The Astrophysical Journal Supplement Series, 272:33 (19pp), 2024 June https://doi.org/10.3847/1538-4365/ad3e7d

Quantifying the Escape of Lyα at $z \approx 5$–6: A Census of Lyα Escape Fraction with Hα-emitting Galaxies Spectroscopically Confirmed by JWST and VLT/MUSE

Xiaojing Lin[1,2], Zheng Cai[1], Yunjing Wu[1,2], Zihao Li[1], Fengwu Sun[2], Xiaohui Fan[2], Zuyi Chen[2], Mingyu Li[1], Fuyan Bian[3], Yuanhang Ning[1], Linhua Jiang[4,5], Gustavo Bruzual[6], Stephane Charlot[7], and Jacopo Chevallard[8]
[1] Department of Astronomy, Tsinghua University, Beijing 100084, People's Republic of China; zcai@mail.tsinghua.edu.cn
[2] Steward Observatory, University of Arizona, 933 N Cherry Avenue, Tucson, AZ 85721, USA
[3] European Southern Observatory, Alonso de Córdova 3107, Casilla 19001, Vitacura, Santiago 19, Chile
[4] Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China
[5] Department of Astronomy, School of Physics, Peking University, Beijing 100871, People's Republic of China
[6] Institute of Radio Astronomy and Astrophysics, National Autonomous University of Mexico, San José de la Huerta 58089 Morelia, Michoacán, México
[7] Sorbonne Université, CNRS, UMR7095, Institut d'Astrophysique de Paris, F-75014, Paris, France
[8] Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
Received 2023 December 19; revised 2024 March 22; accepted 2024 April 10; published 2024 May 27

Abstract

The James Webb Space Telescope provides an unprecedented opportunity for unbiased surveys of Hα-emitting galaxies at $z > 4$ with the NIRCam's wide-field slitless spectroscopy (WFSS). In this work, we present a census of Lyα escape fraction ($f_{\rm esc,Ly\alpha}$) of 165 star-forming galaxies at $z = 4.9$–6.3, utilizing their Hα emission directly measured from FRESCO NIRCam/WFSS data. We search for Lyα emission of each Hα-emitting galaxy in the Very Large Telescope/MUSE data. The overall $f_{\rm esc,Ly\alpha}$ measured by stacking is 0.090 ± 0.006. We find that $f_{\rm esc,Ly\alpha}$ displays a strong dependence on the observed UV slope ($\beta_{\rm obs}$) and $E(B-V)$, such that the bluest galaxies ($\beta_{\rm obs} \sim -2.5$) have the largest escape fractions ($f_{\rm esc,Ly\alpha} \approx 0.6$), indicative of the crucial role of dust and gas in modulating the escape of Lyα photons. $f_{\rm esc,Ly\alpha}$ is less well related to other parameters, including the UV luminosity and stellar mass, and the variation in $f_{\rm esc,Ly\alpha}$ with them can be explained by their underlying coupling with $E(B-V)$ or $\beta_{\rm obs}$. Our results suggest a tentative decline in $f_{\rm esc,Ly\alpha}$ at $z \gtrsim 5$, implying increasing intergalactic medium attenuation toward higher redshift. Furthermore, the dependence of $f_{\rm esc,Ly\alpha}$ on $\beta_{\rm obs}$ is proportional to that of the ionizing photon escape fraction ($f_{\rm esc,LyC}$), indicating that the escape of Lyα and ionizing photon may be regulated by similar physical processes. With $f_{\rm esc,Ly\alpha}$ as a proxy to $f_{\rm esc,LyC}$, we infer that UV-faint ($M_{\rm UV} > -16$) galaxies contribute >70% of the total ionizing emissivity at $z = 5$–6. If these relations hold during the epoch of reionization, UV-faint galaxies can contribute the majority of UV photon budget to reionize the Universe.

Unified Astronomy Thesaurus concepts: High-redshift galaxies (734); Reionization (1383)

1. Introduction

The epoch of reionization (EoR) represents a phase transition when the intergalactic medium (IGM) evolved rapidly from mostly neutral to highly ionized (Fan et al. 2006; Robertson 2022). Although it is generally acknowledged that EoR is completed by $z \sim 5$–6 (e.g., Kulkarni et al. 2019; Zhu et al. 2021; Jin et al. 2023), its onset and major contributors remain elusive. Recent measurements on quasar luminosity functions (LFs) at $z \sim 6$ suggest that the quasar population only contributes a negligible fraction ($\lesssim$7%–10%) of total photons to keep the Universe ionized (Matsuoka et al. 2018; Jiang et al. 2022). Young star-forming galaxies are thought to be the main producers of ionizing photons for reionization. However, one of the biggest uncertainties is whether a vast number of faint galaxies (e.g., Finkelstein et al. 2019) or a small population of bright oligarchs (e.g., Naidu et al. 2020) dominate the cosmic ionizing emissivity during the EoR. To address these questions, extensive studies in the past decades focused on three key quantities observationally: the ionizing photon production efficiency $\xi_{\rm ion}$, the UV luminosity density $\rho_{\rm UV}$, and the escape fraction of Lyman-continuum (LyC) photons $f_{\rm esc,LyC}$ (Robertson et al. 2013; Bouwens et al. 2016).

Among the above three elements, the biggest uncertainty is from the estimates of $f_{\rm esc,LyC}$ at high redshifts (Vanzella et al. 2010; Bian & Fan 2020; Naidu et al. 2022). Directly observing the leakage of LyC photons at $z \gtrsim 4$ (Steidel et al. 2018) is difficult due to the high opacity of the IGM. Therefore, many works studied low-redshift galaxies with similar properties (e.g., Leitherer et al. 2016; Flury et al. 2022a; Saldana-Lopez et al. 2022) or indirect tracers of $f_{\rm esc,LyC}$ (e.g., Flury et al. 2022b; Xu et al. 2022; Choustikov et al. 2023), although it is still unclear whether low-redshift analogs can be good representatives of the main contributors to reionization (e.g., Katz et al. 2020; Schaerer et al. 2022), and whether those indirect indicators of $f_{\rm esc,LyC}$ derived from very small sample could be generalized (e.g., Katz et al. 2022; Saxena et al. 2022). The typically required $f_{\rm esc,LyC}$ for reionization to end at $z \sim 5$–6 is approximately $\gtrsim$5% (Robertson et al. 2015; Finkelstein et al. 2019), and recent work have constrained the average $f_{\rm esc,LyC}$ to be $\lesssim$10% through various indirect methods (Steidel et al. 2018; Pahl et al. 2021; Saldana-Lopez et al. 2023).

