

# Closing in on the sources of cosmic reionization: First results from the GLASS-JWST program

S. Mascia<sup>1,2</sup>, L. Pentericci<sup>1</sup>, A. Calabrò<sup>1</sup>, T. Treu<sup>3</sup>, P. Santini<sup>1</sup>, L. Yang<sup>4</sup>, L. Napolitano<sup>1</sup>, G. Roberts-Borsani<sup>3</sup>, P. Bergamini<sup>5,6</sup>, C. Grillo<sup>5,7</sup>, P. Rosati<sup>6,8</sup>, B. Vulcani<sup>9</sup>, M. Castellano<sup>1</sup>, K. Boyett<sup>10,11</sup>, A. Fontana<sup>1</sup>, K. Glazebrook<sup>17</sup>, A. Henry<sup>12,13</sup>, C. Mason<sup>14,15</sup>, E. Merlin<sup>1</sup>, T. Morishita<sup>16</sup>, T. Nanayakkara<sup>17</sup>, D. Paris<sup>1</sup>, N. Roy<sup>13</sup>, H. Williams<sup>18</sup>, X. Wang<sup>19,20,21</sup>, G. Brammer<sup>14,15</sup>, M. Bradač<sup>23,24</sup>, W. Chen<sup>18</sup>, P. L. Kelly<sup>18</sup>, A. M. Koekemoer<sup>12</sup>, M. Trenti<sup>10,11</sup>, and R. A. Windhorst<sup>22</sup>

<sup>1</sup> INAF – Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy  
e-mail: sara.mascia@inaf.it

<sup>2</sup> Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, 00133 Roma, Italy

<sup>3</sup> Department of Physics and Astronomy, University of California, 430 Portola Plaza, Los Angeles, CA 90095, USA

<sup>4</sup> Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, Kashiwa 277-8583, Japan

<sup>5</sup> Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy

<sup>6</sup> INAF – OAS, Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, 40129 Bologna, Italy

<sup>7</sup> INAF – IASF Milano, Via A. Corti 12, 20133 Milano, Italy

<sup>8</sup> Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Saragat 1, 44122 Ferrara, Italy

<sup>9</sup> INAF Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, 35122 Padova, Italy

<sup>10</sup> School of Physics, University of Melbourne, Parkville 3010, VIC, Australia

<sup>11</sup> ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Stromlo, ACT 2611, Australia

<sup>12</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>13</sup> Center for Astrophysical Sciences, Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

<sup>14</sup> Cosmic Dawn Center (DAWN), Denmark

<sup>15</sup> Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 København N, Denmark

<sup>16</sup> IPAC, California Institute of Technology, MC 314-6, 1200 E. California Boulevard, Pasadena, CA 91125, USA

<sup>17</sup> Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia

<sup>18</sup> Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA

<sup>19</sup> School of Astronomy and Space Science, University of Chinese Academy of Sciences (UCAS), Beijing 100049, PR China

<sup>20</sup> National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, PR China

<sup>21</sup> Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, PR China

<sup>22</sup> School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA

<sup>23</sup> Department of Mathematics and Physics, University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia

<sup>24</sup> Department of Physics and Astronomy, University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA

Received 9 January 2023 / Accepted 16 February 2023

## ABSTRACT

The escape fraction of Lyman-continuum (LyC) photons ( $f_{\text{esc}}$ ) is a key parameter for determining the sources of cosmic reionization at  $z \geq 6$ . At these redshifts, owing to the opacity of the intergalactic medium, the LyC emission cannot be measured directly. However, LyC leakers during the epoch of reionization could be identified using indirect indicators that have been extensively tested at low and intermediate redshifts. These include a high [OIII]/[OII] flux ratio, high star-formation surface density, and compact sizes. In this work, we present observations of 29  $4.5 \leq z \leq 8$  gravitationally lensed galaxies in the Abell 2744 cluster field. From a combined analysis of JWST-NIRSpec and NIRCam data, we accurately derived their physical and spectroscopic properties: our galaxies have low masses ( $\log(M_*) \sim 8.5$ ), blue UV spectral slopes ( $\beta \sim -2.1$ ), compact sizes ( $r_e \sim 0.3\text{--}0.5$  kpc), and high [OIII]/[OII] flux ratios. We confirm that these properties are similar to those characterizing low-redshift LyC leakers. Indirectly inferring the fraction of escaping ionizing photons, we find that more than 80% of our galaxies have predicted  $f_{\text{esc}}$  values larger than 0.05, indicating that they would be considered leakers. The average predicted  $f_{\text{esc}}$  value of our sample is 0.12, suggesting that similar galaxies at  $z \geq 6$  have provided a substantial contribution to cosmic reionization.

**Key words.** galaxies: evolution – galaxies: formation – early Universe – galaxies: clusters: intracluster medium

## 1. Introduction

The Lyman-continuum (LyC,  $\lambda < 912 \text{ \AA}$ ) photons escaping from star-forming galaxies into the neutral intergalactic medium (IGM) can account for the photon budget required to complete reionization only if a substantial fraction of them escape from the galaxies' interstellar and circumgalactic media (ISM and CGM). Given the number density of star-forming galaxies in

the Epoch of Reionization (EoR), an average LyC escape fraction ( $f_{\text{esc}}$ ) of  $\sim 10\%$  across all galaxies would be required (e.g., Finkelstein et al. 2019; Robertson et al. 2015) to both reionize the Universe by  $z = 6$  and match the Thomson optical depth of electron scattering observed in the cosmic microwave background (CMB, Planck Collaboration VI 2020).

However, at  $z \geq 4.5$  it is impossible to detect the LyC photons escaping from galaxies, since they are absorbed and scattered by

the IGM along the line of sight (Inoue et al. 2014). Therefore, efforts at low redshift, where LyC can be detected, have been focused on identifying other observable properties that trace physical conditions facilitating the escape of LyC photons. These indirect indicators could then be used in the EoR to identify the cosmic ionizers. Several indirect diagnostics have been proposed (e.g., Yamanaka et al. 2020; Izotov et al. 2018b; Marchi et al. 2018; Verhamme et al. 2017), but they are all characterized by a large scatter. One of the best indicators is the presence of a strong Ly $\alpha$  emission (e.g., Pahl et al. 2021; Gazagnes et al. 2020), often characterized by two emission peaks with a small velocity separation. However, at  $z > 6.5$ , this line is attenuated due to its resonant nature and by the increasingly neutral IGM as we approach the EoR (Pentericci et al. 2018; Mason et al. 2019; Jung et al. 2020; Ouchi et al. 2020; Bolan et al. 2022). Emission from the [MgII] $\lambda\lambda 2796, 2803$  doublet has been proposed as a promising LyC proxy (Chisholm et al. 2020), as the escape of this line is controlled by resonant scattering in the same low column-density gas as the Ly $\alpha$  (see also Xu et al. 2022; Izotov et al. 2022).

More recently, the nebular C IV emission line, requiring comparably high ionization energies to the He II line ( $E > 47.9$  eV and  $> 54.4$  eV respectively), has attracted attention since its presence might be strongly linked to the escape of Lyman continuum photons from galaxies (e.g., Schaerer et al. 2022; Senchyna et al. 2022; Saxena et al. 2022; Mascia et al. 2023), although this line is in general much fainter than Ly $\alpha$ . Another very popular indicator is the emission line ratio [O III] $\lambda\lambda 4959, 5007$ /[O II] $\lambda 3727$  (hereafter, O32), as Nakajima & Ouchi (2014) first found evidence for its correlation with  $f_{\text{esc}}$ : high values of O32 would reflect partially incomplete HII regions, where some LyC photons could escape from (Marchi et al. 2018). Later on, it was found that this correlation is characterized by a substantial scatter, as highlighted by low-redshift studies (e.g., Izotov et al. 2018b; Nakajima et al. 2020; Flury et al. 2022). As a result, a high O32 flux ratio is still a necessary condition for a significant measurement of  $f_{\text{esc}}$ , although it is not sufficient by itself to define a LyC leaker, as viewing angles might play a role as well as variation in metallicity and ionization parameter (e.g., Bassett et al. 2019; Katz et al. 2020). Further additional properties, such as low values of Balmer lines' rest-frame equivalent widths ( $\text{EW}_0$ ), are thus required. As a matter of fact, measuring low values of EW in these lines could indicate lower optical depth in the HII region (e.g., Bergvall et al. 2013). However, many local LyC emitters exhibiting high values of Balmer emission line  $\text{EW}_0$ s are present in the literature (Izotov et al. 2016a,b, 2018b,a). A possible explanation for this discrepancy lies in the fact that Balmer-line  $\text{EW}_0$ s are sensitive not only to HII region size, but also to starburst age and star formation history (Zackrisson et al. 2017; Binggeli et al. 2018; Alavi et al. 2020). It is thus imperative to pair the EW diagnostic with another indicator in order to prevent degeneracy. Finally, another diagnostic for a high  $f_{\text{esc}}$  is the SFR surface density ( $\Sigma_{\text{SFR}}$ ): feedback from star formation can blow bubbles or chimneys into the host galaxy's ISM, linking a high  $\Sigma_{\text{SFR}}$  value to a high  $f_{\text{esc}}$ . This connection is supported by the detection of some compact LyC emitters (LCEs, e.g., Izotov et al. 2018b; Marchi et al. 2018), even though compactness may not be a defining characteristic of all of them (e.g., Marchi et al. 2018).

Recently, Flury et al. (2022) presented the first statistical test of many of the diagnostics just described on a large sample of low-redshift LCEs. Their conclusion is that indicators based on Ly $\alpha$  emission properties (such as peak velocity separation and equivalent width) perform well, and that diagnostics such as the [O III]/[O II] flux ratio and the SFR surface density,  $\Sigma_{\text{SFR}}$ , could also be used to determine if a galaxy is an LCE or not.

With the advent of JWST, we now have the opportunity to constrain LyC diagnostics during the epoch of reionization with the first NIRSpec observations of high-redshift galaxies. In this paper, we make use of NIRSpec spectra obtained in the Abell 2744 cluster region for a sample of galaxies at  $4.5 \leq z \leq 8$  observed as part of the JWST Early Release Science program GLASS (Treu et al. 2022) and the JWST Director Discretionary Time program (JWST-GO-2756; PI: Chen; Roberts-Borsani et al. 2022a). With the wavelength coverage of 0.9–5.3  $\mu\text{m}$  offered by the NIRSpec observations, we can confirm not only the redshifts of tens of photometrically selected candidates using intense emission lines, but we can also measure optical line ratios and rest-frame EWs with high precision. Our unique data set also includes deep JWST/NIRCam images, enabling us to better characterize those cosmic reionizer candidates based on their physical properties.

