to non-ionizing flux, depending on the SED shape and normalization. This is analysed in detail on a source-by-source basis in Appendix A.

4 THE ESCAPE FRACTION OF IONIZING PHOTONS

4.1 How to measure $f_{\rm esc}$?

Assuming that LyC photons escape through holes in the ISM (and hence that HII regions are ionization bounded from which no ionizing photons escape), the escape fraction, $f_{\rm esc}$, can be measured directly from the ratio of observed to produced LyC luminosity (averaged over the solid angle of the measured aperture).

In this framework, produced LyC photons either escape the ISM, ionize neutral gas (leading to recombination radiation) or are absorbed by dust (e.g. Bergvall et al. 2006). The number of produced ionizing photons per second, $Q_{\rm ion}$, can be estimated from the strength of the (dust corrected) Hα emission line as follows:

$$L_{\rm H\alpha} = Q_{\rm ion}\, c_{\rm H\alpha}\, (1 - f_{\rm esc} - f_{\rm dust}) \quad (2)$$

where $Q_{\rm ion}$ is in s^{-1}, $L_{\rm H\alpha}$ is in erg s^{-1}, $f_{\rm esc}$ is the fraction of produced ionizing photons that escapes the galaxy and $f_{\rm dust}$ is the fraction of produced ionizing photons that is absorbed by dust. For case B recombinations with a temperature of $T = 10\,000$ K, $c_{\rm H\alpha} = 1.36 \times 10^{-12}$ erg (e.g. Kennicutt 1998; Schaerer 2003). Since the dust attenuation curve at wavelengths below 912 Å is highly uncertain, we follow the approach of Rutkowski et al. (2016), who use $f_{\rm dust} = 0.5$, which is based on the mean value derived by Inoue (2002) in local galaxies.

Rest-frame LyC photons are redshifted into the *NUV* filter at $z = 2.2$. However, the IGM between $z = 2.2$ and our telescopes is not transparent to LyC photons (see Fig. 2), such that we need to correct the observed LyC luminosity for IGM absorption. The observed luminosity in the *NUV* filter (L_{NUV}) is then related to the produced number of ionizing photons as:

$$L_{NUV} = Q_{\rm ion}\, \epsilon\, f_{\rm esc}\, T_{\rm IGM,NUV} \quad (3)$$

Here, ϵ is the average energy of an ionizing photon observed in the *NUV* filter (which traces rest-frame wavelengths from 550 to 880 Å, see Fig. 2). Using the STARBURST99 models as described in Section 3.2, we find that ϵ is a strong function of age, but that it is strongly correlated with the EW of the Hα line (which itself is also a strong function of age). For the range of Hα EWs in our sample, $\epsilon = 17.04^{+0.45}_{-0.26}$ eV. We therefore take $\epsilon = 17.0$ eV.

$T_{\rm IGM,NUV}$ is the absorption of LyC photons due to the intervening IGM, convolved with the *NUV* filter. Note that $T_{\rm IGM} = e^{-\tau_{\rm IGM}}$, where $\tau_{\rm IGM}$ is the optical depth to LyC photons in the IGM, see e.g Vanzella et al. (2012). The IGM transmission depends on the wavelength and redshift. According to the model of Inoue et al. (2014), the mean IGM transmission for LyC radiation at $\lambda \sim 750$ Å for a source at $z = 2.2$ is $T_{\rm IGM} \approx 40$ per cent. We convolve the IGM transmission as a function of observed wavelength for a source at $z = 2.2$ with the normalized transmission of the *NUV* filter, see Fig. 2. This results in a bandpass-averaged $T_{\rm IGM,NUV} = 40.4$ per cent.

Combining equations (2) and (3) results in:

$$f_{\rm esc} = \frac{1 - f_{\rm dust}}{(1 + \alpha \frac{L_{\rm H\alpha}}{L_{NUV}})} \quad (4)$$

where we define $\alpha = \epsilon\, c_{\rm H\alpha}^{-1}\, T_{\rm IGM,NUV}$. Combining our assumed values, we estimate $\alpha = 8.09$. We note that ϵ and $c_{\rm H\alpha}$ are relatively insensitive to systematic uncertainties, while $f_{\rm dust}$ and $T_{\rm IGM}$ are highly uncertain for individual sources.

In addition to the absolute escape fraction of ionizing radiation, it is common to define the relative escape fraction of LyC photons to UV (∼1500 Å) photons, since these are most commonly observed in high-redshift galaxies. Following Steidel, Pettini & Adelberger (2001), the relative escape fraction, $f_{\rm esc}^{\rm rel}$, is defined as:

$$f_{\rm esc}^{\rm rel} = f_{\rm esc} e^{\tau_{\rm dust,\,UV}} = \frac{(L_{\rm UV}/L_{NUV})_{\rm int}}{(L_{\rm UV}/L_{NUV})_{\rm obs}}\, T_{\rm IGM,NUV}^{-1} \quad (5)$$

In this equation, $L_{\rm UV}$ is the luminosity in the observed *V* band, $e^{\tau_{\rm dust,\,UV}}$ is the correction for dust (see Section 2.1.3) and we adopt an intrinsic ratio of $(L_{\rm UV}/L_{NUV})_{\rm int} = 5$ (e.g. Siana et al. 2007). The relative escape fraction can be related to the absolute escape fraction when the dust attenuation for $L_{\rm UV}$, $A_{\rm UV}$, is known: $f_{\rm esc} = f_{\rm esc}^{\rm rel} \times 10^{-0.4 A_{\rm UV}}$.

4.2 Individual *NUV* detections

By matching our sample of HAEs and LAEs with the public *GALEX* EM cleaned catalogue (e.g. Zamojski et al. 2007; Conseil et al. 2011), we find that 33 HAEs and 5 LAEs have a detection with $NUV < 26$ within a separation of 1 arcsec. However, most of these matches are identified as spurious, foreground sources or significantly contaminated inside the large PSF-FWHM of *NUV* imaging (see Appendix A). Yet, seven of these matches (of which five are AGN) are in the CLEAN subsample without a clear foreground source and could thus potentially be LyC leakers. Because it is known that foreground contamination has been a major problem in studies of LyC leakage at $z \sim 3$ (e.g. Mostardi et al. 2015; Siana et al. 2015), we can only confirm the reality of these candidate LyC leakers with high-resolution UV imaging with *HST*. We list the individual detections in Appendix A, but caution the reader that any further investigation requires these candidates to be confirmed first.

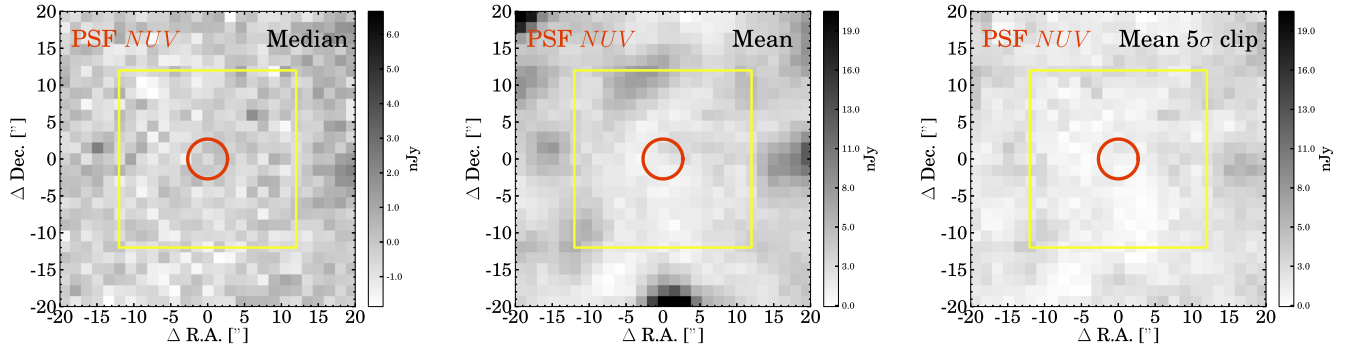

Figure 3. 20 arcsec × 20 arcsec thumbnail images of the *NUV* stack for CLEAN, star-forming HAEs in COSMOS, for three different stacking methods. The red circle shows the PSF-FWHM of *NUV* on the central position. The yellow box is the box which is used to measure the depth of the stack. Note that the range of the colour-bar of the median stack is different than the colour-bar of the mean stacks because the median stack is deeper.

Table 1. Stacked measurements for subsamples of HAEs and LAEs at $z = 2.2$. # indicates the number of objects in each subsample. We further show the general characteristics of the subsample with observed Hα luminosity (corrected for [NII] contribution, see Section 2.1.2), the Hα extinction with the $E(B-V)$ value and a Calzetti law, the median stellar mass and UV slope (β) inferred from $V-R$ colours. The *NUV* column shows the limits on the *NUV* magnitude. L_{1500} is the rest-frame 1500 Å luminosity obtained from the *V* band. The absolute f_{esc} is measured from Hα and the *NUV* as described in Section 4.1. $f_{esc,\,rel}$ is the relative escape fraction of ionizing photons to UV photons and is measured from *NUV* and L_{1500}. Note that with a Calzetti law $A_{UV} = 3.1A_{H\alpha}$. CLEAN subsamples are samples without foreground/neighbouring source within the *NUV* PSF (2.″7).

Subsample	#	$L_{H\alpha,obs}$ (erg s^{-1})	$A_{H\alpha}$ (mag)	β	M_{star} $\log_{10}(M_\odot)$	*NUV* 1σ AB	L_{1500} (erg s^{-1}Hz^{-1})	f_{esc} (per cent)	f_{esc}^{rel} (per cent)
Median stacking									
COSMOS no AGN CLEAN	191	1.60×10^{42}	1.23	−1.97	9.55	29.7	5.78×10^{28}	<2.8	<92.5
Mean stacking									
COSMOS no AGN CLEAN						27.9		<11.7	<465.4
–5σ clip						28.7		<6.4	<231.0

4.3 Stacks of HAEs

The majority of our sources are undetected in the *NUV* imaging. Therefore, in order to constrain f_{esc} for typical star-forming galaxies, we stack *NUV* thumbnails of our full sample of HAEs in COSMOS and also stack various subsets. We create thumbnails of 40 arcsec × 40 arcsec centred on the position of the NB_K (Hα) detection and stack these by either median or mean combining the counts in each pixel. While median stacking results in optimal noise properties and is not dominated by outliers, it assumes that the underlying population is uniform, which is likely not the case. Mean stacking is much more sensitive to outliers (such as, for example, luminous AGN), but would give a more meaningful result as it gives the average f_{esc}, which is the important quantity in assessing the ionizing photon output of the entire galaxy population.

We measure the depth by randomly placing 100 000 empty apertures with a radius of 0.67 ×PSF-FWHM (similar to e.g. Cowie, Barger & Trouille 2009; Rutkowski et al. 2016) in a box of 24 arcsec × 24 arcsec around the centre of the thumbnail (see Fig. 3) and quote the 1σ standard deviation as the depth. Apertures with a detection of *NUV* < 26 AB magnitude are masked (this is particularly important for mean stacking). Counts are converted to AB magnitudes with the photometric zero-point of 20.08 (Cowie et al. 2009). For mean stacking, we experiment with an iterative 5σ clipping method in order to have the background not dominated by a few luminous sources. To do this, we compute the standard deviation of the counts of the stacked sample in each pixel and ignore 5σ outliers in computing the mean value of each pixel. This is iterated five times, although we note that most of the mean values already converge after a single iteration.

By stacking only sources from the CLEAN sample and by removing X-ray AGN, the limiting *NUV* magnitude of the stack of CLEAN HAEs is *NUV* ≈ 29.7 AB (see Table 1), which translates into an upper limit of $f_{esc} < 2.8$ per cent. Mean stacking gives shallower constraints $f_{esc} < 11.7$ per cent because the noise does not decrease as rapidly by stacking more sources, possibly because of a contribution from faint background or companion sources below the detection limit. This is improved somewhat by our iterative 5σ clipping ($f_{esc} < 6.4$ per cent), which effectively masks out the contribution from bright pixels. We show the stacked thumbnails of this sample in Fig. 3.

The median (mean) upper limit on the relative escape fraction, $f_{esc,\,rel}$, is much higher (<92.5(231) per cent). However, if we correct for the dust attenuation with the Calzetti et al. (2000) law, we find $A_{UV} \approx 3.8$ and a dust corrected inferred escape fraction of <2.8(7.0) per cent, in good agreement with our direct estimate from Hα, although we note that the additional uncertainty due to this dust correction is large.

We have experimented by stacking subsets of galaxies in bins of stellar mass, SFR and UV magnitude or LAEs, but all result in a non-detection in the *NUV*, all with weaker upper limits than the stack of CLEAN HAEs.