To directly constrain the $f_{\rm esc,LyC}$ for galaxies at $z \gtrsim 4$, one of the most promising ways is to study the escape of Lyα. Numerous studies have shown a strong connection between the escape of ionizing photons and Lyα photon (Chisholm et al. 2018; Marchi et al. 2018; Maji et al. 2022). Their escapes are both determined by the distribution and geometry of neutral gas

in the interstellar medium (ISM), and facilitated by optically thin channels within the gas reservoir (Dijkstra et al. 2016; Gazagnes et al. 2020; Lin et al. 2023). These ionized channels could be carved by feedback from massive young stars, which also modulate the production of LyC and Lyα photons (Vanzella et al. 2022; Kim et al. 2023).

A number of previous efforts have been made to probe the escape of Lyα ($f_{\rm esc,Ly\alpha}$), which could be derived by comparing Lyα and Hα emissions. Before the launch of James Webb Space Telescope (JWST), Hα of galaxies at $z \gtrsim 5$ could be estimated only through the Spitzer broadband color (e.g., Faisst et al. 2019; Lam et al. 2019; Stefanon et al. 2022). Hence, the measurements of $f_{\rm esc,Ly\alpha}$ in the prelaunch era suffer from large modeling and photometric uncertainties. Improved estimates of the Hα flux excess have been done recently using JWST narrowband observations (Ning et al. 2023; Simmonds et al. 2023). Moreover, the spectroscopic capability of JWST in the infrared regime enables us to detect and accurately measure Hα emission lines in faint, high-redshift galaxies. The JWST Near Infrared Spectrograph (NIRSpec) can capture both the Lyα and Hα emission of high-redshift galaxies simultaneously (Chen et al. 2024; Saxena et al. 2024) and thus quantify $f_{\rm esc,Ly\alpha}$, but a census of $f_{\rm esc,Ly\alpha}$ by NIRSpec is difficult due to the limited sample size it can observe and the trade-off between high sensitivity and high spectral resolution. The wide-field slitless spectroscopy (WFSS) of JWST/NIRCam (Greene et al. 2016; Rieke et al. 2023), instead, offers a unique capability to map the spectra of all sources within its field of view (FoV; Sun et al. 2022, 2023), ideal to build a large sample of emission line galaxies through blind spectroscopic surveys. Combining the Hα emission lines captured by WFSS with the corresponding Lyα lines observed by ground-based facilities, we are able to, for the first time, conduct a direct census of $f_{\rm esc,Ly\alpha}$ with an unbiased sample limited by Hα flux.

In this work, we utilize JWST/NIRCam WFSS around 4 μm to detect Hα emission over the redshift range of $z \approx 4.9$–6.3, when the Universe was experiencing rapid transition, and galaxies were undergoing substantial mass, dust, and chemical content assembly (Madau & Dickinson 2014; Davidzon et al. 2017; Faisst et al. 2019). The ground-based Very Large Telescope (VLT)/Multi Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010), working in the full optical domain, perfectly covers their Lyα so that we can obtain more precise measurements of $f_{\rm esc,Ly\alpha}$.

The paper is organized as follows. In Section 2, we outline the photometric and spectroscopic data used and sample selection based on Hα emission lines. We describe the measurements of galaxy properties, search for Lyα lines, and methods to calculate $f_{\rm esc,Ly\alpha}$ in Section 3. In Section 4, we present the distribution of $f_{\rm esc,Ly\alpha}$, its dependence on galaxy properties and redshift evolution. We then investigate the connection between LyC and Lyα escape in Section 5, and discuss the implication for reionization. The final summary is presented in Section 6. We show more details about the analysis and address potential systematics in Appendix A. Throughout this work, a flat ΛCDM cosmology is assumed, with $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\Lambda,0} = 0.7$, and $\Omega_{m,0} = 0.3$.

2. Data and Sample

We describe the data set that we use in this work and the construction of our sample in this section. We introduce all the photometric and spectroscopic data, including JWST, Hubble Space Telescope (HST), and VLT/MUSE, in Section 2.1. The selection of Hα emitters (HAEs) is presented in Section 2.2.

2.1. Photometric and Spectroscopic Data

2.1.1. JWST/NIRCAM Imaging and Slitless Grism in the GOODS-S Field

The JWST "First Reionization Epoch Spectroscopically Complete Observations" (FRESCO) Survey (Oesch et al. 2023, GO-1895; PI Oesch), covers 62 arcmin^2 in each of the extragalactic legacy GOODS/CANDELS fields with deep NIRCam imaging and WFSS observations. In GOODS-South field, FRESCO obtains 4×2 pointings with 8×7043 s exposures using the F444W grism, accompanied by 8×3522 s exposures of F210M imaging and 11×4456 s of F182M imaging. Additional 8×934 s exposures of F444W imaging are taken to associate the slitless spectra with individual objects.

The imaging data were first reduced up to `stage2` using the standard JWST pipeline v1.7.2[9] and the calibration reference files "jwst_1014.pmap." The calibrated single exposures (`*_cal.fits`) were further processed by GRIZLI.[10] GRIZLI mitigated 1/f noise, alleviated "snowball" artifacts from cosmic rays, and converted the world coordinate system (WCS) information in the headers of each exposure to a simple imaging polynomial format so that exposures can be drizzled and combined with ASTRODRIZZLE. The WCS of each exposure image was registered using GAIA DR3 (Gaia Collaboration et al. 2023), and the images were finally drizzled with a pixel scale of 0.″031 using `pixfrac=0.8`. An additional background subtraction using PHOTUTILS.BACKGROUND2D was performed on the final mosaics after masking bright sources.

For the WFSS data, after the standard `stage1` steps, flat-fielding was performed using the flat images of direct imaging since the WFSS flat reference files are not available yet in CRDS. 1/f noise is removed only along columns as the spectra are dispersed along rows. The WCS for each individual exposure was assigned using `CALWEBB.assign_wcs`. There were astrometric offsets between this assigned WCS and that of the fully reduced F444W images. To associate the WFSS spectra and individual galaxies accurately, we calculated the offsets using accompanied short-wavelength exposures (i.e., the F182M and F210M imaging taken simultaneously with the WFSS) and corrected them for each WFSS exposure. These preprocessed files, as well as a source catalog based on the final F444W mosaic, were provided as inputs for GRIZLI for spectral extraction using the updated sensitivity functions and tracing models as described in Sun et al. (2023).[11] To facilitate emission line searching, we performed running median filtering along the dispersion direction on single calibrated grism exposures as described in Kashino et al. (2023), and extracted continuum-removed 1D spectra aiming for emitter searching only.