This paper is organized as follows. We present the data set in Sect. 2. We characterize the selected sample and compare the physical and spectroscopic properties with cosmological models and other samples at lower redshifts in Sect. 3. In Sect. 4, we indirectly infer  $f_{\text{esc}}$  for the high-redshift sample and in Sect. 5, we summarize our key conclusions. Throughout this work, we assume a flat  $\Lambda$ CDM cosmology with  $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $\Omega_{\text{m}} = 0.307$  (Planck Collaboration VI 2020) and the Chabrier (2003) initial mass function. All magnitudes are expressed in the AB system (Oke & Gunn 1983).

## 2. Observations and method

The final target list of the GLASS-JWST program and the way in which targets have been prioritized will be presented in a future paper. However, Morishita et al. (2022) have already described the observations and data reduction strategy. Here, we present a brief summary, highlighting the points that are most relevant for this work and describing the methods used to study the properties of the galaxies in our sample.

### 2.1. JWST/NIRSpec MSA observations and data reduction

Our spectra were acquired through NIRSpec MSA observations in two programs: the GLASS-JWST Early Release Science Program (PID 1324, PI: Treu; Treu et al. 2022) and a JWST DDT program (PID 2756, PI: W. Chen; Roberts-Borsani et al. 2022a), which obtained NIRSpec observations for a subset of targets residing over the central regions of the Frontier Field galaxy cluster Abell 2744.

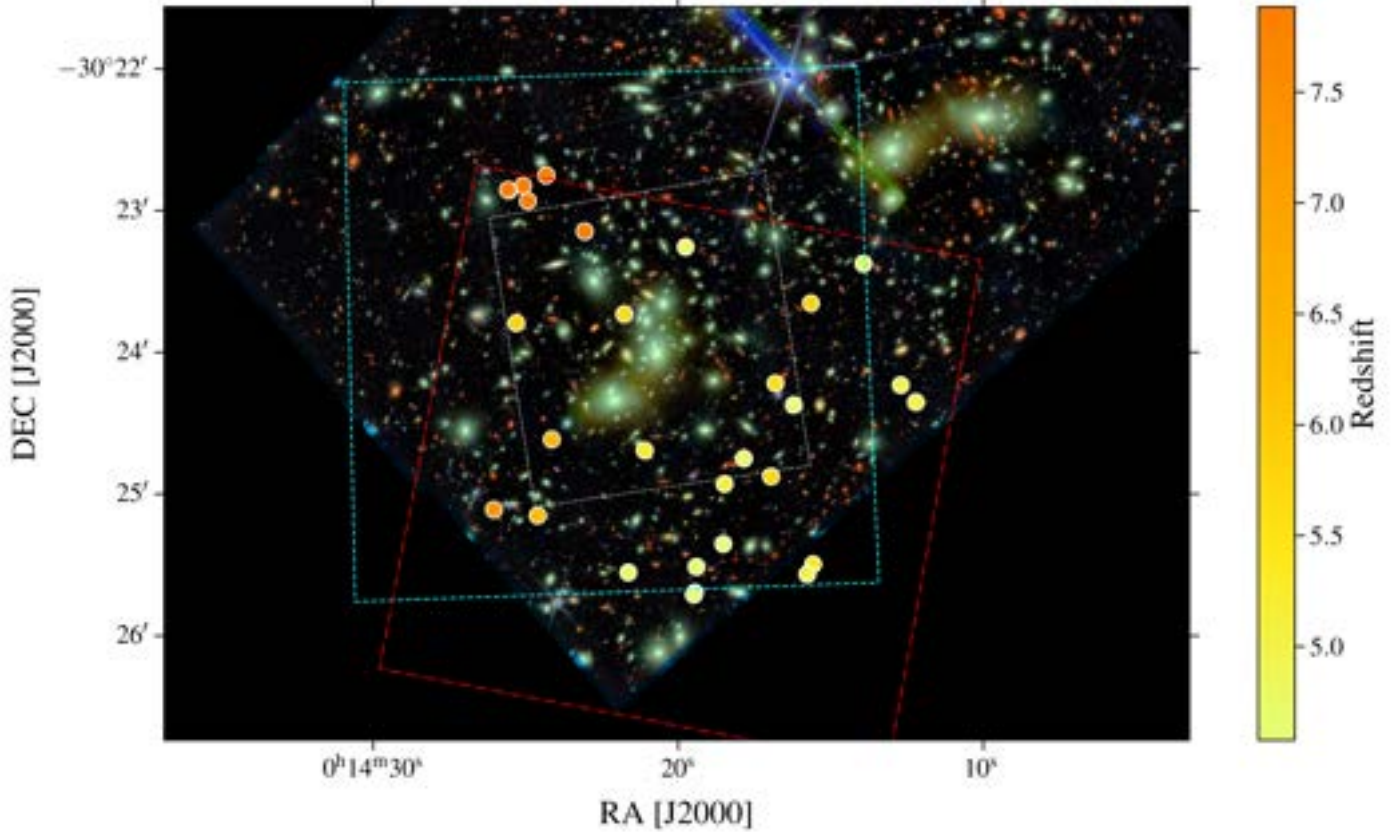
The GLASS-JWST observations were carried out on November 10, 2022, with three spectral configurations (G140H/F100LP, G235H/F170LP, and G395H/F290LP). These configurations cover wavelengths between 1–5.14  $\mu\text{m}$ , at  $R \sim 2000$ –3000. We exposed each of the three high-resolution gratings for a total of 4.9 h. Specifically, in this work, we use the G235H/F170LP and G395H/F290LP observations, which contain the bright emission lines we will analyze.

DDT NIRSpec observations were carried out on October 23 2022, using the CLEAR filter+PRISM configuration, which provides continuous wavelength coverage of 0.6–5.3  $\mu\text{m}$  at  $R \sim 30$ –300 spectral resolution. The on-source exposure time is 1.23 h.

Data were reduced using the official STScI JWST pipeline (ver.1.8.2)<sup>1</sup> for Level 1 data products and the `msaexp`<sup>2</sup> code for Level 2 and 3 data products, which is based on the STScI

<sup>1</sup> <https://github.com/spacetelescope/jwst>

<sup>2</sup> <https://github.com/gbrammer/msaexp>



**Fig. 1.** Spatial location of the 29 selected sources, color coded by their spectroscopic redshift. They are superimposed on the RGB image of the UNCOVER program, made with NIRCcam filters (blue =  $F115W + F150W$ , green =  $F200W + F277W$ , and red =  $F356W + F410M + F444W$ ). The MUSE footprint is shown in white, the NIRSpec GLASS-JWST pointing is shown in cyan, and the NIRSpec DDT pointing is shown in red.

pipeline but also includes additional correction routines. In summary, we initially reduced the uncalibrated data using the `Detector1Pipeline` routine and the latest set of reference files (`jwst_1023.pmap`) to correct for detector-level artifacts and convert them to count-rate images. Then, we applied custom preprocessing routines from `msaexp` to remove residual  $1/f$  noise that is not corrected by the IRS2 readout, to identify and remove “snowballs”, and to remove bias exposure by exposure before running STScI routines from `Spec2Pipeline` for the final 2D cutout images. To perform WCS registration, flat-fielding, path-loss corrections, and flux calibration, these routines include `AssignWcs`, `Extract2dStep`, `FlatFieldStep`, `PathLossStep`, and `PhotomStep`. Of note, our chosen reference files include an in-flight flux calibration, accounting for NIRSpec’s better-than-expected throughput at blue wavelengths. Local background subtraction was performed using a three-shutter nod pattern before the resulting images are drizzled onto a common grid. We optimally extracted the spectra using an inverse-variance weighted kernel, which is derived by summing the 2D spectrum along the dispersion axis and fitting the signal along the spatial axis to a Gaussian profile. We visually inspected all kernels to make sure spurious events are not included. As a result, the kernel extracts the 1D spectrum along the dispersion axis. The final step was to verify the default wavelength calibration for the gratings, which is accurate within  $1 \text{ \AA}$  (Williams et al., in prep.).

## 2.2. Imaging data

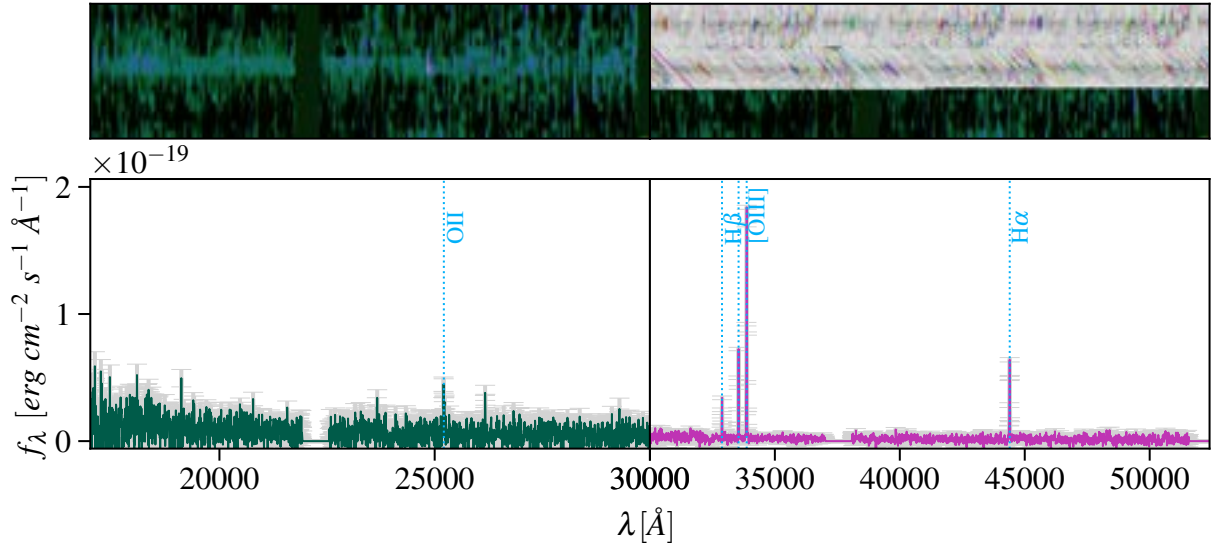
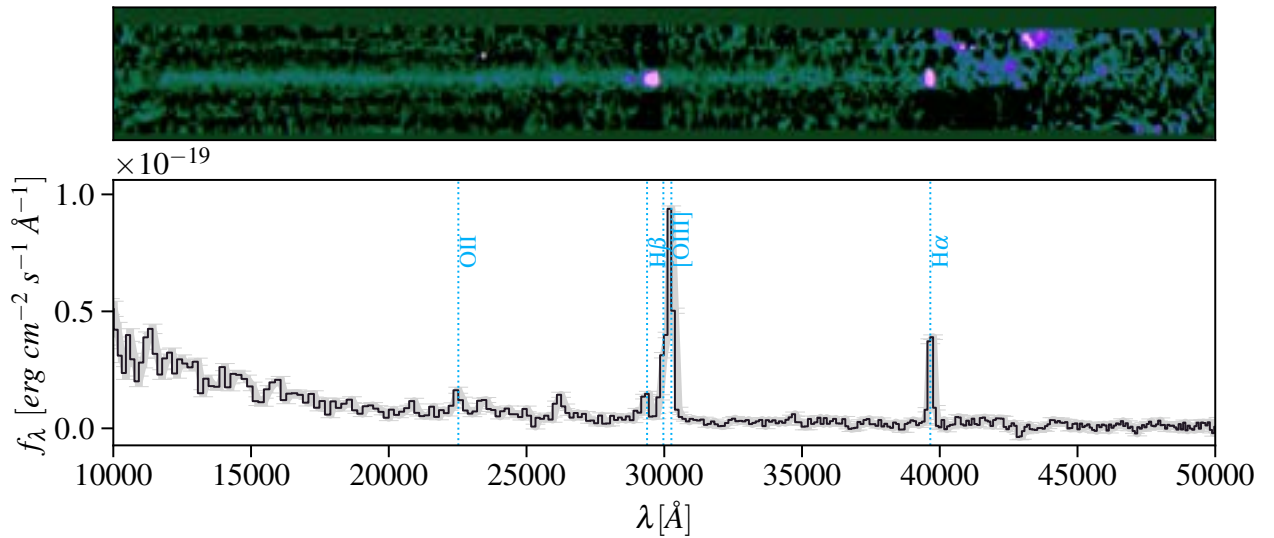
Deep NIRCcam images were acquired from the GO program UNCOVER (GO 2561; PI I. Labbe) and included observa-