4.3.1 *Systematic uncertainty due to the dust correction*

In this subsection, we investigate how sensitive our results are to the method used to correct for dust, which is the most important systematic uncertainty. In Table 1, we have used the SED inferred value of $E(B-V)$ to infer $A_{H\alpha}$: $A_{H\alpha} = E(B-V) \times k_{H\alpha}$, where

$k_{H\alpha} = 3.3277$ following Calzetti et al. (2000), which results in $A_{H\alpha} = 1.23$. However, it is also possible to infer $A_{H\alpha}$ from a relation with the UV slope (e.g. Meurer et al. 1999), such that $A_{H\alpha} = 0.641(\beta + 2.23)$, for $\beta > -2.23$ and $A_{H\alpha} = 0$ for $\beta < -2.23$. Finally, we also use the relation between $A_{H\alpha}$ and stellar mass from Garn & Best (2010), which is: $A_{H\alpha} = 0.91 + 0.77X + 0.11X^2 - 0.09X^3$, where $X = \log_{10}(M_{star}/10^{10}\ M_{\odot})$. Note that we assume a Calzetti et al. (2000) dust law in all these prescriptions.

It is immediately clear that there is a large systematic uncertainty in the dust correction, as for our full sample of HAEs we infer $A_{H\alpha} = 0.70$ with the Garn & Best (2010) prescription and $A_{H\alpha} = 0.19$ following Meurer et al. (1999), meaning that the systematic uncertainty due to dust can be as large as a factor of 3. Thus, these different dust corrections result in different upper limits on f_{esc}. For the CLEAN, star-forming HAE sample, the upper limit on f_{esc} from median stacking increases to $f_{esc} < 4.4\,(6.6)$ per cent, using the attenuation based on stellar mass (β). With a simple 1 mag of extinction for Hα, $f_{esc} < 3.4$ per cent and without correcting for dust results in $f_{esc} < 7.7$ per cent.

In addition to the uncertainty in the dust correction of the Hα luminosity, another uncertainty in our method is the f_{dust} parameter introduced in equation (2). The dust attenuation curve at wavelengths below 912 Å is highly uncertain, such that our estimate of f_{dust} is uncertain as well. However, because our limits on f_{esc} from the median stack are low, the results do not change significantly by altering f_{dust}: if $f_{dust} = 0.75(0.25)$, we find $f_{esc} < 1.4(4.1)$ per cent. This means that as long as the limit is low, our result is not very sensitive to the exact value of f_{dust}.

5 CONSTRAINING f_{esc} OF HAES FROM THE IONIZING BACKGROUND

In addition to constraining f_{esc} directly, we can obtain an indirect measurement of f_{esc} by using the ionizing emissivity, measured from quasar absorption studies, as a constraint. The emissivity is defined as the number of escaping ionizing photons per second per comoving volume:

$$\dot{N}_{ion} = \langle f_{esc} \rangle \times \Phi(H\alpha) \times c_{H\alpha}^{-1} \quad (6)$$

Here, $\dot{N}_{ion}$ is in s^{-1} Mpc^{-3}, $\langle f_{esc} \rangle$ is the escape fraction averaged over the entire galaxy population, Φ(Hα) is the Hα luminosity density in erg s^{-1} Mpc^{-3} and $c_{H\alpha}$ is the recombination coefficient as in equation (2).

We first check whether our derived emissivity using our upper limit on f_{esc} for HAEs is consistent with published measurements of the emissivity. The Hα luminosity density is measured in Sobral et al. (2013) as the full integral of the Hα luminosity function, with a global dust correction of $A_{H\alpha} = 1.0$. Using the mean limit on f_{esc} for our CLEAN sample of HAEs (so $f_{esc} \leq 6.4$ per cent), we find that $\dot{N}_{ion} \leq 1.3^{+0.2}_{-0.2} \times 10^{51}$ s^{-1} Mpc^{-3}, where the errors come from the uncertainty in the Hα LF. We note that these numbers are relatively independent on the dust correction method because while a smaller dust attenuation would decrease the Hα luminosity density, it would also raise the upper limit on the escape fraction, thus almost cancelling out. These upper limits on $\dot{N}_{ion}$ are consistent with the measured emissivity at $z = 2.4$ of Becker & Bolton (2013), who measured $\dot{N}_{ion} = 0.90^{+1.60}_{-0.52} \times 10^{51}$ s^{-1} Mpc^{-3} (combined systematic and measurement errors) using the latest measurements of the IGM temperature and opacity to Lyα and LyC photons.

Now, by isolating $\langle f_{esc} \rangle$ in equation (6), we can estimate the globally averaged escape fraction. If we assume that there is no evolution in the emissivity from Becker & Bolton (2013) between $z = 2.2$ and 2.4 and that the Hα luminosity function captures all sources of ionizing photons, we find that $\langle f_{esc} \rangle = 4.4^{+7.1}_{-2.0}$ per cent for $A_{H\alpha} = 1.0$. There are a number of systematic uncertainties that we will address now and add to the error bars of our final estimate. These uncertainties are: (i) integration limit of the Hα LF, (ii) the dust attenuation to L(Hα), (iii) the conversion from L(Hα) to ionizing numbers and (iv) the [NII] correction to the observed Hα luminosity.

Integrating the Hα LF only to SFR $\approx$3 M$_{\odot}$ yr^{-1}, we find $\langle f_{esc} \rangle = 6.7^{+10.8}_{-3.1}$ per cent, such that the systematic uncertainty is of order 50 per cent. If $A_{H\alpha} = 0.7$, which is the median value when we correct for dust using stellar mass, and which may be more representative of fainter HAEs (as faint sources are expected to have less dust), the escape fraction is somewhat higher, with $\langle f_{esc} \rangle = 5.9^{+9.3}_{-2.6}$ per cent. These numbers are summarized in Table 2. The uncertainty in $c_{H\alpha}$ is relatively small, as $c_{H\alpha}$ depends only modestly on the density and the temperature. For example, in the case of a temperature of $T = 30000$ K, $c_{H\alpha}$ decreases only by $\approx$10 per cent (Schaerer 2002). This adds a 10 per cent uncertainty in the escape fraction. As explained in Section 2.1.2, there is an uncertainty in the measured Hα luminosity due to the contribution of the [NII] doublet to the observed NB flux, for which we correct using a relation with observed EW. By comparing this method with the method from Erb et al. (2006), which is based on the observed mass–metallicity relation of a sample of LBGs at $z \sim 2$, we find that the inferred Hα luminosity density would conservatively be 10 per cent higher, such that this correction adds another 10 per cent systematic uncertainty in the escape fraction.

For our final estimate of $\langle f_{esc} \rangle$, we use the dust correction based on stellar mass, fully integrate the Hα luminosity function and add a 10 per cent error in quadrate for the systematic uncertainty in each of the parameters as described above, 50 per cent due to the uncertain integration limits and add a 40 per cent error due to the systematics in the dust attenuation. This results in $\langle f_{esc} \rangle = 5.9^{+14.5}_{-4.2}$ per cent at $z = 2.2$.

An additional contribution to the ionizing emissivity from rarer sources than sources with number densities $<10^{-5}$ Mpc^{-3} such as quasars, would lower the escape fraction for HAEs. While Madau & Haardt (2015) argue that the ionizing budget at $z \approx 2$–3 is dominated by quasars, this measurement may be overestimated by assuming quasars have a 100 per cent escape fraction. Recently, Micheva, Iwata & Inoue (2016) obtained a much lower emissivity (up to a factor of 10) from quasars by directly measuring f_{esc} for a sample of $z \sim 3$ AGN. Using a large sample of quasars at $z = 3.6$–4.0, Cristiani et al. (2016), measure a mean $\langle f_{esc,quasar} \rangle \approx 70$ per cent, which means that quasars do not dominate the ionizing background at $z \approx 4$. When we include a quasar contribution from Madau & Haardt (2015) in the most conservative way (meaning that we assume $f_{esc} = 100$ per cent for quasars), we find that $\langle f_{esc} \rangle = 0.5^{+3.6}_{-0.5}$ per cent. If the escape fraction for quasars is 70 per cent, $\langle f_{esc} \rangle = 1.6^{+5.4}_{-1.3}$ per cent, such that a non-zero contribution from star-forming galaxies is not ruled out.

We note that, these measurements of $\langle f_{esc} \rangle$ contain significantly less (systematic) uncertainties than measurements based on the integral of the UV luminosity function (e.g. Becker & Bolton 2013; Khaire et al. 2016). This is because: (i) UV selected galaxy samples do not necessarily span the entire range of SFGs (e.g. Oteo et al. 2015) and might thus miss dusty star-forming galaxies and (ii) there are additional uncertainties in converting non-ionizing UV luminosity to intrinsic LyC luminosity (in particular the dust corrections in ξ_{ion} and uncertainties in the detailed SED models in $(L_{UV}/L_{NUV})_{int}$). An issue is that Hα is very challenging to observe

Table 2. Measurements of $\langle f_{\rm esc}\rangle$, the escape fraction of ionizing photons averaged over the galaxy population at $z \approx 2$–5. Constraints on the IGM emissivity from absorption studies by Becker & Bolton (2013) have been used to infer the global escape fraction. For $z = 2.2$, we have used the Hα luminosity function from Sobral et al. (2013). We have used the analytical formula from Madau & Haardt (2015) to estimate the contribution from quasars to the ionizing emissivity, which assumes that $f_{\rm esc,quasars} = 100$ per cent. At $z = 3.8$ and 4.9 we have used the SFR function from Smit et al. (2015).

Sample	Method	$\langle f_{\rm esc}\rangle$ (per cent)
This paper		
HAEs $z = 2.2$	Full SFR integration, $A_{\rm H\alpha} = 1.0$	$4.4^{+7.1}_{-2.0}$
HAEs $z = 2.2$	SFR >3 $M_\odot$/yr, $A_{\rm H\alpha} = 1.0$	$6.7^{+10.8}_{-3.1}$
HAEs $z = 2.2$	Full SFR integration, $A_{\rm H\alpha} = 0.7$	$5.9^{+9.3}_{-2.6}$
HAEs $z = 2.2$	Final estimate: full integration, $A_{\rm H\alpha} = 0.7$, conservative systematic errors	$\mathbf{5.9^{+14.5}_{-4.2}}$
HAEs $z = 2.2$	Full SFR integration, $A_{\rm H\alpha} = 1.0$, QSO contribution	$0.5^{+3.6}_{-0.5}$
LBGs $z = 3.8$	Full SFR integration, Hα from *Spitzer*/IRAC	$2.7^{+7.2}_{-2.3}$
LBGs $z = 3.8$	Full SFR integration, Hα from *Spitzer*/IRAC, QSO contribution	$0.0^{+3.0}_{-0.0}$
LBGs $z = 4.9$	Full SFR integration, Hα from *Spitzer*/IRAC	$6.0^{+13.9}_{-5.2}$
LBGs $z = 4.9$	Full SFR integration, Hα from *Spitzer*/IRAC, QSO contribution	$2.1^{+6.2}_{-2.1}$
Literature		
Cristiani et al. (2016) $z = 3.8$	Integrated LBG LF + contribution from QSOs	$5.3^{+2.7}_{-1.2}$

at $z \gtrsim 2.8$ and that a potential spectroscopic follow-up study of UV selected galaxies with the *JWST* might yield biased results.

5.1 Redshift evolution

Using the methodology described in Section 5, we also compute the average $f_{\rm esc}$ at $z = 3.8$ and 4.9 by using the SFR functions of Smit et al. (2015), which are derived from UV luminosity functions, a Meurer et al. (1999) dust correction and a general offset to correct for the difference between SFR(UV) and SFR(Hα), estimated from *Spitzer*/IRAC photometry. This offset is implicitly related to the value of $\xi_{\rm ion}$ from Bouwens et al. (2016), which is estimated from the same measurements. We combine these SFR functions, converted to the Hα luminosity function as in Section 2.1.2, with the IGM emissivity from Becker & Bolton (2013) at $z = 4.0$ and 4.75, respectively. Similarly to the Hα luminosity density, we use the analytical integral of the Schechter function. Similarly as at $z = 2.2$, we conservatively increase the error bars by a factor $\sqrt{2}$ to take systematic uncertainties into account. This results in $\langle f_{\rm esc}\rangle = 2.7^{+7.2}_{-2.3}$ and $6.0^{+13.9}_{-5.2}$ per cent at $z \approx 4$ and 5, respectively, see Table 2. When including a (maximum) quasar contribution from Madau & Haardt (2015) as described above, we find $\langle f_{\rm esc}\rangle = 0.0^{+3.0}_{-0.0}$ per cent at $z \approx 4$ and $\langle f_{\rm esc}\rangle = 2.1^{+6.2}_{-2.1}$ per cent.