2.1.2. HST Imaging in GOODS-S Field

The high-level products of the Hubble Legacy Fields[12] (HLF; Whitaker et al. 2019) in GOODS-S include images from ultraviolet to infrared bands, observed over the past 18 yr, with

[9] https://github.com/spacetelescope/jwst
[10] https://github.com/gbrammer/grizli/
[11] https://github.com/fengwusun/nircam_grism
[12] http://archive.stsci.edu/hlsps/hlf

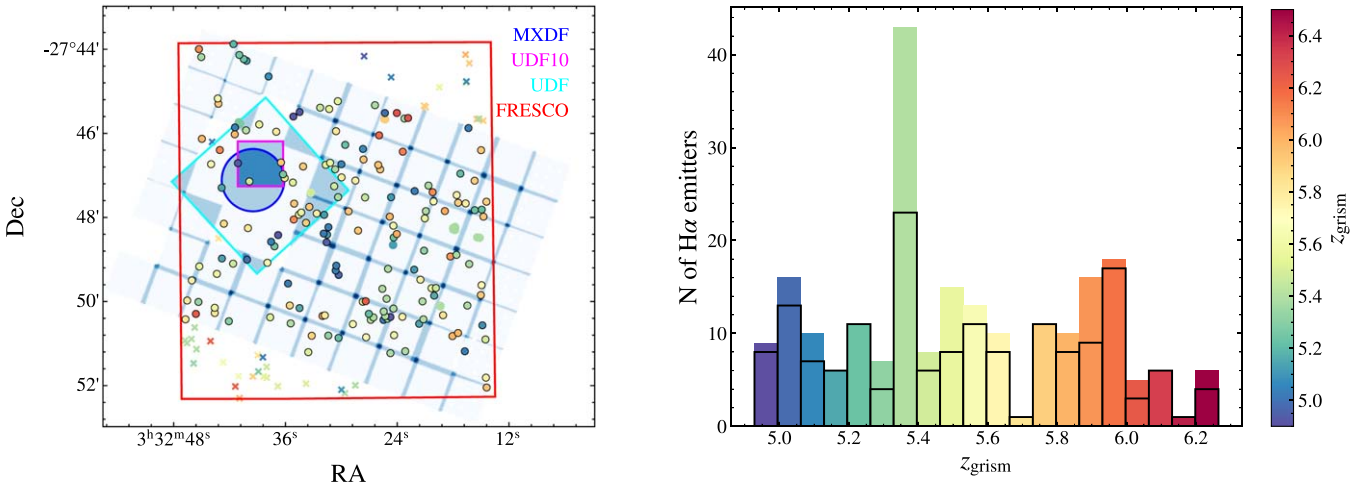

Figure 1. The spatial and redshift distribution of Hα emitters. Left: the distribution of Hα sample in the footprints of multiband legacy data in GOODS-S. The blue-shaded region is covered by MUSE observations, including the MUSE-Wide (Urrutia et al. 2019) and MUSE Hubble UDF survey (Bacon et al. 2023), color-coded by the number of overlapping data cubes (darker color indicates more data cubes). The cyan, blue, and magenta frames represent the footprints of three fields in MUSE Hubble UDF survey: the 10 hr UDF, 31 hr UDF-10, and the >100 hr MXDF; the red rectangle shows the FoV of FRESCO JWST/NIRCam F444W imaging. Color-coded markers indicate the position of our Hα emitters. The color indicates their redshift in exactly the same scale as that on the right panel. An overdensity at $z \approx 5.3$–5.4 is obvious, which has been reported in Helton et al. (2023). The crosses denote Hα emitters not covered by MUSE (41 in total); dots are within MUSE coverage (197 in total), and dots with black edges are selected sources we use for analysis (157 in total). Right: the redshift distribution of our Hα emitters. The distribution of selected Hα sources is wrapped with black edges.

7491 individual exposures and 6.4 million s in total. All the HLF image mosaics are tied to an absolute GAIA DR2 (Gaia Collaboration et al. 2018) reference frame. In this work, we use four optical images in the Advanced Camera for Surveys (ACS)/WFC F435W, F606W, F775W, F814W bands at 30 mas pixel^{-1}, and four infrared images in the WFC3/IR F105W, F125W, F140W, and F160W bands at 60 mas pixel^{-1}.

2.1.3. Source Extraction and Photometry

We first resampled all HST images onto a fixed grid with a pixel scale of 0.″031. For HST mosaics, we constructed empirical point-spread functions (PSFs) by stacking stars with good signal-to-noise ratios (S/Ns) in the fields. For JWST mosaics, the centers of most bright stars are saturated especially in the F182M and F210M images, so it is hard to build empirical PSFs that can accurately reproduce the complicated PSF shapes with these limited numbers of stars in the final mosaics. Therefore, we adopted the modeled PSFs provided by WEBBPSF.[13] We generate PSF matching kernels to match the PSFs of HST WFC3/IR F160W or JWST/NIRCam F444W, using PHOTUTILS,[14] with high-frequency noise filters in Aniano et al. (2011). All HST ACS/WFC mosaics (F435W, F606W, F775W, F814W), JWST/NIRCam F182M, F210M mosaics were convolved to match the PSF of JWST/NIRCam F444W with corresponding kernels; all HST WFC3/IR mosaics (F105W, F125W, F140W) were matched to the PSF of F160W.

We ran SOURCE EXTRACTOR (Bertin & Arnouts 1996) in dual mode for each band, using the F444W mosaics as the detection image. We started with Kron aperture photometry with Kron factors $k = 1.2$. We measured the flux uncertainty for each object by randomly positioning the same Kron aperture in empty regions nearby. The measured flux and uncertainties were further corrected to total values by multiplying the ratio of flux inside the $k = 2.5$ versus $k = 1.2$ Kron apertures on the detection image. Finally, we correct for flux outside the $k = 2.5$ radii using the PSF of F444W. Note that for HST WFCS/IR bands we homogenized the images to the PSF of F160W, slightly larger than that of JWST/NIRCam F444W; thus, we enlarged the aperture size on these HST IR mosaics by a factor of 1.2, which ensures that all the flux from objects would be included as tested on the F160W PSF.