tions in the  $F115W$ ,  $F150W$ ,  $F200W$ ,  $F277W$ ,  $F356W$ ,  $F410M$ , and  $F444W$  filters. The imaging data were reduced using the STScI JWST pipeline and the latest versions of photometric zero points and reference files. A detailed description of all the reduction and calibration steps is presented in Merlin et al. (2022). Of the 29 sources analyzed in this work (see next section), 27 were observed within the UNCOVER pointings. Their positions and the UNCOVER footprints are presented in Fig. 1.

## 2.3. Emission line and redshift identifications

The focus of the study is on all sources at  $z \geq 4.5$ . Specifically, we analyzed the spectra of: 13 galaxies with a spectroscopic redshift larger than 4.5 previously confirmed by the MUSE observations (Mahler et al. 2018; Richard et al. 2021), all showing Ly $\alpha$  in emission in their optical spectra (the footprint of the MUSE observations is also shown in Fig. 1); 29 galaxies with a photometric redshift in the range 4.5–8, of which 5 were selected as part of the  $z \approx 7.9$  candidate protocluster and whose confirmation has been recently presented by Morishita et al. (2022). Finally, 23 of the 42 galaxies were observed as part of GLASS-JWST, while 19 were part of the DDT program.

Spectra were visually examined for detectable optical lines using the spectroscopic or photometric redshift information. For photometric sources we determined the spectroscopic redshift when possible using the H $\beta$ , [O III] $\lambda\lambda 4959, 5007$ , and (when present) H $\alpha$  lines. In 29 cases, the [O II] $\lambda 3727$ , [O III] $\lambda\lambda 4959, 5007$  and H $\beta$  were detected and their line fluxes were measured. We also measured the H $\alpha$  line in 17 out of 29 cases, since it falls within the observed spectrum (examples are

ID: 110000,  $z = 5.765$ 

ID: 150015,  $z = 5.041$ 


**Fig. 2.** Example 2D and 1D spectra of two representative galaxies in our sample. Top: 2D (top) and 1D (bottom) NIRSpect GLASS-JWST spectrum of 110 000 at  $z = 5.765$  (green for the G235H/F170LP and purple for the G395H/F290LP configuration) and  $1\sigma$  uncertainty (gray). Bottom: 2D (top) and 1D (bottom) NIRSpect DDT spectrum of 150015 at  $z = 5.041$  (black) and  $1\sigma$  uncertainty (gray). Vertical dashed lines mark the position of well-detected rest-optical emission lines. The continuum emission is also detected as seen in the 2D spectrum. On the  $x$ -axis in the bottom panel, the observed wavelength ( $\text{\AA}$ ) is reported.

shown in Fig. 2). For this part of our analysis, we used the latest version of the `specutils`<sup>3</sup> packages in PYTHON.

From our initial target list of 42 galaxies, there were eight sources with a spectroscopic redshift confirmed from previous MUSE observations, which were placed in the MSA, but for which we are unable to confirm any emission line from JWST data. Of these, five were observed as part of GLASS-JWST and three during the DDT. Most of these sources are extremely faint (fainter than  $28 F_{150W}$ ), thus, their redshift confirmation was solely based on the presence of faint Ly $\alpha$  emission (flux on the order of  $1\text{--}3 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$ ) in the MUSE cubes. We were also unable to detect any emission lines for five galaxies with a

photometric redshift between 4.5 and 7: these galaxies are also faint ( $m_{F_{150W}} \approx 27.2\text{--}28$ ). However, in these cases it is possible that the photometric redshift was incorrect and this is why we are unable to confirm any emission line in their spectra.

In Table 1, we report the GLASS-JWST IDs, the coordinates and the spectroscopic redshifts of all sources together with their  $m_{F_{150W}}$  magnitudes derived from the imaging data when present.

#### 2.4. Dust correction and flux measurements

We measured the total flux of all detected lines (Balmer lines, [OII], and of [OIII]) with a Gaussian fit of each line component. From the flux measurement we subtracted a constant continuum emission, which is estimated as the signal averaged over regions in which there are no lines near the one being measured. When the signal-to-noise ratio (S/N) of [OII] or H $\beta \leq 2$ , we set  $2\sigma$

<sup>3</sup> <https://specutils.readthedocs.io/en/stable/index.html>

**Table 1.** Sample and spectroscopic measurements.

Program	Glass ID	RA [deg]	Dec [deg]	$z_{\text{spec}}$	$m_{F150W}$ [mag]	$EW_0(H\beta)$ [Å]	R23	O32
DDT	10025	3.59609	-30.38581	7.875	26.7	$139 \pm 30$	$8 \pm 3$	$6.6 \pm 1.4$
	40196	3.55803	-30.38962	4.760	25.5	$167 \pm 42$	>11	>90
	50017	3.57006	-30.40369	5.550	26.3	$89 \pm 17$	$8 \pm 3$	$9 \pm 2$
	70002	3.60544	-30.3966	5.771	27.0	>210	>12	>8
	80075	3.56755	-30.40623	4.632	26.9	$173 \pm 38$	$3.4 \pm 1.5$	$9 \pm 2$
	100004	3.60657	-30.38093	7.884	26.9	>130	>5	>5
	110003	3.59069	-30.39554	5.660	25.1	$99 \pm 19$	>4	>8
	150015	3.55085	-30.40590	5.041	25.7	$268 \pm 53$	$4.1 \pm 1.8$	$7.2 \pm 1.9$
	150053	3.58120	-30.42853	4.580	26.7	$212 \pm 51$	>5	>21
	160170	3.56570	-30.42613	4.895	–	$190 \pm 43$	$5 \pm 2$	$8 \pm 2$
	160281	3.59015	-30.42593	4.715	25.6	$231 \pm 52$	$5 \pm 2$	$5.3 \pm 1.4$
	160284	3.58083	-30.42526	4.700	24.8	$130 \pm 27$	>5	>8
	160345	3.55293	-30.40389	5.020	27.2	$110 \pm 44$	>4	>5
	GLASS-JWST	10000	3.60134	-30.37923	7.884	25.5	$76 \pm 13$	$7 \pm 2$
10021		3.60851	-30.41854	7.288	25.4	$104 \pm 24$	$12 \pm 5$	$13 \pm 5$
50002		3.57700	-30.41552	5.135	27.8	$69 \pm 23$	$12 \pm 6$	$19 \pm 7$
50038		3.56520	-30.39426	5.773	25.5	$92 \pm 19$	$10 \pm 4$	$9 \pm 3$
70018		3.58790	-30.41159	5.284	27.1	$155 \pm 32$	>4	>12
80070		3.58232	-30.38765	4.798	26.4	$135 \pm 25$	$12 \pm 5$	$53 \pm 20$
80085		3.57435	-30.41253	4.728	28.1 <sup>(a)</sup>	$168 \pm 24$	$2.5 \pm 1.5$	$19 \pm 11$
100001		3.60385	-30.38223	7.875	26.5	$39 \pm 8$	$10 \pm 4$	$3.2 \pm 1.0$
100003		3.60451	-30.38044	7.880	26.2	$85 \pm 18$	$11 \pm 4$	$21 \pm 7$
100005		3.60646	-30.38099	7.883	26.7	$33 \pm 15$	$21 \pm 12$	$2.9 \pm 1.0$
110000		3.57064	-30.41464	5.765	25.8	$57 \pm 13$	$9 \pm 4$	$6 \pm 2$
150008		3.60253	-30.41923	6.230	25.8	$141 \pm 30$	>7	>20
150029		3.57717	-30.42258	4.585	26.1	$205 \pm 45$	$8 \pm 3$	$17 \pm 6$
160122		3.5649	-30.42496	5.333	–	$149 \pm 25$	$5 \pm 2$	$18 \pm 7$
160275		3.58107	-30.42830	4.578	26.2	$5 \pm 1$	–	–
400009		3.60059	-30.41027	6.376	26.8	$35 \pm 7$	–	–

**Notes.** <sup>(a)</sup>From MUSE catalog, magnitude  $F140W$  HST. Missing O32 and R23 values are due to [OII] undetection because of its position in un-acquired part of the spectrum.  $H\beta$  and [OIII] are always detected.

as an upper limit. Prior to carry out a quantitative analysis, it is necessary to consider corrections for dust reddening. For 22 galaxies,  $H\alpha$  and  $H\beta$  are both available: for these, we calculated the correction for dust extinction on the basis of the Balmer decrement, assuming a Calzetti et al. (2000) extinction law and an intrinsic ratio  $H\alpha/H\beta = 2.86$  (see e.g., Domínguez et al. 2013; Kashino et al. 2013; Price et al. 2014), which is valid for an electric temperature of 10 000 K. For the other 7 sources where  $H\alpha$  is not in the observed range, we applied the average correction  $E(B-V)_{\text{neb}} \sim 0.1$  derived from all other sources.