As illustrated in Fig. 4, the global escape fraction is low at $z \approx 2$–5. While dust has been corrected for with different methods at $z = 2.2$ and $z \approx 4$–5, we note that the differences between different dust correction methods are not expected to be very large at $z \approx 4$–5. This is because higher redshift galaxies typically have lower mass, which results in a higher agreement between dust correction methods based on either $M_{\rm star}$ or β. One potentially important caveat is that our computation assumes that the Hα and UV luminosity functions include all sources of ionizing photons in addition to quasars. An additional contribution of ionizing photons from galaxies which have potentially been missed by a UV selection (for example, submillimeter galaxies) would decrease the global $f_{\rm esc}$. Such a bias is likely more important at $z \approx 3$–5 than $z \approx 2$ because the $z \approx 2$ sample is selected with Hα which is able to recover submillimeter galaxies. Even under current uncertainties, we rule out a globally averaged $\langle f_{\rm esc}\rangle > 20$ per cent at redshifts lower than $z \approx 5$.

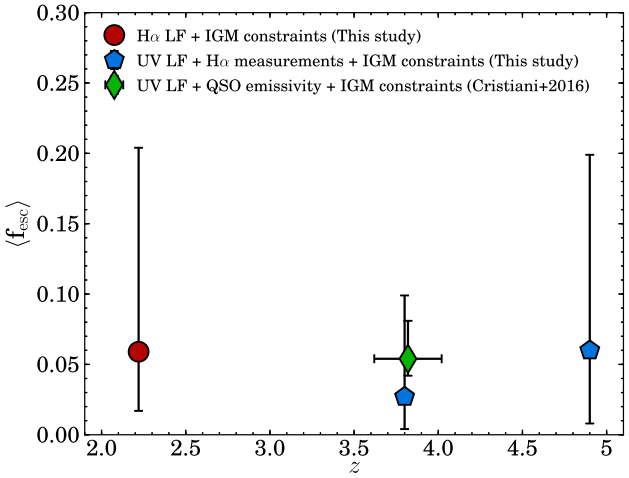

Figure 4. Evolution of the globally averaged $\langle f_{\rm esc}\rangle$, which is obtained by forcing the emissivity of the integrated Hα ($z = 2.2$) and UV ($z \approx 4$–5) LF to be equal to the emissivity measured by IGM absorption models from Becker & Bolton 2013. The $z \approx 4$–5 results are based on a UV luminosity function which is then corrected to an SFR function with Hα measurements from *Spitzer*/IRAC, which implicitly means using a value of $\xi_{\rm ion}$ (SFR functions are presented in Smit et al. 2015, but see also Bouwens et al. 2016). The error bars of red and blue symbols include estimates of the systematic uncertainties. The green diamond shows the estimated value by Cristiani et al. (2016), who combined IGM constraints with a UV LBG and the emissivity of QSOs at $z = 3.6$–4.0.

These indirectly derived escape fractions of ~4 per cent at $z \approx 2$–5 are consistent with recently published upper limits from Sandberg et al. (2015) at $z = 2.2$ and similar to strict upper limits on $f_{\rm esc}$ at $z \sim 1$ measured by Rutkowski et al. (2016), see also Cowie et al. (2009); Bridge et al. (2010). Recently, Cristiani et al. (2016) estimated that galaxies have on average $\langle f_{\rm esc}\rangle = 5.3^{+2.7}_{-1.2}$ per cent at $z \approx 4$ by combining IGM constraints with the UV luminosity function from Bouwens et al. (2011) and by including the contribution from quasars to the total emissivity. This result is still consistent within the error bars with our estimate using the Madau & Haardt (2015) quasar contribution and Smit et al. (2015) SFR function. Part of this is because we use a different conversion from UV luminosity to the

number of produced ionizing photons based on Hα estimates with *Spitzer*/IRAC, and because our computation assumes $f_{\rm esc,\,quasars} = 100$ per cent, while Cristiani et al. (2016) use $f_{\rm esc,\,quasars} \approx 70$ per cent.

Furthermore, our results are also consistent with observations from Chen, Prochaska & Gnedin (2007) who find a mean escape fraction of 2 ± 2 per cent averaged over galaxy viewing angles using spectroscopy of the afterglow of a sample of γ-ray bursts at $z > 2$. Grazian et al. (2016) measures a strict median upper limit of $f_{\rm esc}^{\rm rel} < 2$ per cent at $z = 3.3$, although this limit is for relatively luminous LBGs and not for the entire population of SFGs. This would potentially indicate that the majority of LyC photons escape from galaxies with lower luminosity, or galaxies missed by a Lyman-break selection, i.e. Cooke et al. (2014) or that they come from just a subset of the population, and thus the median $f_{\rm esc}$ can even be close to zero. Khaire et al. (2016) find that $f_{\rm esc}$ must evolve from ≈5 to 20 per cent between $z = 3$ and 5, which is allowed within the errors. However, we note that they assume that the number of produced ionizing photons per unit UV luminosity does not evolve with redshift. In Section 6.5, we find that there is evolution of this number by roughly a factor 1.5, such that the required evolution of $f_{\rm esc}$ would only be a factor ≈3. While our results indicate little to no evolution in the average escape fraction up to $z \approx 5$, this does not rule out an increasing $f_{\rm esc}$ at $z > 5$, where theoretical models expect an evolving $f_{\rm esc}$ (e.g. Kuhlen & Faucher-Giguère 2012; Ferrara & Loeb 2013; Mitra, Ferrara & Choudhury 2013; Khaire et al. 2016; Sharma et al. 2016; Price, Trac & Cen 2016), see also a recent observational claim of evolving $f_{\rm esc}$ with redshift (Smith et al. 2016).

Finally, we stress that a low $\langle f_{\rm esc} \rangle$ is not inconsistent with the recent detection of the high $f_{\rm esc}$ of above 50 per cent from a galaxy at $z \approx 3$ (De Barros et al. 2016; Vanzella et al. 2016), which may simply reflect that there is a broad distribution of escape fractions. For example, if only a small fraction (<5 per cent) of galaxies are LyC leakers with $f_{\rm esc} \approx 75$ per cent, the average $f_{\rm esc}$ over the galaxy population is ≈4 per cent, consistent with the indirect measurement, even if $f_{\rm esc} = 0$ for all other galaxies. Such a scenario would be the case if the escape of LyC photons is a very stochastic process, for example, if it is highly direction or time dependent. This can be tested with deeper LyC limits on individual galaxies over a complete selection of star-forming galaxies.

6 THE IONIZING PROPERTIES OF STAR-FORMING GALAXIES AT $z = 2.2$

6.1 How to measure $\xi_{\rm ion}$?

The number of ionizing photons produced per unit UV luminosity, $\xi_{\rm ion}$, is used to convert the observed UV luminosity of high-redshift galaxies to the number of produced ionizing photons. It can thus be interpreted as the production efficiency of ionizing photons. $\xi_{\rm ion}$ is defined as:

$$\xi_{\rm ion} = Q_{\rm ion}/L_{\rm UV,int} \quad (7)$$

As described in the previous section, $Q_{\rm ion}$ (in s^{-1}) can be measured directly from the dust-corrected Hα luminosity by rewriting equation (2) and assuming $f_{\rm esc} = 0$. $L_{\rm UV,int}$ (in erg s^{-1} Hz^{-1}) is obtained by correcting the observed UV magnitudes for dust attenuation. With a Calzetti et al. (2000) attenuation curve $A_{\rm UV} = 3.1A_{\rm H\alpha}$.

In our definition of $\xi_{\rm ion}$, we assume that the escape fraction of ionizing photons is ≈0. Our direct constraint of $f_{\rm esc} \lesssim 6$ per cent and our indirect global measurement of $f_{\rm esc} \approx 5$ per cent validate this assumption. If the average escape fraction is $f_{\rm esc} = 10$ per cent, $\xi_{\rm ion}$ is higher by a factor 1.11 (so only 0.04 dex), such that $\xi_{\rm ion}$ is relatively insensitive to the escape fraction as long as the escape fraction is low. We also note that the $\xi_{\rm ion}$ measurements at $z \approx 4$–5 from Bouwens et al. (2016) are validated by our finding that the global escape fraction at $z < 5$ is consistent with being very low, <5 per cent.

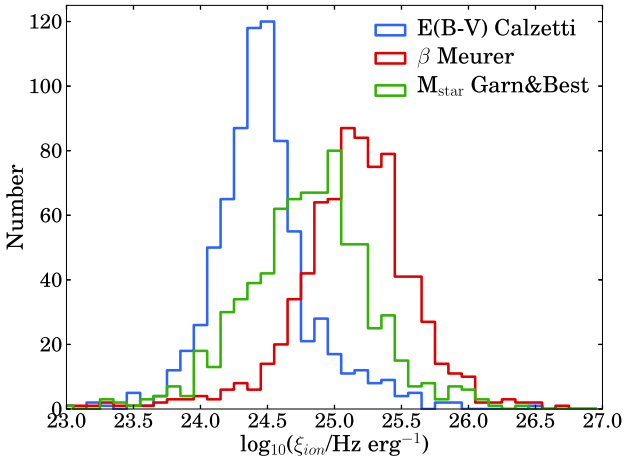

Figure 5. Histogram of the values of $\xi_{\rm ion}$ for HAEs with three different methods to correct for dust attenuation. The blue histogram shows values of $\xi_{\rm ion}$ when dust is corrected with the $E(B-V)$ value from the SED in combination with a Calzetti law (see Section 2.1). The red histogram is corrected for dust with the Meurer et al. (1999) prescription based on the UV slope and the green histogram is corrected for dust with the prescription from Garn & Best (2010) based on a relation between dust attenuation and stellar mass. As can be seen, the measured values of $\xi_{\rm ion}$ differ significantly, with the highest values found when correcting for dust with the UV slope. When the nebular attenuation is higher than the stellar attenuation, $\xi_{\rm ion}$ would shift to higher values.

6.2 $\xi_{\rm ion}$ at $z = 2.2$

We show our measured values of $\xi_{\rm ion}$ for HAEs in Fig. 5 and Table 3, where dust attenuation is corrected with three different methods based either on the $E(B-V)$ value of the SED fit, the UV slope β or the stellar mass. It can be seen that the average value of $\xi_{\rm ion}$ is very sensitive to the dust correction method, ranging from $\xi_{\rm ion} = 10^{24.39 \pm 0.04}$ Hz erg^{-1} for the SED method to $\xi_{\rm ion} = 10^{25.11 \pm 0.04}$ Hz erg^{-1} for the β method. For the dust correction based on stellar mass, the value lies in between, with $\xi_{\rm ion} = 10^{24.85 \pm 0.04}$ Hz erg^{-1}. In the case of a higher nebular attenuation than the stellar attenuation, as for example by a factor ≈2 as in the original Calzetti et al. (2000) prescription, $\xi_{\rm ion}$ increases by 0.4 dex to $\xi_{\rm ion} = 10^{24.79 \pm 0.04}$ Hz erg^{-1} when correcting for dust with the SED fit.

We note that independent (stacking) measurements of the dust attenuation from *Herschel* and Balmer decrements at $z \sim 1$–2 indicate that dust attenuations agree very well with the Garn & Best (2010) prescription (e.g. Sobral et al. 2012; Ibar et al. 2013; Buat et al. 2015; Pannella et al. 2015), thus favouring the intermediate value of $\xi_{\rm ion}$. Without correcting $\xi_{\rm ion}$ for dust, we find $\xi_{\rm ion} = 10^{25.41 \pm 0.05}$ Hz erg^{-1}. With 1 mag of extinction for Hα, as for example used in the conversion of the Hα luminosity density to an SFR density in Sobral et al. (2013), $\xi_{\rm ion} = 10^{24.57 \pm 0.04}$ Hz erg^{-1}.

Since individual Hα measurements for LAEs are uncertain due to the difference in filter transmissions depending on the exact redshift (see Matthee et al. 2016), we only investigate $\xi_{\rm ion}$ for our sample of LAEs in the stacks described in Sobral et al. (2016a). With stacking, we measure the median Hα flux of LAEs convolved

Table 3. Ionizing properties of HAEs and LAEs for various methods to correct for dust attenuations and different subsets. We show the median stellar mass of each subsample. Errors on $\xi_{\rm ion}$ are computed as $\sigma_{\xi_{\rm ion}}/\sqrt{N}$, where $\sigma_{\xi_{\rm ion}}$ is the median measurement error of $\xi_{\rm ion}$ and N the number of sources. For the Bouwens et al. (2016) measurements, we show only dust corrections with a Calzetti et al. (2000) curve. The subsample of 'low-mass' HAEs has $M_{\rm star} = 10^{9.0-9.4}$ M$_\odot$. 'UV faint' HAEs have $M_{1500} > -19$.