2.1.4. VLT/MUSE in GOODS-S Field

MUSE is an integral-field spectrograph in the optical wavelength range on the VLT. MUSE has an FoV of 1 arcmin^2, a spatial sampling of 0.″2, and a spectral range from 4750 to 9350 Å with a mean resolution of 3000. Its high sensitivity to emission lines in the full optical domain makes it an ideal instrument to detect high-redshift Lyα emitters (LAEs) over $z = 2.9$–6.7.

The GOODS-S field is covered by the legacy MUSE-Wide survey (Urrutia et al. 2019) and MUSE Hubble Ultra Deep Field (UDF) survey (Bacon et al. 2023). MUSE-Wide comprises 60×1 arcmin^2 pointings, each with a depth of 1 hr. The MUSE UDF survey includes 3×3 arcmin^2 mosaic of nine 10 hr depth fields (MOSAIC), a 1×1 arcmin^2 31 hr depth field (UDF-10), and the circular MUSE eXtremly Deep Field (MXDF). The MXDF covers a radius of 41″ and 31″ for respectively 10+ and 100+ hr of depth, reaching a final maximum depth of 141 hr, and is the deepest optical spectroscopic survey so far. In Figure 1, we show the footprints of FRESCO, MUSE-Wide, and MUSE UDF. The overlap region between FRESCO and the two MUSE surveys is about 50 arcmin^2, including ∼43 arcmin^2 of the MUSE-Wide, ∼9 arcmin^2 of the 10 hr MOSAIC, and the whole UDF-10 and MXDF.

We adopt the publicly released DR1[15] data for MUSE-Wide and AMUSED[16] data for MUSE UDF. For MUSE-Wide, DR1

[13] https://github.com/spacetelescope/webbpsf
[14] HTTPS://PHOTUTILS.READTHEDOCS.IO/EN/STABLE/PSF_MATCHING.HTML
[15] https://musewide.aip.de/
[16] https://amused.univ-lyon1.fr/project/UDF/HUDF/

only releases 44 of the 60 data cubes, so we reduced the remaining 16 data cubes with our custom pipeline. This includes the standard pipeline[17] `v2.8.4` (Weilbacher et al. 2020) and improved the sky subtraction using ZAP[18] (Soto et al. 2016). We aligned all the data cubes to the Hubble ACS astrometry as done by the MUSE-Wide team. The offset between the HST astrometry and the GAIA DR2, adopted in WFSS and imaging, is $\Delta\mathrm{R.A.} = +0\farcs094 \pm 0\farcs042$, and $\Delta\mathrm{decl.} = -0\farcs26 \pm 0\farcs10$ (Bacon et al. 2023).

2.2. Selection of Hα Emitters at z = 5–6

2.2.1. Selecting High-redshift Emitters Both Photometrically and Spectroscopically

The F444W WFSS of JWST/NIRCam is used to conduct unbiased surveys of HAEs at $z = 5$–6. We performed emission line searching on the 1D continuum-removed spectra first. Among all the emitter candidates, we identified possible high-redshift emitter candidates by applying the following cuts:

(1) 1D emission line with S/N $\geqslant$ 4,
(2) S/N <2 in F435W,
(3) F606W-F105W>1 and F606W-F182M>1,
(4) F182M-F444W>-0.1.

The second criterion requires a complete 912 Å dropout for $z > 4.9$ galaxies at F435W. The third criterion describes the Lyα break features; we require at least 1 mag drop in F606W that is blueward of the wavelength of Lyα. The fourth criterion is the F444W excess due to Hα emission lines when considering possible Balmer jumps. This is the minimum excess for $z = 4.9$–6.7 mock galaxies in `JAGUAR` (Williams et al. 2018) with Hα flux larger than 1.6×10^{-18} erg s^{-1} cm^{-2} (4σ depth of the FRESCO F444W grism data according to Oesch et al. 2023). We also obtained photometric redshifts for each source using EAZY (Brammer et al. 2008) with the `corr_sfhz_13` templates, which contain redshift-dependent star formation histories. We excluded objects with robust photo-z, i.e., $P(z_p > 4.5) = 0$, as low-redshift objects. We finally constructed a $z > 4.9$ emitter catalog by careful visual inspection, during which we removed artifacts on imaging and grism data, false detection by the search algorithms, and misidentified lines from the contamination of other sources.

2.2.2. HAE Sample

The high-redshift emitter catalog described above comprises all $z > 4.9$ candidates, including $z > 7$ [O III] emitters. Bright [O III] emitters are easily identified by their doublets at rest-frame 4959 and 5007 Å. However, as the [O III] 4959 Å lines are ∼3 times fainter than the 5007 Å, and Hβ can be even fainter, in many cases, we can only detect a single 5007 Å emission while the 4959 Å lines remain undetected (see Appendix A for more details). Since our science goal is more sensitive to the sample purity than completeness, we exclude all possible [O III] candidates based on both photometry and spectroscopy. We first assumed the brightest single emission line in the 1D spectra to be 5007 Å lines and force-fitted 4959 lines. This would effectively identify [O III] emitters with marginally detected 4959 lines. We then further extend the [O III] sample by photometric cuts as follows:

1. S/N <2 in both F606W and F775W,
2. F606W–F115W>1.7 and F814W–F115W>1.7,
3. F115W–F210M<1.0 and F115W–F210M<1.0,
4. F814W–F115W>(F115W-F210M)+1.5.

Similar criteria have been applied in Endsley et al. (2022) to select $z > 6.5$ emitters with high [O III] equivalent width (EW). Note that we do not require F814W to be completely undetected, allowing for Lyα transmission spikes within its wide wavelength coverage (Kakiichi et al. 2018).

We find 74 *possible* [O III] candidates. 29 of the 74 [O III] candidates reside in the JADES-DEEP footprint (Eisenstein et al. 2023; Rieke et al. 2023). Among the 29 sources, 21 sources have photo-$z > 7$ based on multiple bands of JADES imaging, with [O III] emission in F444W; eight of them have photo-z of $5 - 6$ with Hα in F444W. We argue that our criteria for [O III] could, in most cases, successfully remove $z > 7$ [O III] emitters and thus keep an Hα sample with high purity.