With the dust corrected values, we calculated the R23 = ([OIII] + [OII])/H $\beta$  and O32 line fluxes ratios and the H $\beta$  rest-frame  $EW_0$ s. We list all these values in in Table 1. It is important to highlight that the O32 values slightly depend on the dust correction. The assumed temperature may be too low for high- $z$  star-forming galaxies (e.g., Curti et al. 2023; Nakajima et al. 2023), however the  $H\alpha/H\beta$  ratio varies by less than 5% for a 20 000 K temperature (Dopita & Sutherland 2003) and therefore the O32 variation are well below the current uncertainties.

### 2.5. Measurements of physical parameters

Following Santini et al. (2023), we measured the stellar masses  $M_{\star, \text{obs}}$ , the observed absolute UV magnitudes at 1500 Å ( $M_{1500, \text{obs}}$ ), the star formation rates (SFRs), and dust reddening

$E(B-V)$  by fitting synthetic stellar templates to the seven-band NIRCcam photometry and the released HST photometry (Castellano et al. 2016) with ZPHOT (Fontana et al. 2000). We adopted Bruzual & Charlot (2003) models and assumed delayed exponentially declining star formation histories – SFH( $t$ )  $\propto (t^2/\tau) \cdot \exp(-t/\tau)$  – with  $\tau$  ranging from 0.1 to 7 Gyr. The age ranges from 10 Myr to the age of the Universe at each galaxy redshift, while metallicity can assume values of 0.02, 0.2 or 1 times Solar metallicity. For the dust extinction, we used the Calzetti et al. (2000) law with  $E(B-V)$  ranging from 0 to 1.1. We computed  $1\sigma$  uncertainties on the physical parameters by retaining, for each object, the minimum and maximum fitted masses among all the solutions with a probability  $P(\chi^2) > 32\%$  of being correct, fixing the redshift to the best-fit value.

As a final step, we needed to adjust the stellar masses and  $M_{1500}$  and the SFR values for the lensing amplification factor. The magnification factor  $\mu$  was derived by combining the LM-model (Bergamini et al. 2023), the model used by Roberts-Borsani et al. (2022b), with a new spatially more extensive model that fully covers the JWST field of view (Bergamini et al., in prep.). The  $\mu$  values range from 1.6 to 12, with a median value of  $\sim 2$ . The results on our sources' stellar masses,  $M_{\star, \text{true}}$ , their  $M_{1500}$ , and their lensing magnification factors  $\mu$  are reported in Table 2.

**Table 2.** Physical parameters derived from multi-band photometry.

Program	Glass ID	$\mu^{(*)}$	$\log_{10}(M_{\star})$ [ $M_{\odot}$ ]	$r_e$ [kpc]	$\log_{10}(\Sigma_{\text{SFR}})$ [ $M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ ]	$M_{1500}$ [mag]	$\beta$	
DDT	10025	$2.60^{+0.08}_{-0.07}$	$8.03^{+0.19}_{-0.14}$	0.40	$0.4 \pm 0.2$	$-19.00^{+0.01}_{-0.03}$	$-2.08 \pm 0.45$	
	40196	$2.95^{+0.21}_{-0.72}$	$7.87^{+0.13}_{-0.13}$	0.33	$1.6 \pm 1.0$	$-19.41^{+0.02}_{-0.03}$	$-2.03 \pm 0.48$	
	50017	$2.27^{+0.07}_{-0.16}$	$8.78^{+0.23}_{-0.69}$	0.43	$0.7 \pm 0.1$	$-19.13^{+0.10}_{-0.19}$	$-1.77 \pm 0.25$	
	70002	$2.19^{+0.05}_{-0.05}$	$7.13^{+0.01}_{-0.15}$	0.12	$2.0 \pm 1.3$	$-18.44^{+0.01}_{-0.01}$	$-2.54 \pm 0.26$	
	80075	$2.02^{+0.06}_{-0.17}$	$6.76^{+0.23}_{-0.15}$	0.07	$2.3 \pm 1.7$	$-17.37^{+0.01}_{-0.03}$	$-2.26 \pm 0.23$	
	100004	$1.92^{+0.06}_{-0.08}$	$9.17^{+0.14}_{-0.28}$	0.40	$0.2 \pm 0.3$	$-19.76^{+0.11}_{-0.10}$	$-1.88 \pm 0.44$	
	110003	$11.6^{+0.9}_{-1.0}$	$7.92^{+0.71}_{-0.06}$	0.48	$1.2 \pm 0.5$	$-20.52^{+0.15}_{-0.03}$	$-2.51 \pm 0.24$	
	150015	$1.79^{+0.05}_{-0.29}$	$7.87^{+0.34}_{-0.44}$	0.38	$1.2 \pm 0.7$	$-19.52^{+0.09}_{-0.15}$	$-2.54 \pm 0.49$	
	150053	$1.75^{+0.03}_{-0.14}$	$7.85^{+0.48}_{-0.45}$	0.80	$-0.1 \pm 0.6$	$-19.26^{+0.08}_{-0.15}$	$-1.25 \pm 0.21$	
	160170	$1.58^{+0.04}_{-0.13}$	–	–	–	–	–	
	160281	$1.96^{+0.03}_{-0.15}$	$7.93^{+0.24}_{-0.08}$	1.65	$-0.3 \pm 0.9$	$-20.00^{+0.01}_{-0.11}$	$-2.26 \pm 0.51$	
	160284	$1.85^{+0.03}_{-0.15}$	$8.85^{+0.01}_{-0.10}$	0.53	$1.2 \pm 0.6$	$-20.16^{+0.02}_{-0.01}$	$-2.05 \pm 0.48$	
	160345	$1.86^{+0.05}_{-0.30}$	$7.95^{+0.09}_{-0.81}$	0.45	$0.5 \pm 0.1$	$-18.67^{+0.06}_{-0.12}$	$-2.49 \pm 0.49$	
	GLASS-JWST	10000	$2.13^{+0.07}_{-0.11}$	$8.11^{+0.27}_{-0.01}$	0.20	$2.0 \pm 1.3$	$-19.71^{+0.10}_{-0.02}$	$-2.27 \pm 0.46$
		10021	$2.17^{+0.08}_{-0.03}$	$8.70^{+0.11}_{-0.35}$	0.68	$0.8 \pm 0.2$	$-21.18^{+0.08}_{-0.02}$	$-2.25 \pm 0.48$
50002		$2.15^{+0.06}_{-0.13}$	$8.57^{+0.22}_{-0.11}$	0.70	$-1.4 \pm 2.0$	$-18.78^{+0.06}_{-0.11}$	$-1.66 \pm 0.66$	
50038		$2.66^{+0.12}_{-0.32}$	$9.23^{+0.03}_{-0.39}$	0.48	$-1.7 \pm 2.0$	$-19.76^{+0.03}_{-0.09}$	$-1.79 \pm 0.25$	
70018		$6.89^{+0.21}_{-0.30}$	$8.05^{+0.01}_{-0.14}$	0.19	$1.9 \pm 1.3$	$-18.00^{+0.04}_{-0.04}$	$-2.02 \pm 0.49$	
80070		$5.45^{+0.26}_{-0.27}$	$7.42^{+0.21}_{-0.06}$	0.59	$0.4 \pm 0.2$	$-18.24^{+0.01}_{-0.03}$	$-1.92 \pm 0.49$	
80085		$2.11^{+0.05}_{-0.11}$	–	–	–	–	$-2.43^{(a)}$	
100001		$1.99^{+0.06}_{-0.09}$	$9.48^{+0.04}_{-0.13}$	0.50	$0.34 \pm 0.3$	$-20.45^{+0.04}_{-0.02}$	$-1.63 \pm 0.48$	
100003		$1.96^{+0.06}_{-0.09}$	$8.29^{+0.31}_{-0.15}$	0.15	$2.3 \pm 1.7$	$-20.15^{+0.15}_{-0.02}$	$-2.51 \pm 0.48$	
100005		$1.92^{+0.06}_{-0.08}$	$8.63^{+0.17}_{-0.27}$	0.25	$1.5 \pm 1.2$	$-19.85^{+0.11}_{-0.07}$	$-2.55 \pm 0.48$	
110000		$1.90^{+0.05}_{-0.14}$	$9.10^{+0.05}_{-0.62}$	0.54	$0.5 \pm 0.2$	$-20.13^{+0.04}_{-0.17}$	$-1.70 \pm 0.23$	
150008		$2.60^{+0.03}_{-0.04}$	$8.41^{+0.35}_{-0.19}$	0.40	$1.5 \pm 0.7$	$-19.33^{+0.14}_{-0.06}$	$-2.10 \pm 0.25$	
150029		$1.86^{+0.04}_{-0.14}$	$8.55^{+0.07}_{-0.01}$	0.27	$1.7 \pm 1.0$	$-18.75^{+0.01}_{-0.04}$	$-1.85 \pm 0.20$	
160122		$1.59^{+0.04}_{-0.13}$	–	–	–	–	–	
160275		$1.75^{+0.03}_{-0.14}$	$7.73^{+0.11}_{-0.14}$	0.47	$-1.3 \pm 1.8$	$-19.20^{+0.01}_{-0.02}$	$-2.55 \pm 0.20$	
400009	$9.3^{+1.6}_{-0.6}$	$6.77^{+0.54}_{-0.26}$	0.11	$1.0 \pm 0.3$	$-17.05^{+0.14}_{-0.04}$	$-2.17 \pm 0.25$		

**Notes.** All measurements are corrected for magnification where needed. <sup>(\*)</sup>Median magnification of the lens model by Bergamini (in prep.), calculated at the position of the source. Measurements are associated with  $1\sigma$  uncertainties. <sup>(a)</sup>Measured from HST photometry.

## 2.6. Sizes and SFR surface density estimates

The SFRs were derived from the  $H\alpha$  emission line using the standard conversion by Kennicutt (1998), that is,  $\text{SFR}(H\alpha) [M_{\odot} \text{ yr}^{-1}] = 7.9 \times 10^{42} L_{H\alpha}$ , and then corrected for the Chabrier IMF. For those few sources at  $z \geq 6.6$ , where the  $H\alpha$  line falls outside the observable window we used dust-corrected  $H\beta$  fluxes instead. To correct for slit losses and possible residual uncertainties on flux calibration, we normalized the spectra to the  $F444W$  filter by integrating the spectrum under the  $F444W$  bandpass, the closest to the  $H\alpha$  line in our sample. Thanks to the high resolution of the GLASS-JWST spectra, no correction for NII contamination is required for  $H\alpha$  measurements. For galaxies observed by the DDT programs,  $H\alpha$  is blended with the NII doublet. However our sources are all very-low-mass galaxies and, based on low-redshift results, we expect contamination to be less than 10% (e.g., Faisst et al. 2018). Therefore, we did not attempt to make any correction on the fluxes.