Sample	$\langle M_{\rm star}\rangle$ $\log_{10}$ M$_\odot$	$\log_{10}\,\xi_{\rm ion}$ (Hz erg^{-1})	Dust
This paper			
HAEs $z = 2.2$	9.8	24.39 ± 0.04	$E(B-V)$
		25.11 ± 0.04	β
		24.77 ± 0.04	$M_{\rm star}$
		25.41 ± 0.05	No dust
		24.57 ± 0.04	$A_{\rm H\alpha} = 1$
Low mass	9.2	24.49 ± 0.06	$E(B-V)$
		25.22 ± 0.06	β
		24.99 ± 0.06	$M_{\rm star}$
UV faint	10.2	24.93 ± 0.07	$E(B-V)$
		25.39 ± 0.07	β
		25.24 ± 0.07	$M_{\rm star}$
LAEs $z = 2.2$	8.5	24.84 ± 0.09	$E(B-V)$
		25.37 ± 0.09	β
		25.14 ± 0.09	$M_{\rm star}$
		25.39 ± 0.09	No dust
Bouwens et al. (2016)			
LBGs $z = 3.8 - 5.0$	9.2	25.27 ± 0.03	β
LBGs $z = 5.1 - 5.4$	9.2	25.44 ± 0.12	β

through the filter profile and the median UV luminosity by stacking V-band imaging. As seen in Table 3, the median $\xi_{\rm ion}$ is higher than the median $\xi_{\rm ion}$ for HAEs for each dust correction. However, this difference disappears without correcting for dust. Therefore, the higher values of $\xi_{\rm ion}$ for LAEs simply indicate that the median LAE has a bluer UV slope, lower stellar mass and lower $E(B-V)$ than the median HAE. More accurate dust measurements are required to investigate whether $\xi_{\rm ion}$ is really higher for LAEs. We note that ≈40 per cent of the LAEs are undetected in the broad-bands and thus assigned a stellar mass of 10^8 M$_\odot$ and $E(B-V) = 0.1$ when computing the median dust attenuation. Therefore, the $\xi_{\rm ion}$ values for LAEs could be underestimated if the real dust attenuation is even lower.

6.3 Dependence on galaxy properties

In this section, we investigate how $\xi_{\rm ion}$ depends on the galaxy properties that are defined in Section 2.1 and also check whether subsets of galaxies lie in a specific parameter space. As illustrated in Fig. 6 (where we correct for dust with $E(B-V)$), we find that $\xi_{\rm ion}$ does not depend strongly on SFR(Hα) with a Spearman correlation rank ($R_{\rm s}$) of $R_{\rm s} = 0.11$. Such a correlation would naively be expected if the Hα SFRs are not related closely to UV SFRs, since $\xi_{\rm ion} \propto L_{\rm H\alpha}/L_{1500}$ $\propto$ SFR(Hα)/SFR(UV). However, for our sample of galaxies, these SFRs are strongly correlated with only 0.3 dex of scatter, see also Oteo et al. (2015), leading to a relatively constant $\xi_{\rm ion}$ with SFR.

For the same reason, we measure a relatively weak slope of ≈0.25 when we fit a simple linear relation between $\log_{10}(\xi_{\rm ion})$ and M_{1500}, instead of the naively expected value of $\xi_{\rm ion} \propto 0.4 M_{1500}$. At $M_{1500} > -20$, our Hα selection is biased towards high values of Hα relative to the UV, leading to a bias in high values of $\xi_{\rm ion}$. For sources with $M_{1500} < -20$, we measure a slope of ≈0.2. This means that $\xi_{\rm ion}$ does not increase rapidly with decreasing UV luminosity. This is because Hα luminosity and dust attenuation themselves are also related to M_{1500}. Indeed, we find that the Hα luminosity anti-correlates with the UV magnitude and $E(B-V)$ increases for fainter UV magnitudes.

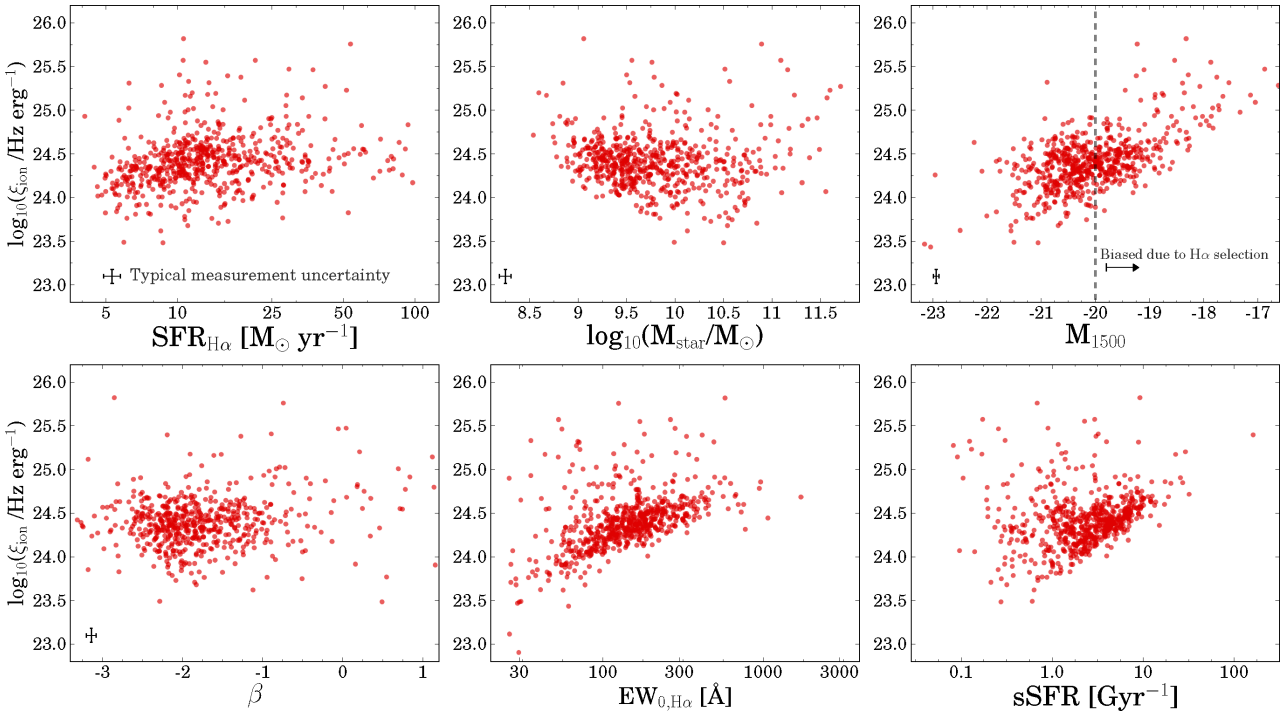

Figure 6. Correlations between $\xi_{\rm ion}$ and galaxy properties for HAEs, when dust is corrected using the SED fitted $E(B-V)$ values. $\xi_{\rm ion}$ does not clearly correlate with SFR(Hα), $M_{\rm star}$ or β. A correlation between $\xi_{\rm ion}$ and M_{1500} is expected of similar strength as seen, based on the definition of $\xi_{\rm ion}$. $\xi_{\rm ion}$ increases strongly with Hα EW and sSFR. High values of $\xi_{\rm ion}$ at low sSFRs are mostly due to the dust correction.

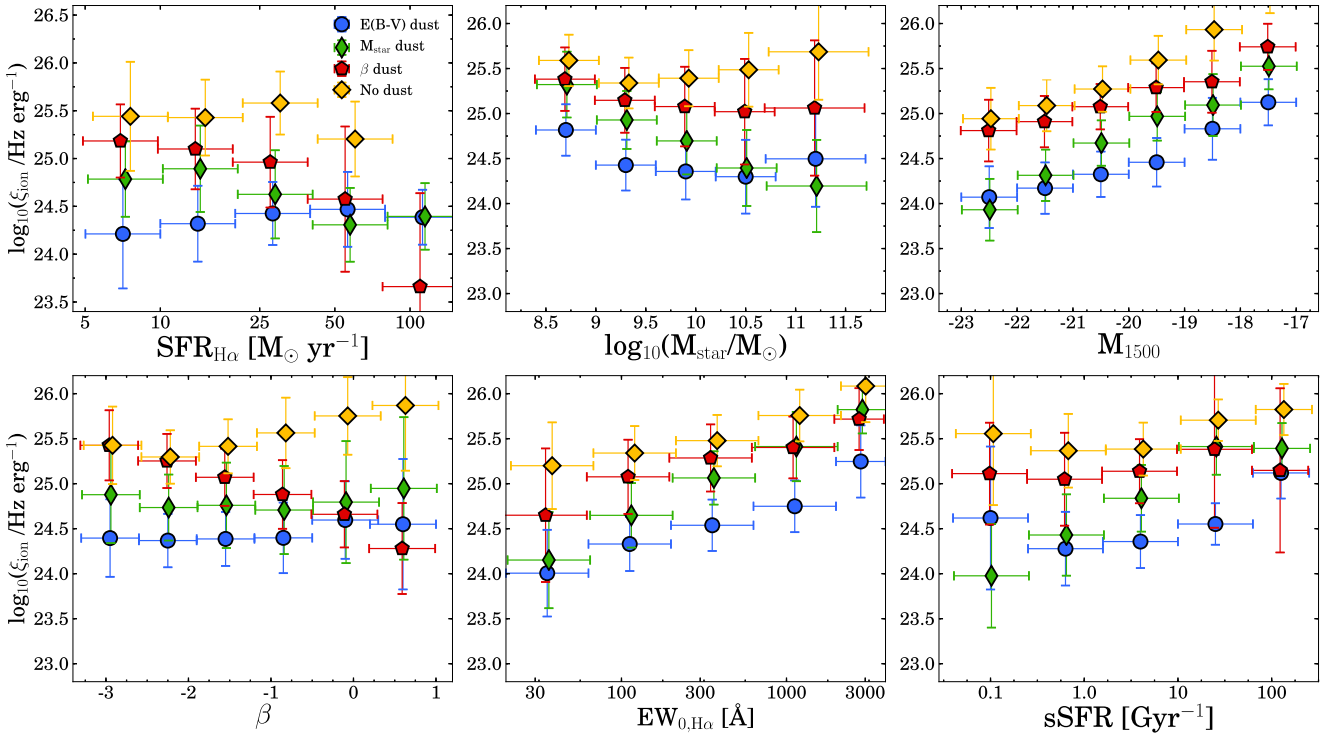

Figure 7. Correlations between $\xi_{\rm ion}$ and galaxy properties for different methods to correct for dust attenuation. To facilitate the comparison, HAEs were binned on the x-axis. The value of $\xi_{\rm ion}$ is the median value in each bin, while the vertical error is the standard deviation. Blue bins show the values where dust is corrected with the $E(B-V)$ value from the SED. The red bins are corrected for dust with the Meurer et al. (1999) prescription based on β and the green bins are corrected for dust with the prescription from Garn & Best (2010) based on stellar mass. Yellow bins show the results where we assume that there is no dust.

The stellar mass and β are not by definition directly related to $\xi_{\rm ion}$. Therefore, a possible upturn of $\xi_{\rm ion}$ at low masses (see the middle-top panel in Fig. 6) may be a real physical effect, although we note that we are not mass-complete below $M_{\rm star} < 10^{10}$ M$_\odot$ and an Hα selected sample of galaxies likely misses low-mass galaxies with lower values of $\xi_{\rm ion}$.

We find that the number of ionizing photons per unit UV luminosity is strongly related to the Hα EW (with a slope of ∼0.6 in log–log space), see Fig. 6. Such a correlation is expected because of our definition of $\xi_{\rm ion}$: (i) the Hα EW increases mildly with increasing Hα (line-)luminosity and (ii) the Hα EW is weakly anti-related with the UV (continuum) luminosity, such that $\xi_{\rm ion}$ increases relatively strongly with EW. Since there is a relation between Hα EW and specific SFR (sSFR = SFR/$M_{\rm star}$, e.g. Fumagalli et al. 2012), we also find that $\xi_{\rm ion}$ increases strongly with increasing sSFR, see Fig. 6.

In Fig. 7, we show the same correlations as discussed above, but now compare the results for different methods to correct for dust. For comparison, we only show the median $\xi_{\rm ion}$ in bins of the property on the x-axis. The vertical error on the bins is the standard deviation of the values of $\xi_{\rm ion}$ in the bin. As $\xi_{\rm ion}$ depends on the dust correction, we find that $\xi_{\rm ion}$ correlates with the galaxy property that was used to correct for dust in the case of β (red symbols) and $M_{\rm star}$ (green symbols). Specific SFR depends on stellar mass, so we also find the strongest correlation between sSFR and $\xi_{\rm ion}$ when $\xi_{\rm ion}$ is corrected for dust with the Garn & Best (2010) prescription. We only find a relation between $\xi_{\rm ion}$ and β when dust is corrected with the Meurer et al. (1999) prescription. For UV magnitude only, the normalization of $\xi_{\rm ion}$ changes with the dust correction method.