Finally, we obtain a sample of 222 HAEs. We identify blended clumpy galaxies with irregular morphology and multiple cores as single objects. These clumpy HAEs have Hα emission lines that cannot be distinguished on the 2D grism spectra, and some of them have unsolved Lyα emission captured by MUSE, even though SEXTRACTOR may label them as different individuals. During the extraction of grism spectra, we use a modified segmentation map as the input of GRIZLI, in which clumpy segments are combined. In our final sample, any two "single" objects are required to be separated from each other by at least 3.5 kpc (i.e., 0.″6 at $z = 5.5$). We measure the total Hα flux on 1D extracted spectra by fitting a "Gaussian + constant" model, where the constant is used to depict the local continuum level around Hα.

Moreover, we label sources having neighbors within 25 kpc (about 4″ at $z = 5$–6) and ±500 km s^{-1}. Due to the significant disparity between the spatial resolutions of MUSE and JWST, we will exclude these galaxies and their close neighbors in the following analysis, since (1) Lyα around individual galaxies can be extended to 3″ (Wisotzki et al. 2016) and will boost Lyα in the overlapping regions of close pairs, introducing systematics during the stack of Lyα as described in Section 3.3.2, and (2) for very close pairs, it is difficult to obtain accurate Hα flux due to the overlap issue along the dispersion direction of grism. Among all the 222 HAEs, 21 of them have close neighbors. Three of the 222 HAEs are bright point sources with hints as being active galactic nuclei (AGNs). There are 184 HAEs that reside in the MUSE footprints, with 17 having close neighbors and two as potential AGNs. Finally, 165 HAEs are selected for the following studies. The spatial and redshift distributions of our selected sample are shown in Figure 1. Detailed properties of these HAEs are presented in Appendix A, which demonstrate that they are typical main-sequence galaxies in this redshift range.

3. Measurements

We describe the methods we use to quantify the properties of our HAE sample in Section 3.1. The search and measurements of Lyα emission using VLT/MUSE data are outlined in Section 3.2. The calculation of $f_{\mathrm{esc,Ly}\alpha}$ for individual galaxies and through a stack analysis is presented in Section 3.3.

3.1. SED Modeling and Galaxy Properties

We infer the physical properties of the 165 $z \approx 5$–6 HAEs by spectral energy distribution (SED) modeling with the Bayesian

[17] https://www.eso.org/sci/software/pipelines/muse/
[18] https://github.com/musevlt/zap

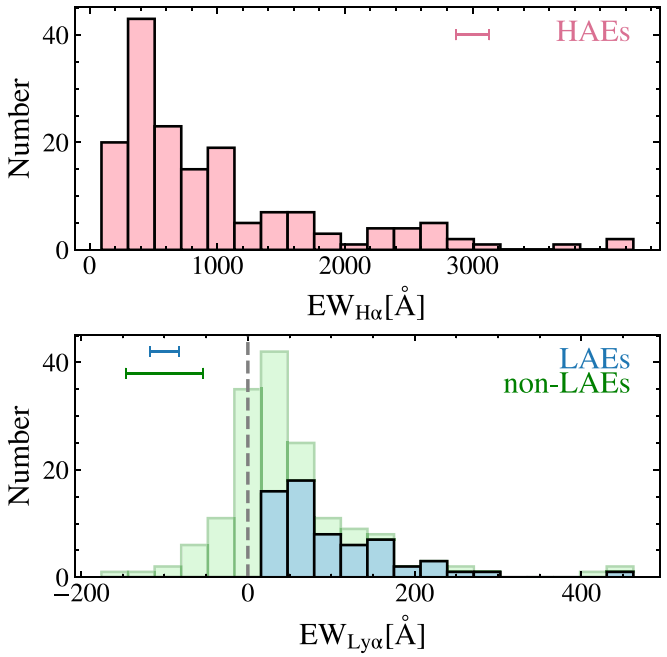

Figure 2. The distribution of $EW_{H\alpha}$ and $EW_{Ly\alpha}$ of the 165 selected HAEs in the overlapping region of FRESCO and MUSE footprints. The EWs are calculated following Equation (5) with f_{cont} from the SED-modeled stellar continuum assuming the Calz dust attenuation law. We label the mean uncertainty of $EW_{H\alpha}$ (pink in the left panel) and $EW_{Ly\alpha}$ (blue for LAEs, and green for non-LAEs in the right panel) as horizontal lines, respectively. For non-LAEs, we note that the uncertainty of forced measured $EW_{Ly\alpha}$ reaches an average of 163 Å when $EW_{Ly\alpha} > 100$ Å, although the mean uncertainty for non-LAEs is about 47 Å. We mark the $EW_{Ly\alpha} = 0$ as the gray dashed line.

code BEAGLE (Chevallard & Charlot 2016), using the measured photometry (see Section 2.1.3) and Hα line flux as inputs. We do not use the Lyα flux as inputs because modeling Lyα and Hα simultaneously requires correctly taking the Lyα escape process into account, and current SED modeling codes cannot do it. We assume a constant star formation history and the Chabrier (2003) initial mass function (IMF) with an upper limit of $100 M_\odot$. We adopt flat priors for metallicity, $0.01 < Z/Z_\odot < 0.2$, and ionization parameter for nebular emission (Gutkin et al. 2016) in log-space, $-3 < \log U < -1$, which is the dimensionless ratio of the number density of H-ionizing photons to that of hydrogen. For dust attenuation, we adopt two attenuation curves: the Small Magellanic Cloud (SMC) dust law (Pei 1992, hereafter SMC) and the Calzetti et al. (2000) law (hereafter Calz) with R_V (ratio of total to selective extinction in V band) being 2.74 and 4.05 respectively. We set the optical depth in the V band varying in the range of 0–0.5 in log-space. Although a large variety of dust attenuation curves for high-redshift galaxies have been used in the literature (Reddy et al. 2015; Scoville et al. 2015; Reddy et al. 2018), these two are commonly suggested to be applicable and utilized broadly in high-redshift studies (e.g., Bouwens et al. 2016; Shivaei et al. 2018; Faisst et al. 2019). In the following analysis, we adopt the Calz as our fiducial dust attenuation law, but we also consider the SMC law for comparison.