The results obtained on the SFR were also compared to the values determined by the SED fitting procedure described in Sect. 2.5, finding a reasonable agreement between the two data

sets, indicating that there are no systematic issues. Obviously, the SED fitting SFR is heavily dependent on the choice of SFH, and also on the dust correction, while the value derived from the  $H\alpha$  luminosity is sensitive to short timescales. That is to say that it only probes the presence of short-lived, massive stars, so their ratio is actually an indication of the burstiness of the galaxy.

We measured the size of each galaxy  $r_e$  in the  $F115W$  band, corresponding to the UV rest-frame of the galaxies. Adopting forward-modeling technique, we assumed that galaxies are well represented by a Sersic profile (Sersic 1968) as Yang et al. (2022). Then we modeled the appearance of the source in the image plane considering lensing and PSF effects. The source reconstruction was performed via python software Lenstruction (see details in Yang et al. 2020), which is built on Lenstronomy (Birrer et al. 2021). In this way we obtained the intrinsic properties of galaxies in the source plane, hence, the intrinsic size. Finally, we calculated the SFR surface density using the relation  $\Sigma_{\text{SFR}} = \text{SFR}/2\pi r_e^2$  (e.g., Naidu et al. 2020; Flury et al. 2022). The SFR surface densities  $\Sigma_{\text{SFR}}$  are listed in Table 2.

## 2.7. UV- $\beta$ slopes

We measured the UV slope of our galaxies from the NIR-Cam photometry and/or the previously available HST photometry (Castellano et al. 2016), with the approach detailed in Calabrò et al. (2021). In brief, we considered all the photometric bands whose entire bandwidths are between a 1216 and 3000 Å rest frame. The former limit is set to exclude the Ly $\alpha$  line and Ly-break, while the latter limit is slightly larger than that adopted in Calabrò et al. (2021) to ensure a larger range.

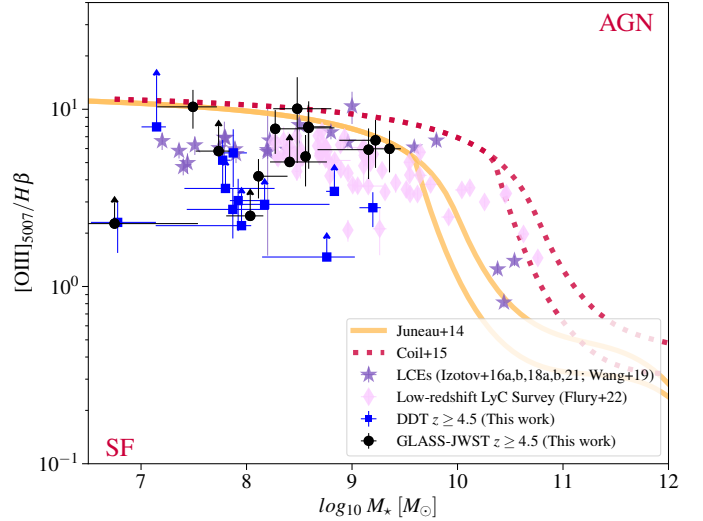
Then, we fitted the selected photometry with a single power-law of the form  $f(\lambda) \propto \lambda^\beta$ . In practice, we fitted two or three photometric bands amongst HST *F814W* and JWST-NIRCam *F115W*, *F150W* or *F200W* depending on the exact redshift of the sources. This choice allows us to uniformly probe the spectral range between 1500 and 3000 Å for most of the galaxies. We measured the  $\beta$  and associated uncertainty for each source using a bootstrap method. By using  $n = 500$  Monte Carlo simulations, the fluxes in each band were varied according to their error. The results provided a resultant slope distribution from which we calculated the mean and standard deviation of  $\beta$  for each galaxy. The results on  $\beta$  with associated errors are reported in Table 2. We find a median  $\beta$  of  $-2.1$ , with a  $1\sigma$  dispersion of 0.4. The UV magnitudes  $M_{1600}$  that can be derived simultaneously with this approach are consistent with the values obtained from the SED fitting and described in the previous section.

## 3. Results

### 3.1. Possible AGN contamination

Our sample was selected solely based on known spectroscopic redshift (via faint and narrow Ly $\alpha$  emission) or through photometric redshifts. It could therefore be affected by AGN contamination. Since we are interested in searching for candidate LyC emitters amongst the star-forming population, we first consider whether the primary source of ionization might not be star formation, but AGN activity instead.

We employed the mass-excitation (MEx) diagram, first proposed by Juneau et al. (2011) to combine the measurements of the [OIII]  $\lambda 5007/H\beta$  emission line ratio with the stellar mass to discriminate between AGNs and star-forming galaxies. This diagram was proposed as an alternative to the classical BPT diagram that compares the [OIII]  $\lambda 5007/H\beta$  to the NII/ $H\alpha$  emission line ratios (Baldwin et al. 1981) when these last two lines fall out of the visibility window and it is not possible to use them to characterize the ionization mechanism in the galaxies – as is the case for our low-resolution DDT spectra where the NII doublet is blended with  $H\alpha$ . Due to the high resolution of the GLASS-JWST spectra, for the galaxies with these observations, we used the dust-corrected flux measurement of the 5007 Å component of the [OIII] doublet. Instead, for the galaxies observed as part of the DDT program for which the doublet cannot be resolved, we determined the [OIII]  $\lambda 5007$  flux, assuming the expected line ratio of 1:3 for the two components fixed by atomic physics. As shown in Fig. 3, the position of our sources in the MEx diagram indicates that our sample contains essentially star-forming galaxies, lying below or around the division line identified by Coil et al. (2015) for  $z = 2.3$  galaxies and AGN from the MOSDEF survey. For reference, we also plot the galaxies at  $z = 0.3-0.4$  from the Low-redshift Lyman Continuum Survey from Flury et al. (2022) and the compilation of low- $z$  LCE also used by the same authors (from Izotov et al. 2016a,b, 2018b,a, 2021; Wang et al. 2019).

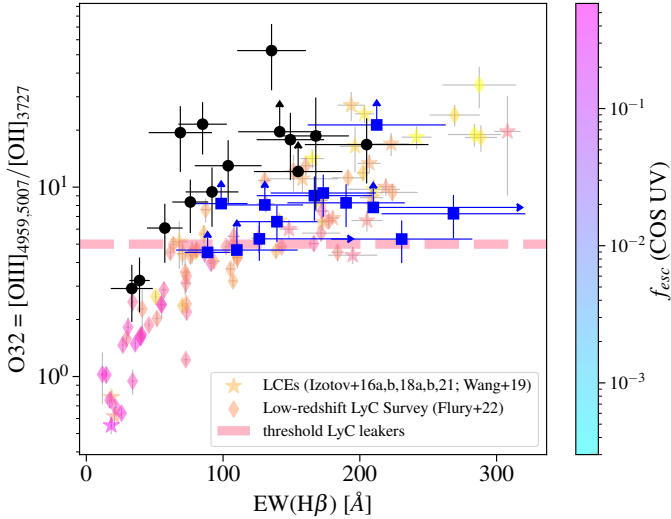


**Fig. 3.** MEx diagram for the sample of galaxies analyzed in this paper: black dots and green squares are for the GLASS-JWST and DDT samples, respectively. For reference, we plot also the galaxies at  $z = 0.3-0.4$  from Flury et al. (2022) (diamonds) and the LCE candidates from previous studies (stars, Izotov et al. 2016a,b, 2018b,a, 2021; Wang et al. 2019). The two orange demarcation lines from Juneau et al. (2014) show the boundaries of the AGN-and-star-forming transition region. All objects above the upper line are AGN-dominated; all galaxies below or rightward of the lower line are presumptively dominated by star formation. We also show the separation from Coil et al. (2015) (dotted lines), which is the adaptation of the Juneau et al. model for galaxies and AGNs at  $z \sim 2.3$  from the MOSDEF survey.

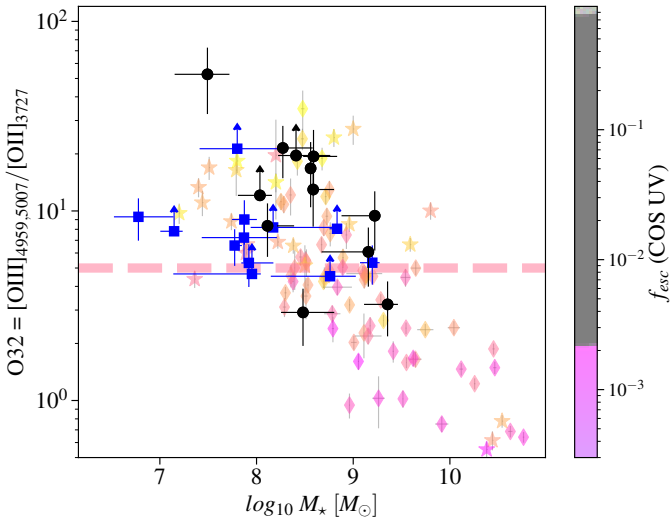
### 3.2. LyC indirect diagnostics

As discussed above, Ly $\alpha$  is possibly the best indirect diagnostic of LyC escape, since the conditions that favour the escape of Ly $\alpha$  photons are often the same that allow for the escape of LyC photons. However, we cannot use this parameter for our sample, since the NIRSpect data do not cover the 1216 Å region for most of our galaxies and only a small subset of our sources have been covered by previous MUSE observations (see also Fig. 1). In addition, for the sources at  $z \geq 7$ , the IGM becomes significantly neutral, thus absorbing the emission even in sources where the line would be intrinsically bright (e.g., Stark et al. 2010; Pentericci et al. 2011; Mason et al. 2018). Indeed, Morishita et al. (2022) recently showed that none of the galaxies in the protocluster candidate at  $z = 7.89$  presents bright Ly $\alpha$  emission and estimate an average neutral hydrogen fraction of the IGM in the region to be  $>0.45$ . Therefore in this work we concentrate on the other most promising diagnostics tested at low redshift by Flury et al. (2022) and available for our galaxies: these authors showed that O32,  $\beta_{1200}$ ,  $r_{50}$ , and  $\Sigma_{\text{SFR}}$  as well as  $\text{EW}_0(H\beta)$  and  $M_{1500}$  exhibit some of the strongest and most significant correlations with  $f_{\text{esc}}$ . This would indicate that certain characteristics, such as concentrated star formation, young stellar populations, and high ionization states, play an essential role in the escape of LyC photons.