It is more interesting to look at correlations between $\xi_{\rm ion}$ and galaxy properties which are not directly related to the computation of $\xi_{\rm ion}$ or the dust correction. Hence, we note that irrespective of the dust correction method, $\xi_{\rm ion}$ appears to be somewhat higher for lower mass galaxies (although this is likely a selection effect as discussed above). Irrespective of the dust correction method, $\xi_{\rm ion}$ increases with increasing Hα EW and fainter M_{1500}, where the particular dust correction method used only sets the normalization. We return to this relation between $\xi_{\rm ion}$ and Hα EW in Section 6.5.

6.4 Redshift evolution of $\xi_{\rm ion}$

Because of its dependency on galaxy properties, it is possible that $\xi_{\rm ion}$ evolves with redshift. In fact, such an evolution is expected as more evolved galaxies (particularly with declining star formation histories) have a relatively stronger UV luminosity than Hα and a higher dust content, likely leading to a lower $\xi_{\rm ion}$ at $z = 2.2$ than at $z > 6$.

By comparing our measurement of $\xi_{\rm ion}$ with those from Bouwens et al. (2016) at $z = 4$–5, we already find such an evolution (see Table 3), although we note that the samples of galaxies are selected differently and that there are many other differences, such as the dust attenuation, typical stellar mass and the Hα measurement. If we mimic a Lyman-break selected sample by only selecting HAEs with $E(B-V) < 0.3$ (typical for UV selected galaxies, e.g. Steidel et al. 2011), we find that $\xi_{\rm ion}$ increases by (maximally) 0.1 dex, such that this does likely not explain the difference in $\xi_{\rm ion}$ at $z = 2.2$ and $z \approx 4$–5 of ≈0.5 dex. Furthermore, our Hα selection is likely biased towards high values of $\xi_{\rm ion}$ for $M_{1500} > -20$, which mitigates the difference on the median $\xi_{\rm ion}$. If we select only low-mass galaxies such that the median stellar mass resembles that of Bouwens et al. (2016), the difference is only ≈0.2 ± 0.1 dex, which still would suggest evolution.

Table 4. Fit parameters for $\log_{10}\xi_{\rm ion} = a + b\,\log_{10}{\rm EW(H}\alpha)$ for different selections and dust corrections.

Sample	$\langle M_{\rm star}\rangle$ $\log_{10}{\rm M}_\odot$	*a*	*b*	Dust
All HAEs	9.8	23.12	0.59	$E(B-V)$
		23.66	0.64	β
		22.60	0.97	$M_{\rm star}$
		23.59	0.45	$A_{\rm H\alpha}=1$
Low mass	9.2	22.64	0.78	$E(B-V)$
		23.68	0.64	β
		23.19	0.77	$M_{\rm star}$
		22.77	0.75	$A_{\rm H\alpha}=1$

We estimate the redshift evolution of $\xi_{\rm ion}$ by combining the relation between $\xi_{\rm ion}$ and Hα EW with the redshift evolution of the Hα EW. Several studies have recently noted that the Hα EW (and related sSFR) increases with increasing redshift (e.g. Fumagalli et al. 2012; Sobral et al. 2014; Smit et al. 2014; Mármol-Queraltó et al. 2016; Faisst et al. 2016; Khostovan et al. 2016). Furthermore, the EW is mildly dependent on stellar mass as EW $\sim {\rm M}_{\rm star}^{-0.25}$ (Sobral et al. 2014; Mármol-Queraltó et al. 2016). In order to estimate the $\xi_{\rm ion}$ using the Hα EW evolution, we:

(i) Select a subset of our HAEs with stellar mass between $10^{9-9.4}$ M$_\odot$, with a median of $M_{\rm star} \approx 10^{9.2}$ M$_\odot$, which is similar to the mass of the sample from Bouwens et al. (2016), see Smit et al. (2015).

(ii) Fit a linear trend between $\log_{10}$(EW) and $\log_{10}(\xi_{\rm ion})$ (with the Garn & Best 2010 prescription to correct for dust attenuation). We note that the trend between EW and $\xi_{\rm ion}$ will be steepened if dust is corrected with a prescription based on stellar mass (since Hα EW anti-correlates with stellar mass, see also Table 4). However, this is validated by several independent observations from either *Herschel* or Balmer decrements which confirm that dust attenuation increases with stellar mass at a wide range of redshifts (Domínguez et al. 2013; Buat et al. 2015; Koyama et al. 2015; Pannella et al. 2015; Sobral et al. 2016b).

Using a simple least-squares algorithm, we find:

$$\log_{10}(\xi_{\rm ion}) = 23.19^{+0.09}_{-0.09} + 0.77^{+0.04}_{-0.04} \times \log_{10}({\rm EW}) \quad (8)$$

(iii) Combine the trend between Hα EW and redshift with the trend between $\xi_{\rm ion}$ and Hα EW. We use the redshift evolution of the Hα EW from Faisst et al. (2016), which has been inferred from fitting SEDs, and measured up to $z \approx 6$. In this parametrization, the slope changes from EW$\approx(1+z)^{1.87}$ at $z < 2.2$ to EW$\approx(1+z)^{1.3}$ at $z > 2.2$. Below $z < 2.2$, this trend is fully consistent with the EW evolution from HiZELS (Sobral et al. 2014), which is measured with NB imaging. Although HiZELS does not have HAEs at $z > 2.2$, the EW evolution of [OIII]+Hβ is found to flatten at $z > 2.2$ as well (Khostovan et al. 2016). We note that we assume that the slope of the Hα EW evolution with redshift does not vary strongly for stellar masses between $10^{9.2}$ and $10^{9.8}$ M$_\odot$, since the following equations are measured at stellar mass $\approx 10^{9.6}$ M$_\odot$ (Faisst et al. 2016), hence:

$${\rm EW}(z) = \begin{cases} 20 \times (1+z)^{1.87}, & z < 2.2 \\ 37.4 \times (1+z)^{1.3}, & z \geq 2.2 \end{cases} \quad (9)$$

This results in:

$$\log_{10}(\xi_{\rm ion}(z)) = \begin{cases} 24.19 + 1.44 \times \log_{10}(1+z), & z < 2.2 \\ 24.40 + 1.00 \times \log_{10}(1+z), & z \geq 2.2 \end{cases} \quad (10)$$

where $\xi_{\rm ion}$ is in Hz erg^{-1}. The error on the normalization is 0.09 Hz erg^{-1} and the error on the slope is 0.18. For our typical mass of $M_{\rm star} = 10^{9.8}$ M$_\odot$, the normalization is roughly 0.2 dex lower and the slope a factor $\approx$1.1 higher compared to the fit at lower stellar masses. This is due to a slightly different relation between $\xi_{\rm ion}$ and EW (see Table 4). The evolving $\xi_{\rm ion}$ is consistent with the typically assumed value of $\xi_{\rm ion} = 10^{25.2\pm0.1}$ Hz erg^{-1} (e.g. Robertson et al. 2013) at $z \approx 2.5$–12 within the 1σ error bars.

We show the inferred evolution of $\xi_{\rm ion}$ with redshift in Fig. 8. The solid and dashed lines use the EW(z) evolution from Faisst et al. (2016), while the dotted line uses the Khostovan et al. (2016) parametrization. The grey shaded region indicates the errors on the redshift evolution of $\xi_{\rm ion}$. Due to the anti-correlation between EW and stellar mass, galaxies with a lower stellar mass have a higher $\xi_{\rm ion}$ (which is then even strengthened by a higher dust attenuation at high masses).

Relatively independent of the dust correction (as discussed in Fig. B1), the median $\xi_{\rm ion}$ increases $\approx$0.2 dex at fixed stellar mass between $z = 2.2$ and 4.5. This can easily explain the 0.2 dex difference between our measurement at $z = 2.2$ and the Bouwens et al. (2016) measurements at $z = 4$–5 (see Fig. 8), such that it is plausible that $\xi_{\rm ion}$ evolves to higher values in the reionization epoch, of roughly $\xi_{\rm ion} \approx 10^{25.4}$ Hz erg^{-1} at $z \approx 8$. Interestingly, LAEs at $z = 2.2$ already have $\xi_{\rm ion}$ similar to the canonical value in the reionization era.

7 IMPLICATIONS FOR REIONIZATION

The product of $f_{\rm esc}\xi_{\rm ion}$ is an important parameter in assessing whether galaxies have provided the photons to reionize the Universe, because these convert the (non-ionizing) UV luminosity density (obtained from integrating the dust-corrected UV luminosity function) to the ionizing emissivity. The typical adopted values are $\xi_{\rm ion} \approx 10^{25.2-25.3}$ Hz erg^{-1} and $f_{\rm esc} \approx 0.1$–0.2 (e.g. Robertson et al. 2015), such that the product is $f_{\rm esc}\xi_{\rm ion} \approx 10^{24.2-24.6}$ Hz erg^{-1}. This is significantly higher than our upper limit of $f_{\rm esc}\xi_{\rm ion} \lesssim 10^{23.5}$ Hz erg^{-1} (using $\langle f_{\rm esc}\rangle$ and $\xi_{\rm ion}$ where dust is corrected with $M_{\rm star}$, see Sections 5 and 6). However, as shown in Section 6.5, we expect $\xi_{\rm ion} \approx 10^{25.4}$ Hz erg^{-1} in the reionization era due to the dependency of $\xi_{\rm ion}$ on EW(Hα), such that escape fractions of $f_{\rm esc} \approx 7$ per cent would suffice for $f_{\rm esc}\xi_{\rm ion} = 10^{24.2}$ Hz erg^{-1}. Becker & Bolton (2013) find an evolution in the product of $f_{\rm esc}\xi_{\rm ion}$ of a factor 4 between $z = 3$ and 5 (similar to Haardt & Madau 2012), which is consistent with our measurements.

Recently, Faisst (2016) inferred that $f_{\rm esc}$ may evolve with redshift by combining a relation between $f_{\rm esc}$ and the [OIII]/[OII] ratio with the inferred redshift evolution of the [OIII]/[OII] ratio. This redshift evolution is estimated from local analogs to high-redshift galaxies selected on Hα EW, such that the redshift evolution of $f_{\rm esc}$ is implicitly coupled to the evolution of Hα EW as in our model of $\xi_{\rm ion}(z)$. Faisst (2016) estimates that $f_{\rm esc}$ evolves from $\approx$2 per cent at $z = 2$ to $\approx$5 per cent at $z = 5$, which is consistent with our measurements of $\langle f_{\rm esc}\rangle$ (see Fig. 4). With this evolving escape fraction, galaxies can provide sufficient amounts of photons to reionize the Universe, consistent with the most recent Cosmic Microwave Background (CMB) constraints (Planck Collaboration XLVII et al. 2016). This calculation assumes $\xi_{\rm ion} = 10^{25.4}$ Hz erg^{-1}, which is the same value our model predicts for $\xi_{\rm ion}$ in the reionization era.

In addition to understanding whether galaxies have reionized the Universe, it is perhaps more interesting to understand which galaxies have been the most important to do so. For example, Sharma et al.