The BEAGLE SED fitting results can provide the stellar mass ($M_\star$), the effective optical depth at the V band (τ_V), the attenuated UV luminosity $L_{UV,att}$, and the intrinsic UV luminosity $L_{UV,int}$ directly. The UV luminosity is defined at rest-frame 1600 Å, and the optical depth could be converted to the color excess $E(B-V)$ with corresponding attenuation curve. We measure the observed UV slope (β_{obs}, not corrected for dust) by fitting the stellar continua of the best-fit SED models with a power-law function ($f_\lambda = \lambda^{\beta_{obs}}$), spanning from 1260 to 2500 Å. The SED-derived UV luminosity and UV slopes are statistically consistent with the fitting results using the photometry directly. We estimate the specific star formation rate (sSFR) for each galaxy with their intrinsic Hα luminosity and SED-derived $M_\star$. The star formation rate (SFR) of each galaxy is converted based on Kennicutt (1998) specialized for the Chabrier (2003) IMF:

$$\mathrm{SFR}[M_\odot \mathrm{yr}^{-1}] = \log L_{H\alpha}[\mathrm{ergs}^{-1}] - 41.35. \quad (1)$$

Note that the conversion factors between the Hα luminosity and SFR for high-redshift galaxies vary among different stellar population models and IMFs, and are highly dependent on the ionizing photon production efficiency ξ_{ion} (e.g., Theios et al. 2019). The conversion factor of subsolar-metallicity galaxies is about 2.5 lower than the canonical conversion factors of low-redshift galaxies (e.g., Reddy et al. 2018; Theios et al. 2019; Shapley et al. 2023). Nevertheless, as we focus on the relative relation between $f_{esc,Ly\alpha}$ and SFR in the following analysis (see Section 4.2), Equation (1) is qualitatively a good proxy to the SFR of our Hα-emitting galaxies.

To derive the sSFR surface density (Σ_{sSFR}) , we fit the F444W surface brightness distribution, i.e., the approximate Hα surface brightness, for each galaxy using single or multiple Sérsic models. The effective radius (r_{eff}), which is the radius of the circular aperture including half of the total flux from the galaxy, is measured in the corresponding compound modeled image after PSF deconvolution.[19] We do not adopt r_{eff} in simple single Sérsic modeling due to the clumpy morphology as well as the highly coupled parameters during modeling. Then, we calculate Σ_{sSFR} as

$$\Sigma_{sSFR} = \mathrm{sSFR}/2\pi r_{eff}^2. \quad (2)$$

We further measure the ionizing efficiency ξ_{ion}, i.e., the production rate of LyC ionizing photons per unit nonionizing UV continuum luminosity. ξ_{ion} can be expressed in terms of $N(H^0)$, the production rate of LyC photons, and $L_{UV,int}$ as

$$\xi_{ion} = \frac{N(H^0)}{L_{UV,\,int}}[\mathrm{s}^{-1}/\mathrm{erg\ s}^{-1}\,\mathrm{Hz}^{-1}] \quad (3)$$

where

$$L_{H\,\alpha,int} = 1.36 \times 10^{-12}(1 - f_{esc,\,LyC})N(H^0). \quad (4)$$

$f_{esc,LyC}$ is the escape fraction of LyC photons from galaxies. Our result, as shown in Section 5, shows that $f_{esc,LyC}$ has negligible impact to the estimate of ξ_{ion}, so we assume $f_{esc,LyC}$ as zero in Equation (4). Therefore, ξ_{ion} we adopt is actually the production efficiency of ionizing photons failing to escape from the galaxy.

We present the EW distribution of our HAE samples in Figure 2 (left panel). The EWs are determined by

$$\mathrm{EW} = \int_{\lambda_0}^{\lambda_1} \frac{f_{line}}{f_{cont}} d\lambda \approx \frac{F_{line}}{f_{cont}}, \quad (5)$$

[19] https://users.obs.carnegiescience.edu/peng/work/galfit/TFAQ.html#effective_radius

where λ_0, λ_1 define the wavelength range for integration, $F_{\rm line}$ is the emission line flux, and $f_{\rm cont}$ is the rest-frame continuum flux density. For Hα, we obtain $f_{\rm cont}$ by fitting the 6530–6600 Å stellar continuum provided by the best-fit SED using power-law models, and extrapolate to the wavelength of Hα. The Hα EW ($\rm EW_{H\alpha}$) of our sample ranges from 93 to 4300 Å, with a median value of 665 ± 67 Å.

3.2. Search for Lyα Emission

We search for Lyα emission for all the HAEs covered by MUSE observations. Instead of the blind line algorithm usually adopted in MUSE 3D data cubes, we now know exactly the position and redshift of our targets and can well predict where the Lyα emission lines could emerge. For each Hα source, we extract $20'' \times 20''$ subcubes and perform median filtering, with a window size of 101 pixels to remove continuum and nearby sources. For sources in overlap regions of multiple pointings, we combine all the individual subcubes weighted by their exposure time.[20] We then search for local surface brightness peaks on 3 pixel width pseudo-narrowband slices and 4×4 pixel boxes, centered on the source position and its rest-frame 1216 Å. The box size is determined by the typical size of UV star-forming region of high-redshift galaxies shown in Kusakabe et al. (2022). We perform the search within a radius of $1\farcs5$ spatially and a spectral range of 1500 km s^{-1}. This radius ensures the inclusion of potential Lyα flux given a possible offset between the Lyα and Hα due to gas distribution and radiation transfer (e.g., Cai et al. 2019; Zhang et al. 2023). For each selected flux peak, we extract a 1D spectrum using the 4×4 pixel box, and fit a skewed Gaussian profile. We visually inspect the extracted spectra and identify 63 HAEs with Lyα emission lines detected at $>2\sigma$. In the following analysis, we refer to the HAEs with Lyα lines detected as "LAEs" and the remaining as "non-LAEs"; but we note that this definition fully depends on the depth of the MUSE observations and is not comparable to the literature definitions (e.g., Lyα EW >10 Å). Among the 63 LAEs, 39 have been reported in Kerutt et al. (2022) or Bacon et al. (2023).

We further extract 1D spectra for all the selected 165 HAEs using $r = 1\farcs5$ circular apertures, and measure the uncertainties by the standard deviation of 500 spectra extracted using equally sized apertures randomly positioned in each subcube. For LAEs, we fix the skewed Gaussian profile derived from the boxy-aperture extracted spectra to maximize S/Ns, and adjust the amplitude to fit the circular-aperture extracted spectra and obtain the total flux. Lyα of two LAEs are overwhelmed by noises in their circular-aperture extracted spectra; we thus adopt the boxy-aperture measured flux for them. Comparing the flux of LAEs we measure with the values reported in Kerutt et al. (2022) and Bacon et al. (2023), the median flux ratio is 1.04 ± 0.38 implying a good consistency. For non-LAE, we measure the integrated flux and uncertainties over ± 1000 km s^{-1} around the wavelength of Lyα within the circular apertures. All non-LAEs do not show emission line peaks around the Lyα wavelength and have integrated Lyα fluxes with S/Ns <3.