We began by analyzing the O32 and the rest-frame  $\text{EW}_0(H\beta)$  relation, as shown in Fig. 4. We plot the values for the JWST high-redshift sample and compare them to the local galaxies with measured  $f_{\text{esc}}$  from previous works (Flury et al. 2022; Izotov et al. 2016a,b, 2018b,a, 2021; Malkan & Malkan 2021; Wang et al. 2019). These values are color-coded as a function of their  $f_{\text{esc}}$  (COS UV) measured values. We can see that the large



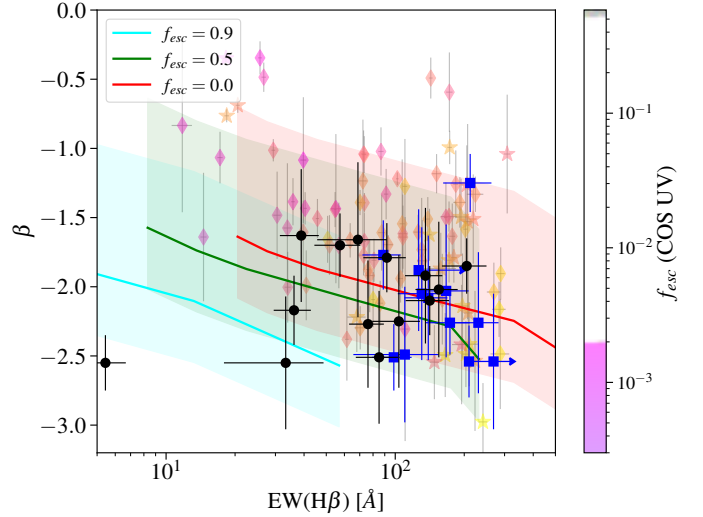
**Fig. 4.** O32 vs. rest-frame  $EW_0(H\beta)$ . Symbols are the same as in Fig. 3. For reference, we also plot the galaxies at  $z = 0.3$ – $0.4$  from Flury et al. (2022); diamonds) and the LCE from Izotov et al. (2016a,b, 2018b,a, 2021), Wang et al. (2019; stars). Symbols are color coded as a function of their measured  $f_{\text{esc}}$ . The pink line ( $O32 = 5$ ) indicates the threshold for LCEs as predicted by Flury et al. (2022).



**Fig. 5.** O32 vs.  $M_*$ . Symbols are as in Fig. 4.

majority of the high-redshift sources have values of O32 larger than 5, which has been indicated as a threshold for LyC leakers (LCEs from here onwards) by Flury et al. (2022). We note that Flury et al. (2022) define leakers as galaxies with an  $f_{\text{esc}} > 0.05$  measured with  $S/N \geq 5$ . Our sample indeed mostly lies in the region of the plot populated by low-redshift leakers. In Fig. 5, we show O32 as a function of total stellar mass, again comparing our sample to the low- $z$  galaxies: our sample perfectly overlaps with the low- $z$  galaxies at the low-mass end, where we find most of the LCEs.

Zackrisson et al. (2013) suggested that galaxies ought to be identified with high  $f_{\text{esc}}$  by combining the UV slope  $\beta$  with the measurement of the EW of a Balmer line, such as  $H\beta$ , and that the predictions are almost independent of the model assumed (i.e., radiation or density bounded nebula). We know that both  $\beta$  and  $EW_0(H\beta)$  are also good indirect indicators according to Flury et al. (2022), although there does not seem to be a direct

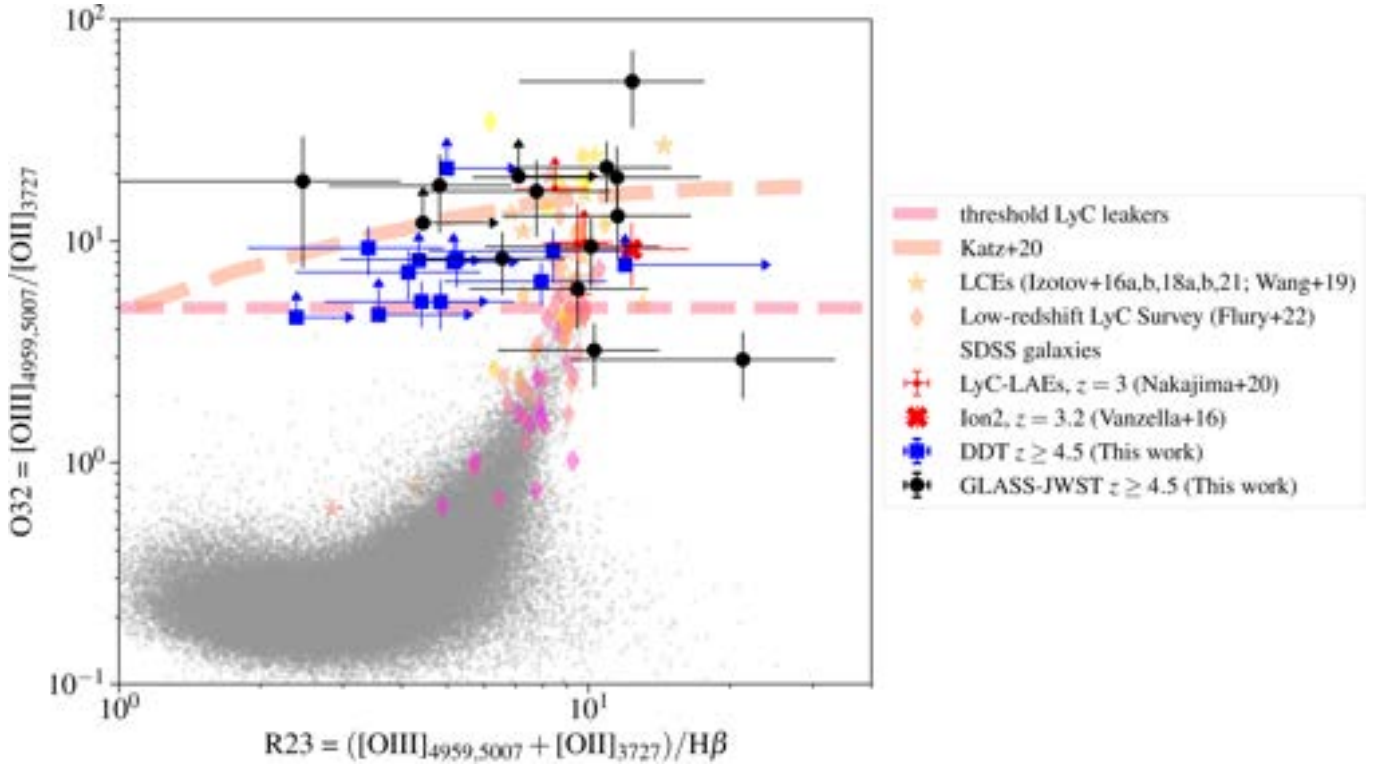


**Fig. 6.**  $\beta$  vs. rest-frame  $EW_0(H\beta)$ . Models are from Zackrisson et al. (2013) and simulate the expected trend for galaxies with an exponential declining SFR ( $Z = 0.02$ , solid lines) and various values of escape fractions. Symbols are as in Fig. 4.

correlation between the two values: Flury et al. (2022) showed that their LCEs do not seem to follow the predictions provided by the models from Zackrisson et al. (2013).

We show the properties of our sample in Fig. 6. Specifically, at a given value of  $EW_0(H\beta)$ , the high-redshift galaxies are, on average, bluer than the majority of the low- $z$  galaxies (consistent with the lower stellar mass probed by our sources); but their slopes are similar to the subset of the low- $z$  sources with moderate-to-high values for  $f_{\text{esc}}$ . The  $\beta$  slopes for our galaxies are perfectly consistent with the average values found for LBGs at  $z \sim 6$  for galaxies with  $M_{\text{UV}}$  in the same range Bouwens et al. (2014). In the figure, we also show the models from Zackrisson et al. (2013) for various values of escape fraction, after correcting the intrinsic  $\beta$  slopes of the models for dust attenuation and assuming the average reddening of our sample derived from the SED fitting,  $E(B - V) = 0.097$  (solid line), and also the maximum and minimum  $E(B - V)$  in our sample (shaded area). We note that our galaxies show more consistency with model predictions that assume moderate-to-low escape fractions  $0.0 \leq f_{\text{esc}} \leq 0.5$ .

Finally, we analyzed the sizes and star formation rate density for our galaxies. Several authors have postulated that concentrated star formation provides the feedback necessary to clear paths in the ISM that allow for the escape of LyC photons. In addition, Marchi et al. (2018) recently found that stacks of galaxies that are UV-compact ( $r_{\text{UV}} \leq 0.30$  kpc) have much higher LyC flux than the average population (see also Izotov et al. 2018b). We find that, with the exception of one galaxy (ID 160281, which has  $r_e = 1.65$  kpc), our targets are all extremely compact with typical  $r_e$  in the rest-frame UV around 0.2–0.6 kpc. These values are similar or even lower than those found for the low- $z$  galaxies and LCEs (Flury et al. 2022); thus, once again the high-redshift sources have equal properties to the low- $z$  LCEs. Since we know that sizes evolve with redshift and galaxies become progressively more compact, we also compared our sample to the general population of star-forming galaxies (LBGs selected) at  $z \approx 6$  analyzed by Shibuya et al. (2015). Specifically, for  $M_{\text{UV}} = -18$  ( $-20$ ), the average UV rest-frame sizes are  $r_e \approx 0.38$  (0.56) kpc, respectively. Thus, indeed our



**Fig. 7.** O32 vs. R23 diagnostic diagram for our sample. Black dots and green squares show the GLASS-JWST and DDT samples, respectively. For reference, we plot also the SDSS galaxies from [Thomas et al. \(2013\)](#) as dark grey points. The LCEs from [Nakajima et al. \(2020\)](#) are shown red dots. The Ion2 at  $z = 3.2$  comes from [Vanzella et al. \(2016\)](#), [de Barros et al. \(2016\)](#) as a red  $x$  sign, the low- $z$  LyC leakers from [Izotov et al. \(2016a,b, 2018a,b\)](#), [Wang et al. \(2019\)](#) as stars, and the LCEs at  $z = 0.3$ – $0.4$  from the Low-redshift Lyman Continuum Survey (as pink diamonds, [Flury et al. 2022](#)). We show also the locus of high-redshift LCEs predicted from cosmological simulations by [Katz et al. \(2020\)](#) as an orange dashed line and the threshold of  $O32 > 5$  determined in [Flury et al. \(2022\)](#) as a pink dashed line.

galaxies are, from this point of view, as compact as the general UV faint population at the same redshift.