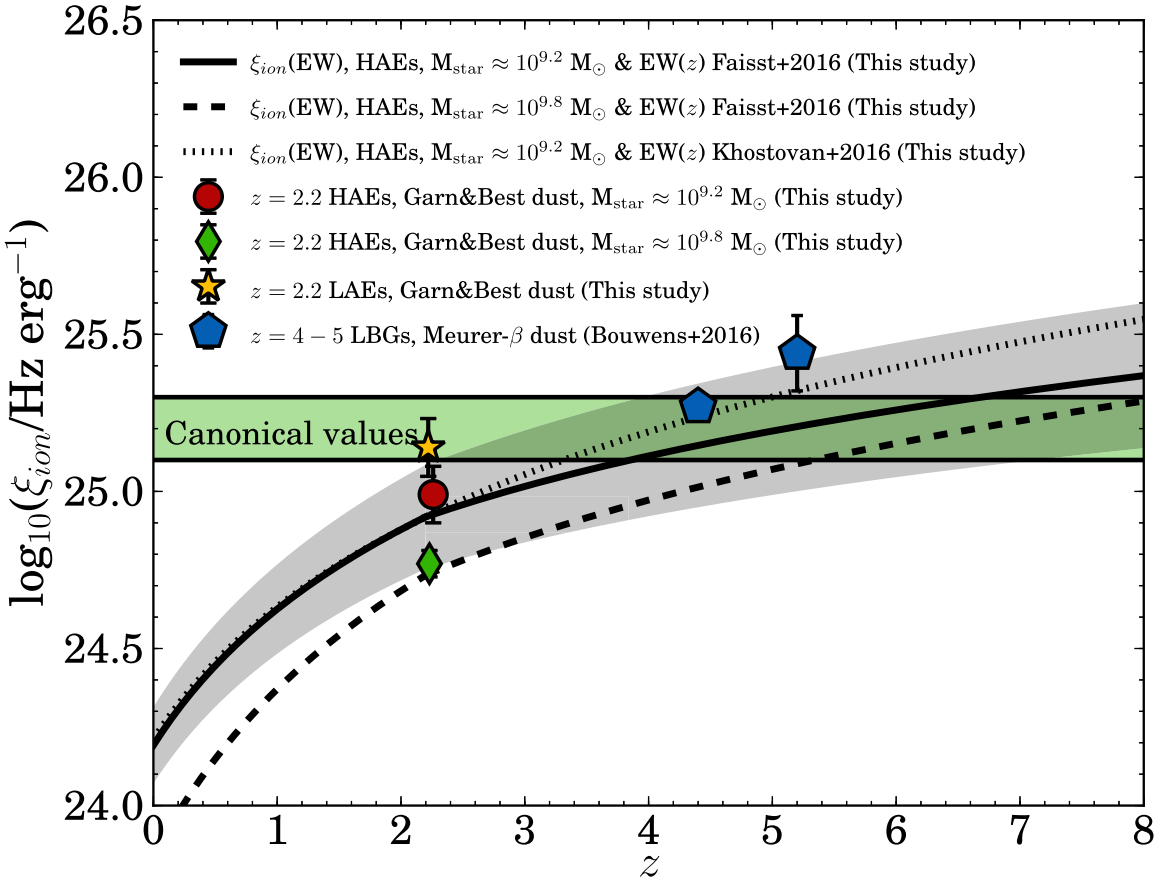

Figure 8. Inferred evolution of $\xi_{\rm ion}$ (corrected for dust with $M_{\rm star}$) with redshift based on our observed trend between $\xi_{\rm ion}$ and Hα EW, for different stellar masses (compare the solid with the dashed line) and EW(z) evolutions (compare the solid with the dotted line). The grey shaded region indicates the errors on the redshift evolution of $\xi_{\rm ion}$. The normalization of $\xi_{\rm ion}$ is higher for lower mass galaxies or LAEs. The green region shows the typically assumed values. The estimated evolution of $\xi_{\rm ion}$ with redshift is consistent with the typically assumed values of $\xi_{\rm ion}$ in the reionization era and with recent measurements at $z = 4$–5.

(2016) argue that the distribution of escape fractions in galaxies is likely very bimodal and dependent on the SFR surface density, which could mean that LyC photons preferentially escape from bright galaxies. Such a scenario may agree better with a late and rapid reionization process such as favoured by the new low optical depth measurement from Planck Collaboration XLVII et al. (2016). We note that the apparent discrepancy between the $f_{\rm esc}$ upper limit from median stacking ($f_{\rm esc} < 2.8$ per cent) and the average $f_{\rm esc}$ from the integrated luminosity density combined with IGM constraints ($\langle f_{\rm esc}\rangle = 5.9$ per cent) can be understood in a scenario where the average $f_{\rm esc}$ is driven by a few galaxies with high $f_{\rm esc}$, or by a scenario where $f_{\rm esc}$ is higher for galaxies below the Hα detection threshold (which corresponds to SFR > 4 $M_\odot$ yr^{-1}), contrarily to a scenario where each typical HAE has an escape fraction of ≈5–6 per cent.

Dijkstra, Gronke & Venkatesan (2016) argue a connection between the escape of Lyα photons and LyC photons, such that LAEs could potentially be important contributors to the photon budget in the reionization era (particularly since we find that $\xi_{\rm ion}$ is higher for LAEs than for more normal star-forming galaxies at $z = 2.2$). Hence, LAEs may have been important contributors of the photons that reionized the Universe.

To make progress we need a detailed understanding of the physical processes which drive $f_{\rm esc}$, for which a significant sample of directly detected LyC leakers at a range of redshifts and galaxy properties is required. It is challenging to measure $f_{\rm esc}$ directly at $z > 3$ (and practically impossible at $z > 5$) due to the increasing optical depth of the IGM with redshift, such that indirect methods to estimate $f_{\rm esc}$ may be more successful (e.g. Jones et al. 2013; Zackrisson, Inoue & Jensen 2013; Verhamme et al. 2015). However, the validity of these methods remains to be evaluated (i.e. Vasei et al. 2016).

8 CONCLUSIONS

We have studied the production and escape of ionizing photons (LyC, $\lambda_0 < 912$ Å) for a large sample of Hα selected galaxies at $z = 2.2$. Thanks to the joint coverage of the rest-frame LyC, UV and Hα (and, in some cases, Lyα), we have been able to reliably estimate the intrinsic LyC luminosity and the number of ionizing photons per unit UV luminosity ($\xi_{\rm ion}$), for which we (indirectly) constrained the escape fraction of ionizing photons ($f_{\rm esc}$). Our results are:

(i) We have stacked the *NUV* thumbnails for all HAEs and subsets of galaxies in order to obtain constraints on $f_{\rm esc}$. None of the stacks shows a direct detection of LyC flux, allowing us to place a median (mean) upper limit of $f_{\rm esc} < 2.8\,(6.4)$ per cent for the stack of star-forming HAEs (Section 4.3). A low escape fraction validates our method to estimate $\xi_{\rm ion}$, the production efficiency of ionizing photons.

(ii) Combining the IGM emissivity measurements from Becker & Bolton (2013) with the integrated Hα luminosity function from Sobral et al. (2013) at $z = 2.2$, we find a globally averaged $\langle f_{\rm esc}\rangle = 5.9^{+14.5}_{-4.2}$ per cent at $z = 2.2$ (Section 5), where the errors include

conservative estimates of the systematic uncertainties. Combined with recent estimates of the quasar (QSO) emissivity at $z \approx 2.2$, we can not fully rule out a non-zero contribution from star-forming galaxies to the ionizing emissivity. We speculate that the apparent discrepancy between the $f_{\rm esc}$ upper limit from median stacking and $\langle f_{\rm esc} \rangle$ can be understood in a scenario where the average $f_{\rm esc}$ is driven by a few galaxies with high $f_{\rm esc}$, or by a scenario where $f_{\rm esc}$ is higher for galaxies below the Hα detection threshold (SFR > 4 M$_\odot$ yr^{-1}).

(iii) Applying a similar analysis to published data at $z \approx 4$–5 results in a relatively constant $f_{\rm esc}$ with redshift (see Table 2 and Fig. 4). We rule out $\langle f_{\rm esc} \rangle > 20$ per cent at redshifts lower than $z \approx 5$. An additional contribution of ionizing photons from rare quasars strengthens this constraint.

(iv) We find that $\xi_{\rm ion}$ increases strongly with increasing sSFR and Hα EW and decreasing UV luminosity, independently on the dust correction method. We find no significant correlations between $\xi_{\rm ion}$ and SFR(Hα), β or $M_{\rm star}$. On average, LAEs have a higher $\xi_{\rm ion}$ than HAEs, a consequence of LAEs having typically bluer UV slopes, lower masses and lower values of $E(B-V)$ (Section 6) – properties which are typical for galaxies at the highest redshift.

(v) The median $\xi_{\rm ion}$ of HAEs at $z = 2.2$ is $\xi_{\rm ion} \approx 10^{24.77 \pm 0.04}$ Hz erg^{-1}, which is ≈0.4 dex lower than the typically assumed values in the reionization era or recent measurements at $z \sim 4$–5 (Bouwens et al. 2016), see Table 3. Only half of this difference is explained by the lower stellar mass and dust attenuation of the galaxies in the Bouwens et al. (2016) sample.

(vi) For LAEs at $z = 2.2$ we find a higher $\xi_{\rm ion} = 10^{25.14 \pm 0.09}$ Hz erg^{-1}, already similar to the typical value assumed in the reionization era. This difference is driven by the fact that LAEs are typically less massive and bluer and thus have less dust than HAEs.

(vii) By combining our trend between $\xi_{\rm ion}$ and Hα EW with the redshift evolution of Hα EW, we find that $\xi_{\rm ion}$ increases with ≈0.2 dex between $z = 2.2$ and $z = 4$–5, resulting in perfect agreement with the results from Bouwens et al. (2016). Extrapolating this trend leads to a median value of $\xi_{\rm ion} \approx 10^{25.4}$ Hz erg^{-1} at $z \sim 8$, slightly higher than the typically assumed value in the reionization epoch (Section 7), such that a relatively low global $f_{\rm esc}$ (consistent with our global estimates at $z \approx 2$–5) would suffice to provide the photons to reionize the Universe.

ACKNOWLEDGEMENTS

We thank the referee for the many helpful and constructive comments which have significantly improved this paper. JM acknowledges the support of a Huygens PhD fellowship from Leiden University. DS acknowledges financial support from the Netherlands Organization for Scientific research (NWO) through a Veni fellowship and from FCT through an FCT Investigator Starting Grant and Start-up Grant (IF/01154/2012/CP0189/CT0010). PNB is grateful for support from the UK STFC via grant ST/M001229/1. IO acknowledges support from the European Research Council in the form of the Advanced Investigator Programme, 321302, COSMICISM. The authors thank Andreas Faisst, Michael Rutkowski and Andreas Sandberg for answering questions related to this work and Daniel Schaerer and Mark Dijkstra for discussions. We acknowledge the work that has been done by both the COSMOS team in assembling such large, state-of-the-art multi-wavelength data set, as this has been crucial for the results presented in this paper. We have benefited greatly from the public available programming language PYTHON, including the NUMPY, MATPLOTLIB, PYFITS, SCIPY (Jones et al. 2001; Hunter 2007; Van Der Walt, Colbert & Varoquaux 2011) and ASTROPY (Astropy Collaboration et al. 2013) packages, the astronomical imaging tools SExtractor and Swarp (Bertin & Arnouts 1996; Bertin 2010) and the Topcat analysis program (Taylor 2013).

REFERENCES

Alexandroff R. M., Heckman T. M., Borthakur S., Overzier R., Leitherer C., 2015, ApJ, 810, 104
Astropy Collaboration et al., 2013, A&A, 558, A33
Atek H. et al., 2015, ApJ, 800, 18
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Bergvall N., Zackrisson E., Andersson B.-G., Arnberg D., Masegosa J., Östlin G., 2006, A&A, 448, 513
Bergvall N., Leitet E., Zackrisson E., Marquart T., 2013, A&A, 554, A38
Bertin E., 2010, SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library, record ascl:1010.068
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. et al., 2011, ApJ, 736, 41
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bouwens R. J. et al., 2012, ApJ, 752, L5
Bouwens R. J. et al., 2015a, ApJ, 803, 34
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015b, ApJ, 811, 140
Bouwens R. J., Smit R., Labbe I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 176
Bowler R. A. A. et al., 2014, MNRAS, 440, 2810
Bridge C. R. et al., 2010, ApJ, 720, 465
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buat V. et al., 2015, A&A, 577, A141
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Castellano M. et al., 2016, A&A, 590, A31
Cen R., Kimm T., 2015, ApJ, 801, L25
Chabrier G., 2003, PASP, 115, 763
Chen H.-W., Prochaska J. X., Gnedin N. Y., 2007, ApJ, 667, L125
Choudhury T. R., Puchwein E., Haehnelt M. G., Bolton J. S., 2015, MNRAS, 452, 261
Civano F. et al., 2012, ApJS, 201, 30
Conseil S., Vibert D., Amouts S., Milliard B., Zamojski M., Liebaria A., Guillaume M., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 107
Cooke J., Ryan-Weber E. V., Garel T., Díaz C. G., 2014, MNRAS, 441, 837
Cowie L. L., Barger A. J., Trouille L., 2009, ApJ, 692, 1476
Cristiani S., Serrano L. M., Fontanot F., Vanzella E., Monaco P., 2016, MNRAS, 462, 2478
De Barros S. et al., 2016, A&A, 585, A51
Deharveng J.-M., Buat V., Le Brun V., Milliard B., Kunth D., Shull J. M., Gry C., 2001, A&A, 375, 805
Dijkstra M., Gronke M., Venkatesan A., 2016, ApJ, 828, 71
Domínguez A. et al., 2013, ApJ, 763, 145
Duncan K., Conselice C. J., 2015, MNRAS, 451, 2030
Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813
Faisst A. L., 2016, ApJ, 829, 99
Faisst A. L. et al., 2016, ApJ, 821, 122
Fan X. et al., 2006, AJ, 132, 117
Ferrara A., Loeb A., 2013, MNRAS, 431, 2826
Finkelstein S. L., 2015, Publ. Astron. Soc. Aust., 33, e037
Finkelstein S. L. et al., 2015, ApJ, 810, 71
Förster Schreiber N. M. et al., 2009, ApJ, 706, 1364
Fumagalli M. et al., 2012, ApJ, 757, L22
Garn T., Best P. N., 2010, MNRAS, 409, 421

Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, MNRAS, 388, 1473
Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, MNRAS, 426, 679
Giallongo E. et al., 2015, A&A, 578, A83
Grazian A. et al., 2016, A&A, 585, A48
Guaita L. et al., 2016, A&A, 587, A133
Haardt F., Madau P., 2012, ApJ, 746, 125
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ibar E. et al., 2013, MNRAS, 434, 3218
Ilbert O. et al., 2009, ApJ, 690, 1236
Inoue A. K., 2002, ApJ, 570, L97
Inoue A. K., Iwata I., Deharveng J.-M., 2006, MNRAS, 371, L1
Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, MNRAS, 442, 1805
Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2015, ApJ, 799, 12
Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitova I., Verhamme A., 2016a, MNRAS, 461, 3683
Izotov Y. I., Orlitová I., Schaerer D., Thuan T. X., Verhamme A., Guseva N. G., Worseck G., 2016b, Nature, 529, 178
Jones E., Oliphant T., Peterson P. et al., 2001, SciPy: Open source scientific tools for Python, http://www.scipy.org/ (accessed 2016 May 9)
Jones T. A., Ellis R. S., Schenker M. A., Stark D. P., 2013, ApJ, 779, 52
Kashino D. et al., 2013, ApJ, 777, L8
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
Khaire V., Srianand R., Choudhury T. R., Gaikwad P., 2016, MNRAS, 457, 4051
Khostovan A. A., Sobral D., Mobasher B., Smail I., Darvish B., Nayyeri H., Hemmati S., Stott J. P., 2016, MNRAS, 463, 2363
Koekemoer A. M. et al., 2007, ApJS, 172, 196
Koyama Y. et al., 2015, MNRAS, 453, 879
Kuhlen M., Faucher-Giguère C.-A., 2012, MNRAS, 423, 862
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. et al., 1999, ApJS, 123, 3
Leitherer C., Hernandez S., Lee J. C., Oey M. S., 2016, ApJ, 823, 64
Lilly S. J. et al., 2009, ApJS, 184, 218
Livermore R. C., Finkelstein S. L., Lotz J. M., 2016, ApJ, preprint (arXiv:1604.06799)
Ma X., Kasen D., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Kereš D., Murray N., 2015, MNRAS, 453, 960
Madau P., Haardt F., 2015, ApJ, 813, L8
Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648
Mármol-Queraltó E., McLure R. J., Cullen F., Dunlop J. S., Fontana A., McLeod D. J., 2016, MNRAS, 460, 3587
Martin C. et al., 2003, in Blades J. C., Siegmund O. H. W., eds, Proc. SPIE Vol. 4854, Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation. SPIE, Bellingham, p. 336
Martin D. C. et al., 2005, ApJ, 619, L1
Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, MNRAS, 451, 400
Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, MNRAS, 458, 449
McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499
McLeod D. J., McLure R. J., Dunlop J. S., Robertson B. E., Ellis R. S., Targett T. A., 2015, MNRAS, 450, 3032
McLure R. J. et al., 2013, MNRAS, 432, 2696
Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64
Micheva G., Iwata I., Inoue A. K., 2016, MNRAS, preprint (arXiv:1604.00102)
Mitra S., Ferrara A., Choudhury T. R., 2013, MNRAS, 428, L1
Mitra S., Choudhury T. R., Ferrara A., 2015, MNRAS, 454, L76
Mostardi R. E., Shapley A. E., Steidel C. C., Trainor R. F., Reddy N. A., Siana B., 2015, ApJ, 810, 107
Mowlavi N., Eggenberger P., Meynet G., Ekström S., Georgy C., Maeder A., Charbonnel C., Eyer L., 2012, A&A, 541, A41
Muzzin A. et al., 2013, ApJS, 206, 8
Nakajima K. et al., 2012, ApJ, 745, 12
Oteo I., Sobral D., Ivison R. J., Smail I., Best P. N., Cepa J., Pérez-García A. M., 2015, MNRAS, 452, 2018
Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544
Pannella M. et al., 2015, ApJ, 807, 141
Pawlik A. H., Schaye J., Dalla Vecchia C., 2015, MNRAS, 451, 1586
Planck Collaboration XLVII et al., 2016, A&A, preprint (arXiv:1605.03507)
Price L. C., Trac H., Cen R., 2016, preprint (arXiv:1605.03970)
Puglisi A. et al., 2016, A&A, 586, A83
Reddy N. A. et al., 2015, ApJ, 806, 259
Reddy N. A., Steidel C. C., Pettini M., Bogosavljević M., 2016, ApJ, 828, 107
Robertson B. E. et al., 2013, ApJ, 768, 71
Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19
Rutkowski M. J. et al., 2016, ApJ, 819, 81
Sandberg A., Östlin G., Melinder J., Bik A., Guaita L., 2015, ApJ, 814, L10
Sanders R. L. et al., 2015, ApJ, 799, 138
Schaerer D., 2002, A&A, 382, 28
Schaerer D., 2003, A&A, 397, 527
Sharma M., Theuns T., Frenk C., Bower R., Crain R., Schaller M., Schaye J., 2016, MNRAS, 458, 94
Shivaei I., Reddy N. A., Steidel C. C., Shapley A. E., 2015, ApJ, 804, 149
Siana B. et al., 2007, ApJ, 668, 62
Siana B. et al., 2015, ApJ, 804, 17
Smit R. et al., 2014, ApJ, 784, 58
Smit R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2015, ApJ, preprint (arXiv:1511.08808)
Smith B. M. et al., 2016, ApJ, preprint, (arXiv:1602.01555)
Sobral D. et al., 2009, MNRAS, 398, L68
Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926
Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, MNRAS, 428, 1128
Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516
Sobral D. et al., 2015, MNRAS, 451, 2303
Sobral D. et al., 2016a, MNRAS, preprint (arXiv:1609.05897)
Sobral D., Stroe A., Koyama Y., Darvish B., Calhau J., Afonso A., Kodama T., Nakata F., 2016b, MNRAS, 458, 3443
Steidel C. C., Pettini M., Adelberger K. L., 2001, ApJ, 546, 665
Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, ApJ, 736, 160
Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012, MNRAS, 426, 935
Taniguchi Y. et al., 2007, ApJS, 172, 9
Taylor M., 2013, Starlink User Note, 253
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, preprint (arXiv:1102.1523)
Vanzella E. et al., 2012, ApJ, 751, 70
Vanzella E. et al., 2016, ApJ, 825, 41
Vasei K. et al., 2016, Astrophys. J., preprint (arXiv:1603.02309)
Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, A&A, 578, A7
Weigel A. K., Schawinski K., Treister E., Urry C. M., Koss M., Trakhtenbrot B., 2015, MNRAS, 448, 3167
Wuyts S. et al., 2011, ApJ, 738, 106
Zackrisson E., Inoue A. K., Jensen H., 2013, ApJ, 777, 39
Zamojski M. A. et al., 2007, ApJS, 172, 468

APPENDIX A: INDIVIDUAL *NUV* DETECTIONS

We search for individual galaxies possibly leaking LyC photons by matching our CLEAN galaxy sample with the public *GALEX* EM cleaned catalogue (e.g. Zamojski et al. 2007; Conseil et al. 2011), which is *U* band detected. In total, we find 19 matches between

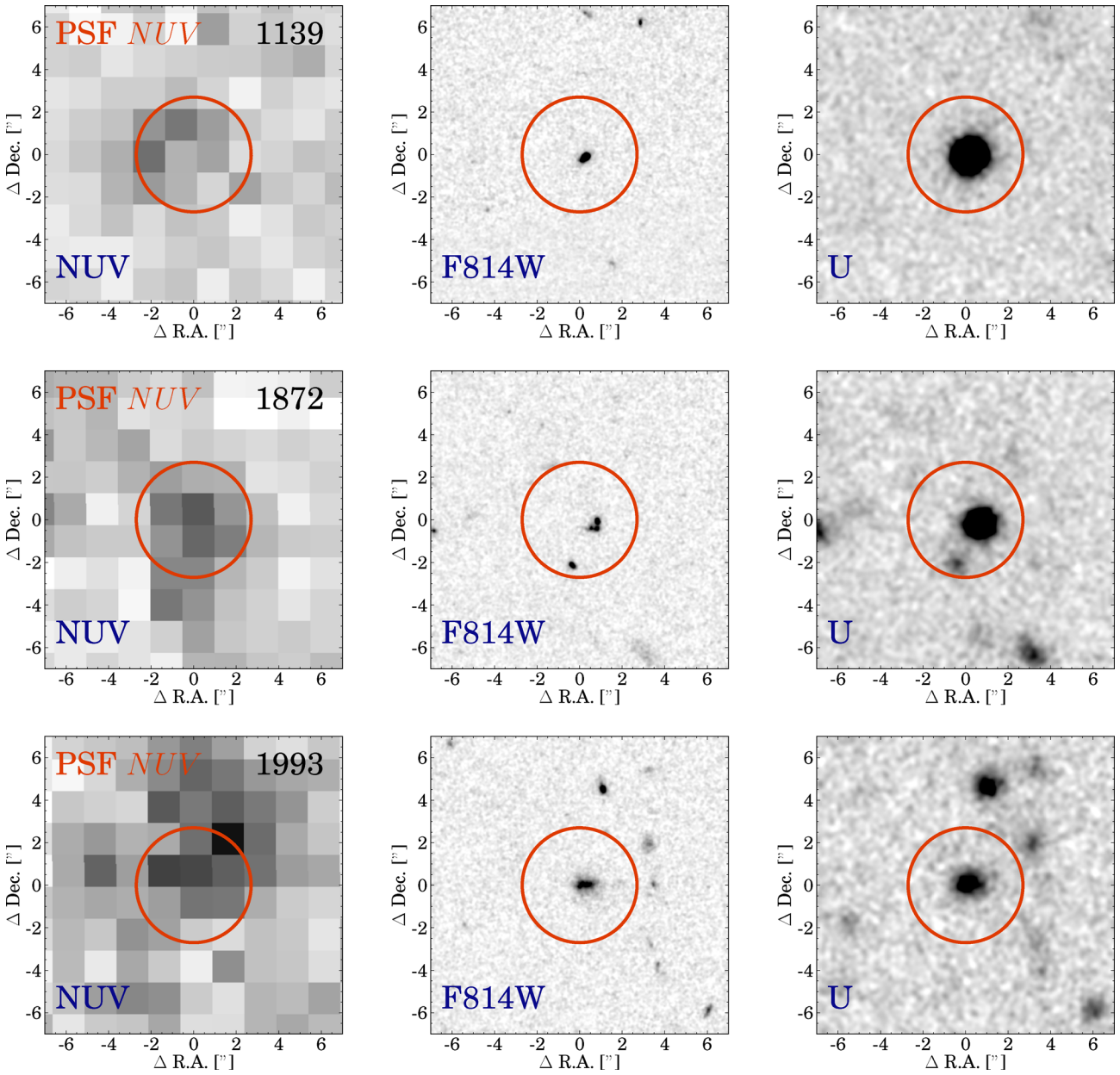

Figure A1. 15 arcsec × 15 arcsec thumbnail images in *NUV*, *F*814*W* and *U* of candidate LyC leaking HAEs and LAEs at $z = 2.2$, centred on the positions of the HAE/LAE. The images are annotated with the IDs of the galaxies in the HiZELS catalogue (Sobral et al. 2013). LAEs are identified with a 'C'. IDs 1139, 1993 and 7801 are detected in both Hα and Lyα. IDs 1139, 7801, C-8 and C-10 are X-ray AGN. All other sources than the central source seen in thumbnails have photometric redshifts of <1.5.

Clean HAEs and *GALEX* sources with *NUV* < 26 within 1 arcsec (33 matches when using all HAEs), and nine matches between LAEs and *GALEX* sources (four out of these nine are also in the HAE sample and we will discuss these as HAEs). By visual inspection of the *HST*/ACS *F814W* and CFHT/*U* band imaging, we mark 8/19 HAEs and 2/5 LAEs as reliable *NUV* detections. The 14 matches that we discarded were either unreliable detections in *NUV* (nine times, caused by local variations in the depth, such that the detections are at 2σ level) or a fake source in *NUV* (five times, caused by artefacts of bright objects). We note however that in most of the remaining 10 *NUV* detections (eight HAEs and two LAEs) the *NUV* photometry is blended with a source at a distance of ≈4 arcsec, see Fig. A1.