[20] Note that the MUSE data cubes we adopt are products of different pipelines, with significant differences in estimating uncertainties. The standard pipeline underestimates the true uncertainties, because it does not consider correlated noise due to resampling. Thus, additional variance calibration procedures are carried out in Urrutia et al. (2019) and Bacon et al. (2023) respectively. To avoid the systematics induced by different variance estimates, we do not use any variance information in this work.

Figure 3 shows six examples of LAEs and non-LAEs among our HAE sample. We show the distribution of Lyα EW ($\rm EW_{Ly\alpha}$) of LAEs and non-LAEs in the right panel of Figure 2. $\rm EW_{Ly\alpha}$ are measured following Equation (5), with $f_{\rm cont}$ obtained by fitting the 1250–1350 Å stellar continuum. Our detected LAEs have rest-frame EW ($\rm EW_{Ly\alpha}$) spanning from 19 Å to 462 Å, with a median of 76 Å. For non-LAE, $\rm EW_{Ly\alpha}$ derived from the forced measured Lyα within ± 1000 km s^{-1} spans from –175 Å to 441 Å, with a median value of 11 Å. Their 3σ upper limits range from 8 Å to 1123 Å with a median value of 91 Å.

3.3. Measurement of Lyα Escape Fraction

In this section, we describe our method to measure $f_{\rm esc,Ly\alpha}$ in detail. We outline the measurement for individual HAEs in Section 3.3.1 and the stack procedure in Section 3.3.2, which is aimed at measuring the overall $f_{\rm esc,Ly\alpha}$ of a specific population.

3.3.1. $f_{esc,Ly\alpha}$ of Individual HAEs

Under the assumption of Case B recombination for $T = 10^4$K, and $n_e \approx 350$ cm^{-3} gas, the Lyα escape fraction could be expressed as (Osterbrock 1989)

$$f_{\rm esc,Ly\alpha} = \frac{F_{\rm Ly\alpha,obs}}{8.7 F_{\rm H\alpha,int}} \quad (6)$$

where $F_{\rm Ly\alpha,obs}$ is the observed Lyα flux, and $F_{\rm H\alpha,int}$ is the intrinsic Hα flux where the dust attenuation effect has been corrected. For LAEs, $f_{\rm esc,Ly\alpha}$ could be directly calculated following Equation (6) with the detected Lyα emission flux as described in Section 3.2; for non-LAEs, we use the integrated flux over ± 1000 km s^{-1} within the $r = 1\farcs5$ circular aperture. We correct the dust attenuation with the Calz and SMC laws and the optical depth in the V band derived from `Beagle` respectively. The differential dust attenuation factor between stars and nebular regions is set as $f_{\rm nebular} = 1$, as suggested by a number of previous studies that $f_{\rm nebular}$ approaches unity at $z > 2$ with Hα EW >100 Å (e.g., Faisst et al. 2019). The distribution of $f_{\rm esc,Ly\alpha}$ of individual HAEs will be discussed in Section 4.1

3.3.2. $f_{esc,Ly\alpha}$ by VLT/MUSE Stack

To calculate the typical $f_{\rm esc,Ly\alpha}$ of a specific population comprising HAEs with both Lyα detected and not detected, we stack the MUSE data cubes and measure the stacked Lyα. We shift the median-filtered subcubes of each HAE to its rest frame and then resample the subcubes into a fixed grid along the spectral axis. The rest-frame subcubes are then normalized based on their intrinsic Hα fluxes, with each voxel in the cube divided by the intrinsic Hα flux. We obtain the median-stacked cube by taking the median value of each voxel element to avoid the effect of outliers and asymmetric distribution of $f_{\rm esc,Ly\alpha}$. The uncertainty of each voxel is estimated by bootstrapping sampling 1000 times.

We extract 1D Lyα spectra using a $r = 2\farcs5$ circular aperture on the median-stacked cubes. This aperture is conservative, larger than what we use for individual HAEs, as we randomly align the subcubes during the stack without considering the asymmetry of Lyα, and thus, the stacked Lyα could be more extended than individuals. On the Lyα pseudo-narrowband