As for the star formation rate densities, the values of  $\log_{10}(\Sigma_{\text{SFR}})$  for our galaxies span a very wide range, with an average  $\Sigma_{\text{SFR}}$  that is slightly higher than the average expected at their redshift, as measured by [Shibuya et al. \(2015\)](#). We note that their values are inferred from the UV and then dust corrected. LyC leakers at low and intermediate redshift, tend to show high  $\Sigma_{\text{SFR}}$  values, as discussed extensively by [Naidu et al. \(2020\)](#), who actually proposed a physically motivated model in which  $f_{\text{esc}}$  would scale as  $\Sigma_{\text{SFR}}^{0.4}$ . [Flury et al. \(2022\)](#) identify a threshold value of  $\Sigma_{\text{SFR}} = 10 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$  above which their LCE fraction changes from 10% to 60%, and where indeed we find half of our high-redshift galaxies.

### 3.3. O32–R23 diagnostics

The O32 vs R23 index diagram (Fig. 7) is widely used to examine the gas-phase metallicity and ionization state both in the local universe (e.g., [Thomas et al. 2013](#); [Izotov et al. 2016a,b, 2018b,a](#)) and at high redshift (e.g., [Flury et al. 2022](#); [Nakajima et al. 2020](#); [Reddy et al. 2022](#); [Vanzella et al. 2019](#)), as both these indices are sensitive to combination of these quantities. [Nakajima et al. \(2020\)](#) showed that  $z \sim 3$  LyC leakers tend to populate the upper right part of this diagram. Recently [Katz et al. \(2020\)](#) used high-resolution cosmological radiation hydrodynamics simulations to examine the properties of LyC leakers deep into the epoch of reionization and found that simulated high-redshift galaxies populate the same

regions of the R23–O32 plane, as the  $z \sim 3$  LyC leakers presented in [Nakajima et al. \(2020\)](#) that tend to have low metallicity. Although they conclude that this plane is not the most useful to differentiate between leakers and non-leakers we note that the  $z = 3$  leakers by [Nakajima et al. \(2020\)](#), with measured values and Ion2 occupy the same region as the  $z = 0.3$  low-redshift leakers. Our galaxies occupy a much broader region, which could reflect either a wider range of metallicity and ionization states or simply the fact that we have large measured uncertainties on the diagnostics.

## 4. Predicting escape fractions of EoR galaxies

In the sections above, we have shown that our high-redshift galaxies have properties that are largely overlapping with those of low- $z$  LCEs. The next step is to try and give an indirect estimate of the  $f_{\text{esc}}$  values for our galaxies, so as to understand if typical low-mass galaxies at  $z \approx 5$ – $7$  could be really the drivers of reionization. We assume that the mechanisms that drive the escape of LyC photons are the same at all redshifts and depend only on the physical properties of the sources.

We proceeded as follows: we used the spectroscopic and physical properties of the 66 galaxies that are part of the [Flury et al. \(2022\)](#) sample, with the additional 22 LCEs from previous studies ([Izotov et al. 2016a,b, 2018b,a, 2021](#); [Wang et al. 2019](#)) to calibrate an empirical relation with the measured  $f_{\text{esc}}$  values. Specifically, we focused on the following properties: O32,  $\text{EW}_0(\text{H}\beta)$ ,  $\beta$ ,  $r_e$  (in kpc),  $\Sigma_{\text{SFR}}$ , and  $M_{1500}$ . For all 88 galaxies, these parameters are accurately measured and, of course, accurate measurements of  $f_{\text{esc}}$  are available. To this end,

we specifically used the  $f_{\text{esc}}$  derived by the COS UV spectral fits (see definition in Flury et al. 2022). Recently Chisholm et al. (2022) followed a somewhat similar approach, also using the same set of low-redshift observations, but limited to the UV- $\beta$  slope and an indirect proxy. They provided a scaling relation between  $\beta$  and  $f_{\text{esc}}$ , although the relation has appreciable scatter that scales with  $f_{\text{esc}}$ .

We first ran the Spearman rank correlation and found that O32,  $r_e$ ,  $\text{EW}_0(\text{H}\beta)$ , and  $\beta$  (in that order) are the properties that are best correlated with  $f_{\text{esc}}$ . We then identified one useful fitting linear relation to obtain an estimate of  $\log_{10}(f_{\text{esc}})$  on the basis of the other four measured physical properties. A fully data-driven regression analysis was carried out by performing a regularized minimization of the root-mean-square error (MSE), computed between the values provided by the equation and the dataset, for several possible combinations of the above properties. We tested both a scheme in which the error on the escape fraction,  $f_{\text{esc}}$ , is not considered and a scheme in which values are weighted by the error. Since O32 and  $\text{EW}_0(\text{H}\beta)$  exhibit a very tight correlation (Spearman correlation between them  $>0.9$ , see also Fig. 4), the information they provide is redundant and therefore we decided to use only O32. We checked that the remaining three parameters are reasonably independent (Spearman correlation  $<0.5$ ) and therefore provide complementary information. We finally found a best-fit relation of the form:

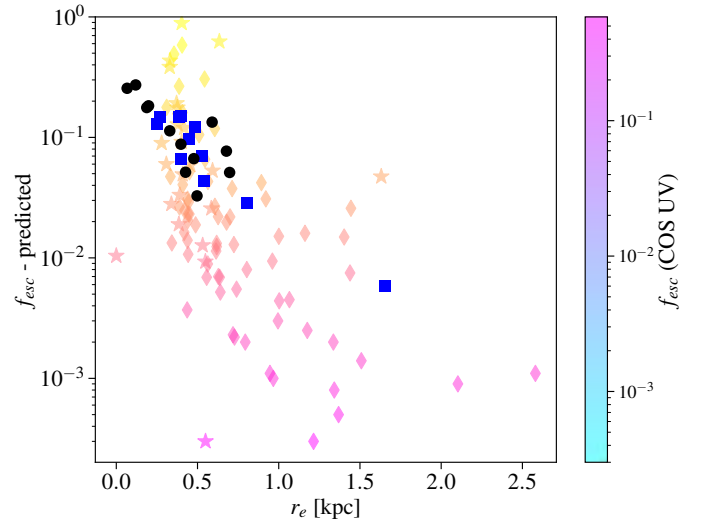
$$\log_{10}(f_{\text{esc}}) = A + B \log_{10}(\text{O32}) + C r_e + D \beta. \quad (1)$$

After identifying the best type of equation, we repeated the minimization process 100 times, following a bootstrap approach every time a random number of sources (between 1 and 25, to avoid being left with too few sources) is randomly removed from the sample. In this way, we could constrain the confidence interval for the above best fit parameters. We find  $A = -1.92[-2.51, -1.71]$ ,  $B = 0.48[0.38, 0.69]$ ,  $C = -0.96[-1.20, -0.62]$ ,  $D = -0.41[-0.58, -0.31]$ , where the values between the parentheses are in the 95th percentile distribution.

Testing the relation on the residuals, we find that it tends to slightly overestimate the  $f_{\text{esc}}$  at very low values (much lower than 0.01), while it tends to underestimate the  $f_{\text{esc}}$  at values that are higher than  $\sim 0.1-0.2$ . This is due to the fact that in the sample used to fit the relation, there are very few galaxies with high  $f_{\text{esc}}$  values and, in general, their measured  $f_{\text{esc}}$  values have higher errors.

We finally applied the above relation to our sample of high-redshift galaxies: out of 29 sources, 3 galaxies do not have the  $r_e$  measurement since they are outside the UNCOVER footprint; for another 2, we do not have an estimate for the O32 parameter. We therefore could apply the best-fitting relation only to 24 sources. In Fig. 8, we show the predicted  $f_{\text{esc}}$  for our sources as a function of  $r_e$ , as well as the low-redshift comparison sample (for which the  $f_{\text{esc}}$  are measured values). Most of our galaxies have predicted  $f_{\text{esc}}$  values larger than 0.05, meaning that they would be considered leakers. The average  $f_{\text{esc}}$  is  $\approx 0.12$  with the bluest and most compact sources having  $f_{\text{esc}}$  as large as 0.2–0.4, which could be lower limits for the reasons discussed above.

Clearly, the main limitations of the above analysis are the fact that the low- $z$  sample used to calibrate the relation is small and, most importantly, it is not evenly populated as it contains mostly objects with rather low  $f_{\text{esc}}$ . Also, it would be important to test these predictions at intermediate redshift ( $z \sim 3-4$ ), that is, much closer to the EoR, where it is possible to both directly detect LyC emission and determine all other physical and spectroscopic properties. However, at the moment, measurements of



**Fig. 8.** Predicted  $f_{\text{esc}}$  vs. size  $r_e$  for our sample (blue and black symbols) and measured  $f_{\text{esc}}$  vs. size  $r_e$  from the literature. Symbols are as in Fig. 4.

$f_{\text{esc}}$  at intermediate redshift are still sparse and most samples lack near-infrared (NIR) spectroscopic follow-up results. In spite of these limitations, a consistent picture emerges from our results, suggesting low-mass galaxies at  $z \sim 5-7$  have mostly properties which indicate moderate values of  $f_{\text{esc}}$ . Interestingly, the average  $f_{\text{esc}}$  value inferred for our sample is equal to the one predicted by Naidu et al. (2020) for galaxies in the mass range  $\log_{10} M_{\star} = 8-9$  at  $z \sim 6$ , according to their simple model, where  $f_{\text{esc}}$  scales with  $\Sigma_{\text{SFR}}$ , which was constrained to reproduce the average observed values for the Steidel et al. (2018) sample at  $z = 3$ .