In order to get a first-order estimate of the contamination from neighbouring sources to the *NUV* flux, we perform the following simulation. First, we simulate the *NUV* flux of the candidate LyC leakers and all sources within 10 arcsec by placing Moffat flux distributions with the PSF-FWHM of *NUV* imaging and $\beta = 3$. These flux distributions are normalized by the *U*-band magnitude of each source, since the catalogue that we use to measure *NUV* imaging uses *U*-band imaging as a prior. We then measure the fraction of the flux that is coming from neighbouring sources within an aperture with radius 0.67 × FWHM centred at the position of the *NUV* detection of the candidate LyC leaker. We find that contamination for most candidates is significant, and remove three candidates for

Table A1. Candidate LyC leakers among the HAE/LAE sample. ID numbers of HAEs refer to the IDs in the HiZELS catalogue (Sobral et al. 2013). IDs indicated with an * are X-ray AGN. The coordinates correspond to the peak of Hα/Lyα emission. The redshift is either spectroscopic (^{s}), photometric (^{p}) or from a dual-NB emission-line confirmation (^{d}). The *NUV* contamination fraction is estimated as described in the text. f_{esc} is corrected for contamination from nearby sources to the *NUV* flux. Because of the absence of Hα measurements for LAEs, we do not estimate the SFR(Hα) or f_{esc}.

ID	RA (J2000)	Dec. (J2000)	Redshift	M_{star} $\log_{10}(M_\odot)$	SFR(Hα) ($M_\odot$ yr^{-1})	M_{1500} (mag)	*NUV* (mag)	*NUV* contamination	f_{esc} (per cent)
1139*	10:00:55.39	+01:59:55.39	2.219^s	10.1	34.8	−21.6	25.9	0.0	30
1872	10:01:56.39	+02:17:36.65	2.22 ± 0.02^p	9.4	9.2	−21.0	25.7	0.14	43
1993	10:02:08.70	+02:21:19.88	2.22 ± 0.01^d	9.6	8.2	−21.3	24.6	0.39	45
2258	10:01:29.69	+02:24:28.50	2.22 ± 0.02^p	10.3	7.3	−21.0	25.1	0.21	43
7801*	10:02:08.55	+01:45:53.60	2.215^s	10.4	43.3	−23.5	24.9	0.05	37
C8*	09:59:34.82	+02:02:49.94	2.182^s	10.9		−22.5	24.6	0.03	
C10*	09:59:05.14	+02:15:29.86	2.222^s	10.6		−23.5	23.7	0.03	

which the estimated contamination is larger than >50 per cent. The remaining candidates have contaminations ranging from 0 to 39 per cent and we subtract this contamination from the measured *NUV* flux when estimating their escape fractions. We estimate the uncertainty in our contamination estimate due to variations in the PSF and in the flux normalization (due to $NUV - U$ colours) as follows: we first simulate the contamination with a Gaussian PSF and Moffat PSFs with increasing β up to $\beta = 7$ and also by correcting the *U*-band magnitude prior with the observed $U - B$ and $NUV - U$ colours. We then estimate the systematic uncertainty by measuring the standard deviation of the contamination rates estimated with the different simulations. For sources with little contamination, the systematic uncertainty in the contamination estimate is of the order 5 per cent.

We test whether the *NUV* detections for these sources could arise solely from flux at $\lambda_0 > 912$ Å in the far red wing of the *NUV* filter (see Section 3.2). For each galaxy, we obtain the best-fitting STARBURST99 model by matching the Hα EW, as Hα EW is most strongly related to the SED shape around 900 Å. We redshift this model to a redshift of 2.22 and normalize the SED to reproduce the *V*-band magnitude (we assume zero dust attenuation, which is a conservative assumption for this analysis, see below) and convolve the model with the mean IGM transmission at $z = 2.22$. Then, we measure the predicted *NUV* magnitude in the case that the flux is only non-zero at $\lambda_0 > 912$ Å. We find that, in all cases, this magnitude is too faint to explain the *NUV* detections, ranging from *NUV* = 30.1 to 32.5, well below the detection limits. In the presence of dust, the attenuation at $\lambda \sim 912$–930 Å is stronger than at $\lambda \sim 1600$ Å (e.g. Reddy et al. 2016), such that the predicted *NUV* magnitude would be even fainter. We test the robustness of this estimate by varying the SED models (lowering the Hα EWs), neglecting the IGM absorption or by perturbing the redshift between $z = 2.20$ and 2.24, but find that this changes the result only by up to 1 mag if all effects are combined. For ID 1139 and 7801, we also test simple AGN power-law models ($f_\lambda \propto \lambda^\beta$) with UV slopes ranging from −2.0 to −2.7, but find that pure non-ionizing flux cannot explain the *NUV* photometry. Therefore, it is unlikely that the *NUV* detections arise purely from flux at $\lambda_0 > 912$ Å, just because the filter transmission at these wavelengths is very low, and the wavelength range constitutes only a fraction of the full filter width.

For the five candidate LyC leakers with Hα measurements, we measure escape fractions ranging from ≈35 to 46 per cent, see Table A1, although we note that these escape fractions are still uncertain due to (i) possible underestimated foreground contamination from sources not detected in *U* (or not detected as individual source due to blending) or with very blue $NUV - U$ colours, (ii) uncertain dust attenuation of the Hα luminosity, (iii) underestimated contribution from flux at $\lambda_0 > 912$ Å due to different SED shapes than expected or (photometric) redshift errors. Observations with higher spatial resolution and detailed spectroscopy are required in order to confirm whether these seven candidates are really leaking LyC photons and at what rate.

Four isolated LyC leaker candidates (including two LAEs) are X-ray AGN, and all have been spectroscopically confirmed at $z = 2.2$ (Lilly et al. 2009; Civano et al. 2012). HiZELS-ID 1993 is detected in two other NBs than the Hα NB: Lyα ($EW_{0,Ly\alpha} = 67$ Å) and [OIII] ($EW_{0,[OIII]} > 100$ Å), and is thus known to be at $z = 2.22 \pm 0.01$ very robustly. ID 1872 and 2258 are selected as HAE at $z = 2.2$ based on their photometry (see Sobral et al. 2013), such that it is possible that they are interlopers (with the second most likely emission-line being [OIII] at $z \sim 3.3$, but other rarer possibilities such as Paschen series lines at $z < 1$). We show thumbnail images of our candidate LyC leakers in the *NUV*, *F*814*W* and *U* bands in Figs A1 and A2.

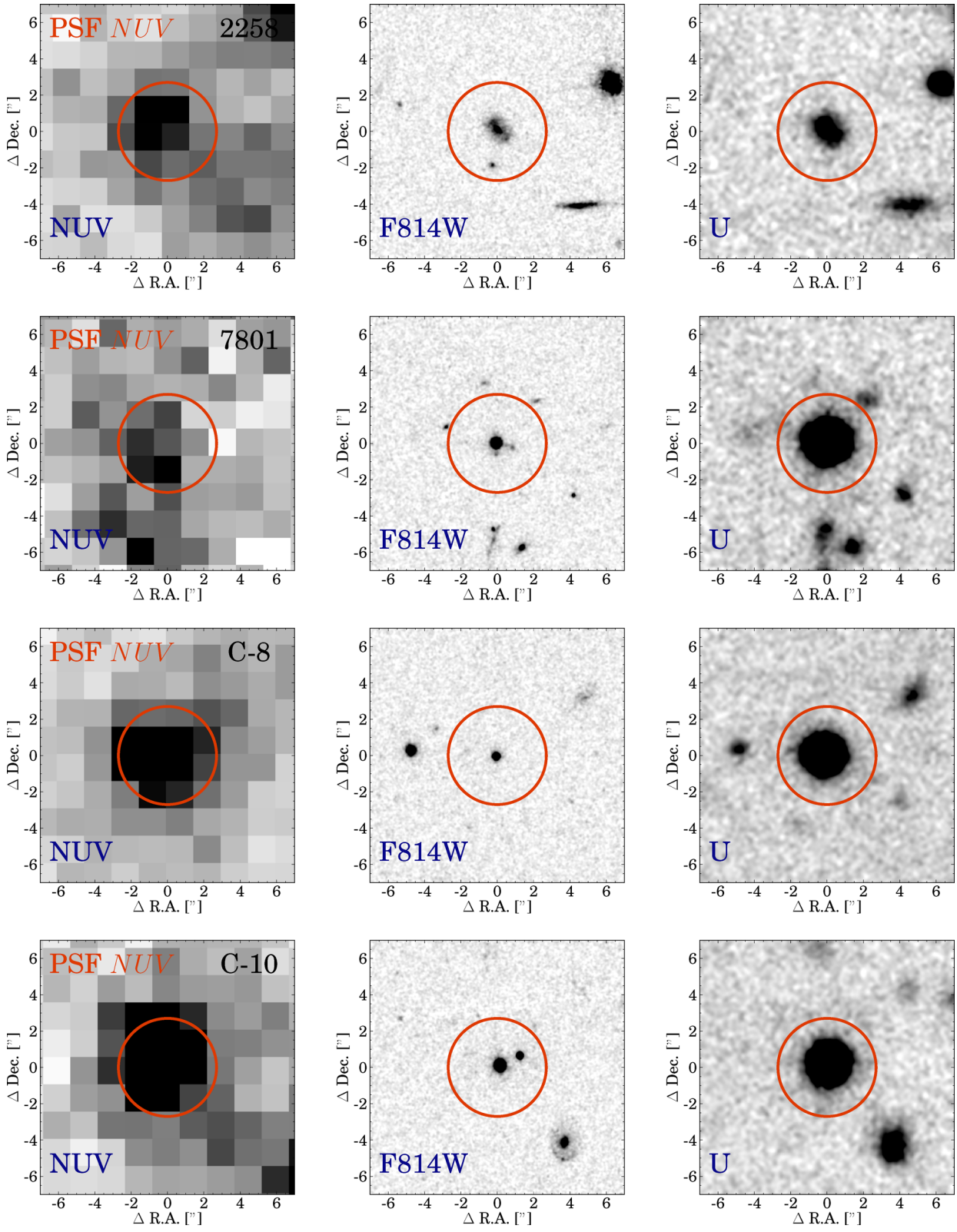

Figure A2. Continued from Fig. A1.

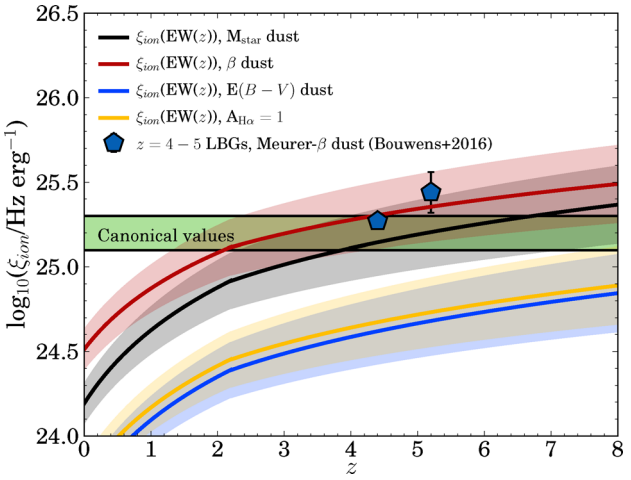

Figure B1. Inferred evolution of $\xi_{\rm ion}$ with redshift based on the EW(Hα) evolution from Faisst et al. (2016) and our observed trend between $\xi_{\rm ion}$ and Hα EW for HAEs with $M_{\rm star} \sim 10^{9.2}$ M$_{\odot}$, for different methods to correct for dust. The black line shows the results when correcting for dust with $M_{\rm star}$, the red line shows dust corrected with β, the blue line shows dust corrected with the $E(B-V)$ values from SED fitting and the yellow line shows the results when we apply a global correction of $A_{\rm H\alpha}=1$. The shaded regions indicate the errors on the redshift evolution of $\xi_{\rm ion}$.

APPENDIX B: REDSHIFT EVOLUTION OF $\xi_{\rm ion}$ WITH DIFFERENT DUST CORRECTIONS

In Fig. B1, we show the inferred redshift evolution of $\xi_{\rm ion}$ when we apply different methods to correct $\xi_{\rm ion}$ for dust. Most of the differences are caused by a varying normalization of $\xi_{\rm ion}$, since we find that the slope of the fit between $\xi_{\rm ion}$ and Hα EW varies only mildly for various dust correction methods, see Table 4. However, we note again that most independent (stacking) observations from Balmer decrements and *Herschel* prefer dust attenuations similar to the dust attenuation we use when correcting for dust with stellar mass.

This paper has been typeset from a TEX/LATEX file prepared by the author.