## 5. Summary and conclusions

Thanks to the magnification power of the Abell 2744 cluster, in this paper, we present the first JWST/NIRSpec observations of 29 gravitationally lensed galaxies with photometric or spectroscopic redshift in the range  $4.5 \leq z \leq 8$ . From a combined NIRSpec and NIRCам analysis, we were able to derive accurate physical properties of the galaxies, including  $M_{\star}$ , SFR,  $r_e$ , UV- $\beta$  slopes, and measurements of the most prominent optical emission lines ([OII], [OIII], H $\beta$ , H $\alpha$ ). We summarize our findings as follows:

- Our sample is composed of purely star-forming galaxies, as inferred from the MEx diagram that excludes the presence of any AGN.
- The galaxies in our sample have blue UV slopes (median  $\beta = -2.08$ ) and are mostly very compact with  $r_e \approx 0.2-0.5$  kpc. These properties are consistent with those of the general population of LBGs at similar redshift and with similarly faint  $M_{\text{UV}}$  (Bouwens et al. 2014; Shibuya et al. 2015).
- Compared to the low- $z$  sample by Flury et al. (2022) as well as to the low- $z$  LCEs from previous studies, our galaxies present properties (in terms of O32,  $\text{EW}_0(\text{H}\beta)$ , UV- $\beta$  slopes,  $r_e$ , and  $\Sigma_{\text{SFR}}$ ) that are entirely consistent with those of low- $z$  galaxies with measured  $f_{\text{esc}}$  larger than 0.05 and which are considered to be LyC leakers.
- Using a linear analysis that employs the minimization of MSE, we found a best fitting relation between  $f_{\text{esc}}$  and the three most correlated and independent parameters (O32, UV- $\beta$ , and  $r_e$ ) for the low-redshift sample of 88 galaxies. Applying this relation to the 24 galaxies from our sample where

these parameters are all measured, we find that 20/24 have predicted escape fractions larger than 0.05, that is, they would be considered leakers and the average  $f_{\text{esc}}$  of our sample is 0.12.

In conclusion, our results show that indeed the average low-mass galaxies around the epoch of reionization have physical and spectroscopic properties consistent with moderate values of escaping ionizing photons ( $f_{\text{esc}} = 0.1\text{--}0.2$ ). With upcoming JWST observations from GTO and GO programs, this analysis could be easily extended to larger samples of galaxies with higher masses and/or at higher redshift to determine if the bulk of the reionizing photons have been provided by even more massive and brighter galaxies (as predicted, e.g., by Sharma et al. 2016; Naidu et al. 2020) or if we need a more substantial contribution by the fainter and more numerous galaxies with  $M_{\text{UV}} > -18$ , as predicted by most other popular models (e.g., Trebitsch et al. 2022; Finkelstein et al. 2019; Atek et al. 2015).

**Acknowledgements.** Support for program JWST-ERS-1324 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127. This work is based on observations made with the NASA/ESA/CSA *James Webb* Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programs GLASS-JWST #1324, JWST DDT #2756 and GO UNCOVER #2561. S.M. thanks the Department of Physics and Astronomy at the University of California, Los Angeles, for a productive and satisfying visit in which the majority of the work presented in this paper was done. We acknowledge support from the INAF Large Grant 2022 “Extragalactic Surveys with JWST” (PI: Pentericci). This research is supported in part by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. L.Y. acknowledges support by JSPS KAKENHI Grant Number JP 21F21325. R.A.W. acknowledges support from NASA JWST Interdisciplinary Scientist grants NAG5-12460, NNX14AN10G and 80NSSC18K0200 from GSFC. M.B. acknowledges support from the Slovenian national research agency ARRS through grant N1-0238. C.M. acknowledges support by the VILLUM FONDEN under grant 37459. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNR140. K.G. and T.N. acknowledge support from Australian Research Council Laureate Fellowship FL180100060.

## References

- Alavi, A., Colbert, J., Teplitz, H. I., et al. 2020, *ApJ*, 904, 59
- Atek, H., Richard, J., Kneib, J.-P., et al. 2015, *ApJ*, 800, 18
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Bassett, R., Ryan-Weber, E. V., Cooke, J., et al. 2019, *MNRAS*, 483, 5223
- Bergamini, P., Acebron, A., Grillo, C., et al. 2023, *A&A*, 670, A60
- Bergvall, N., Leitert, E., Zackrisson, E., & Marquart, T. 2013, *A&A*, 554, A38
- Binggeli, C., Zackrisson, E., Pelckmans, K., et al. 2018, *MNRAS*, 479, 368
- Birrer, S., Shajib, A., Gilman, D., et al. 2021, *J. Open Source Softw.*, 6, 3283
- Bolan, P., Lemaux, B. C., Mason, C., et al. 2022, *MNRAS*, 517, 3263
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, *ApJ*, 793, 115
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calabrò, A., Castellano, M., Pentericci, L., et al. 2021, *A&A*, 646, A39
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Castellano, M., Dayal, P., Pentericci, L., et al. 2016, *ApJ*, 818, L3
- Chabrier, G. 2003, *PASP*, 115, 763
- Chisholm, J., Prochaska, J. X., Schaerer, D., Gazagnes, S., & Henry, A. 2020, *MNRAS*, 498, 2554
- Chisholm, J., Saldana-Lopez, A., Flury, S., et al. 2022, *MNRAS*, 517, 5104
- Coil, A. L., Aird, J., Reddy, N., et al. 2015, *ApJ*, 801, 35
- Curti, M., D’Eugenio, F., Carniani, S., et al. 2023, *MNRAS*, 518, 425
- de Barros, S., Vanzella, E., Amorín, R., et al. 2016, *A&A*, 585, A51
- Domínguez, A., Siana, B., Henry, A. L., et al. 2013, *ApJ*, 763, 145
- Dopita, M. A., & Sutherland, R. S. 2003, *Astrophysics of the Diffuse Universe* (New York: Springer)
- Faisst, A. L., Masters, D., Wang, Y., et al. 2018, *ApJ*, 855, 132
- Finkelstein, S. L., D’Aloisio, A., Paardekooper, J.-P., et al. 2019, *ApJ*, 879, 36
- Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022, *ApJS*, 260, 1
- Fontana, A., D’Odorico, S., Poli, F., et al. 2000, *AJ*, 120, 2206
- Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, *A&A*, 639, A85
- Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, *MNRAS*, 442, 1805
- Izotov, Y. I., Orlová, I., Schaerer, D., et al. 2016a, *Nature*, 529, 178
- Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, *MNRAS*, 461, 3683
- Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, 474, 4514
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, 478, 4851
- Izotov, Y. I., Guseva, N. G., Fricke, K. J., et al. 2021, *A&A*, 646, A138
- Izotov, Y. I., Chisholm, J., Worseck, G., et al. 2022, *MNRAS*, 515, 2864
- Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, *ApJ*, 736, 104
- Juneau, S., Bournaud, F., Charlot, S., et al. 2014, *ApJ*, 788, 88
- Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2020, *ApJ*, 904, 144
- Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, *ApJ*, 777, L8
- Katz, H., Ďurovčíková, D., Kimm, T., et al. 2020, *MNRAS*, 498, 164
- Kennicutt, R. C., Jr. 1998, *ARA&A*, 36, 189
- Mahler, G., Richard, J., Clément, B., et al. 2018, *MNRAS*, 473, 663
- Malkan, M. A., & Malkan, B. K. 2021, *ApJ*, 909, 92
- Marchi, F., Pentericci, L., Guaita, L., et al. 2018, *A&A*, 614, A11
- Mascia, S., Pentericci, L., Saxena, A., et al. 2023, *A&A*, submitted [arXiv:2301.09328]
- Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, *ApJ*, 856, 2
- Mason, C. A., Fontana, A., Treu, T., et al. 2019, *MNRAS*, 485, 3947
- Merlin, E., Bonchi, A., Paris, D., et al. 2022, *ApJ*, 938, L14
- Morishita, T., Roberts-Borsani, G., Treu, T., et al. 2022, *ApJL*, submitted [arXiv:2211.09097]
- Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, *ApJ*, 892, 109
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, 442, 900
- Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, 889, 161
- Nakajima, K., Ouchi, M., Isobe, Y., et al. 2023, *ApJS*, submitted [arXiv:2301.12825]
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713
- Ouchi, M., Ono, Y., & Shibuya, T. 2020, *ARA&A*, 58, 617
- Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., & Reddy, N. A. 2021, *MNRAS*, 505, 2447
- Pentericci, L., Fontana, A., Vanzella, E., et al. 2011, *ApJ*, 743, 132
- Pentericci, L., Vanzella, E., Castellano, M., et al. 2018, *A&A*, 619, A147
- Planck Collaboration VI 2020, *A&A*, 641, A6
- Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, *ApJ*, 788, 86
- Reddy, N. A., Topping, M. W., Shapley, A. E., et al. 2022, *ApJ*, 926, 31
- Richard, J., Claeysens, A., Lagattuta, D., et al. 2021, *A&A*, 646, A83
- Roberts-Borsani, G., Morishita, T., Treu, T., et al. 2022a, *ApJ*, 938, L13
- Roberts-Borsani, G., Treu, T., Chen, W., et al. 2022b, *Nature*, submitted [arXiv:2210.15639]
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, 802, L19
- Santini, P., Fontana, A., Castellano, M., et al. 2023, *ApJ*, 942, L27
- Saxena, A., Cryer, E., Ellis, R. S., et al. 2022, *MNRAS*, 517, 1098
- Schaerer, D., Izotov, Y. I., Worseck, G., et al. 2022, *A&A*, 658, L11
- Senchyna, P., Stark, D. P., Charlot, S., et al. 2022, *ApJ*, 930, 105
- Sersic, J. L. 1968, *Atlas de Galaxias Australes* (Cordoba: Observatorio Astronomico)
- Sharma, M., Theuns, T., Frenk, C., et al. 2016, *MNRAS*, 458, L94
- Shibuya, T., Ouchi, M., & Harikane, Y. 2015, *ApJS*, 219, 15
- Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, *MNRAS*, 408, 1628
- Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, *ApJ*, 869, 123
- Thomas, D., Steele, O., Maraston, C., et al. 2013, *MNRAS*, 431, 1383
- Trebitsch, M., Dayal, P., Chisholm, J., et al. 2022, *A&A*, submitted [arXiv:2212.06177]
- Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, *ApJ*, 935, 110
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, 825, 41
- Vanzella, E., Calura, F., Meneghetti, M., et al. 2019, *MNRAS*, 483, 3618
- Verhamme, A., Orlová, I., Schaerer, D., et al. 2017, *A&A*, 597, A13
- Wang, B., Heckman, T. M., Leitherer, C., et al. 2019, *ApJ*, 885, 57
- Xu, X., Henry, A., Heckman, T., et al. 2022, *ApJ*, 933, 202
- Yamanaka, S., Inoue, A. K., Yamada, T., et al. 2020, *MNRAS*, 498, 3095
- Yang, L., Birrer, S., & Treu, T. 2020, *MNRAS*, 496, 2648
- Yang, L., Leethochawalit, N., Treu, T., et al. 2022, *MNRAS*, 514, 1148
- Zackrisson, E., Inoue, A. K., & Jensen, H. 2013, *ApJ*, 777, 39
- Zackrisson, E., Binggeli, C., Finlator, K., et al. 2017, *ApJ*, 836, 78