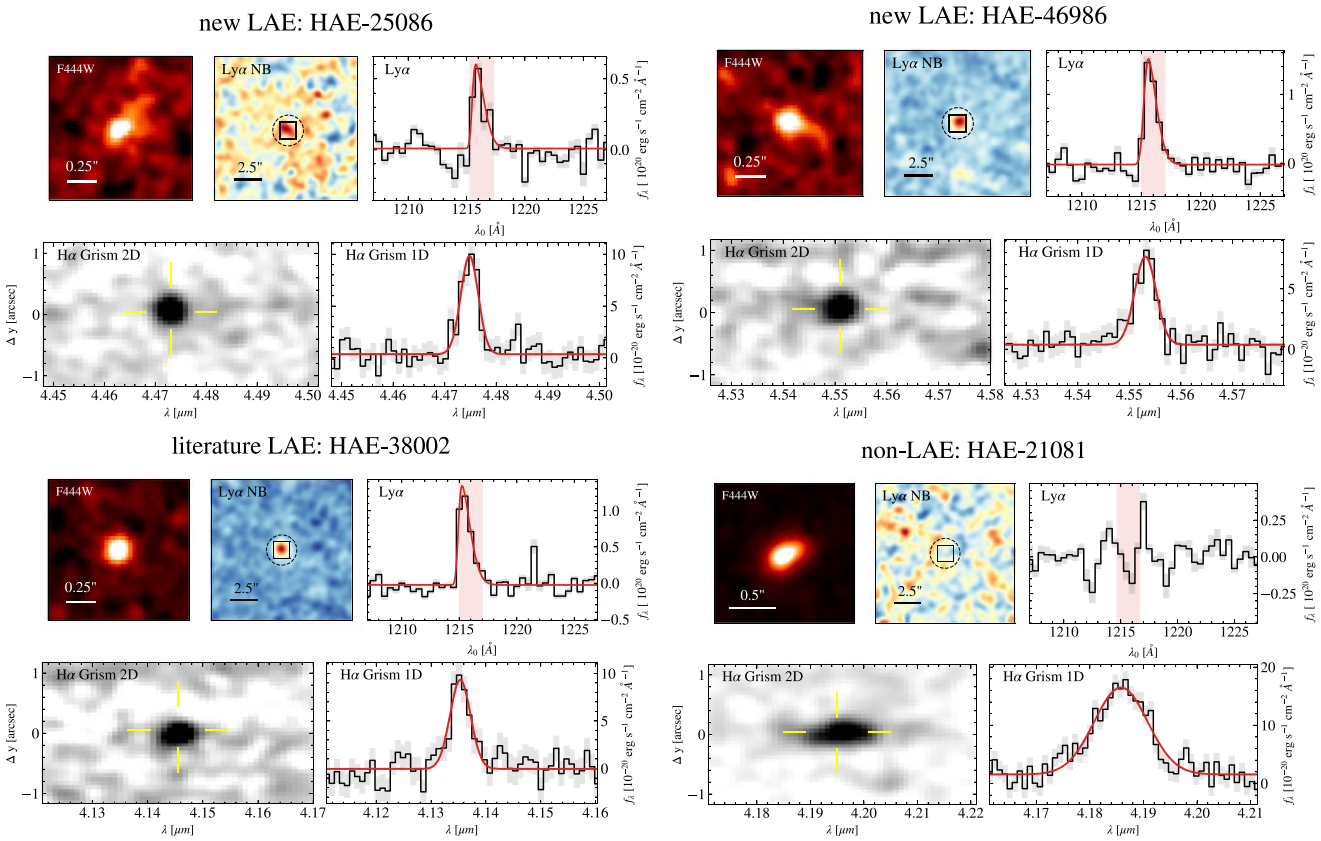

Figure 3. HAE examples with their JWST/NIRcam F444W images (top first), pseudo-Lyα narrowband images (top second), and their Lyα (top third), Hα spectra (bottom). The solid box on the NB image is the 4×4 pixel boxy aperture used to identify Lyα and fit its profile; the dashed circle is the $r = 1\farcs5$ aperture for the total flux. The top third is the boxy-aperture extracted spectrum (black) and its uncertainty (gray shaded). We show the best-fit skewed Gaussian profile in red. The red-shaded region refers to the spectral window (3 pixel width) for Lyα NB. The bottom panels of each example demonstrate the 2D and 1D grism spectrum for Hα. The yellow cross marks the position of Hα emission in the 2D grism spectrum, and the best-fit Gaussian profile is presented in red.

image of the stack of all LAEs integrated over $\pm 1000\,\mathrm{km\,s^{-1}}$, an $r = 2\farcs5$ aperture includes $95\% \pm 12\%$ of the total flux. We fit skewed Gaussian profiles on the extracted 1D spectra to estimate the normalized Lyα flux.

In Figure 4, we show Lyα stacks of all HAEs, LAEs, and non-LAEs. The overall $f_{\mathrm{esc,Ly}\alpha}$ for these HAEs is 0.090 ± 0.006 (0.121 ± 0.009) under the assumption of the Calz (SMC) dust attenuation law. For LAEs, $f_{\mathrm{esc,Ly}\alpha}$ is 0.228 ± 0.012 (0.275 ± 0.013); for non-LAE, $f_{\mathrm{esc,Ly}\alpha}$ is 0.020 ± 0.008 (0.030 ± 0.011). The $f_{\mathrm{esc,Ly}\alpha}$ of LAEs are over twice the average $f_{\mathrm{esc,Ly}\alpha}$ of all HAEs, strongly indicating that previous studies based on Lyα-selected galaxies overestimate the average $f_{\mathrm{esc,Ly}\alpha}$ of the whole galaxy populations.

4. Results

We present the results of our $f_{\mathrm{esc,Ly}\alpha}$ measurements as follows: we describe the distribution of $f_{\mathrm{esc,Ly}\alpha}$ in Section 4.1; the dependence of $f_{\mathrm{esc,Ly}\alpha}$ on various galaxy properties is discussed in Section 4.2; in Section 4.3, we show the redshift evolution of $f_{\mathrm{esc,Ly}\alpha}$.

4.1. The Distribution of $f_{esc,Ly\alpha}$

We measure the $f_{\mathrm{esc,Ly}\alpha}$ of individual HAE following the method in Section 3.3.1. The distributions of $f_{\mathrm{esc,Ly}\alpha}$ of LAEs and non-LAEs are shown in Figure 5. For our detected LAEs, $f_{\mathrm{esc,Ly}\alpha}$ ranges from 0.04 to 1. For non-LAEs, there are negative $f_{\mathrm{esc,Ly}\alpha}$ measurements with large uncertainties; the negative

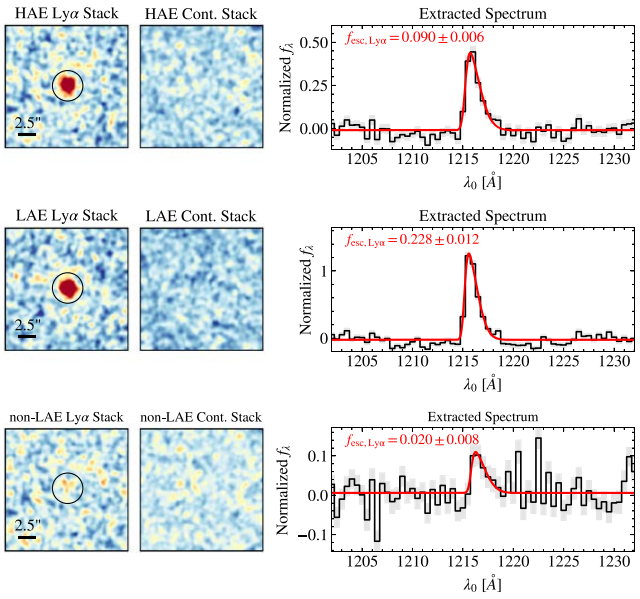

Figure 4. Lyα stacks of MUSE data cubes for all HAEs in our sample, as well as the LAE and non-LAE subsamples. All data cubes are normalized by the intrinsic Hα flux following the Calz dust attenuation law. The psedo-Lyα narrowband (NB) images in the left cover 1213–1219 Å, i.e., slices of ± 3 Å centered on 1216 Å; the continuum (Cont. for short) NB images are calculated from 1230 to 1240 Å slices and scale to the wavelength width of Lyα slices. The extracted spectra based on $r = 2\farcs5$ circular apertures (black solid circle in Lyα NB images) are shown, with the best-fit skewed Gaussian profiles in red.

values are all consistent with zero within 3σ. The mean and median values of the individual $f_{\rm esc,Ly\alpha}$ are 0.13 ± 0.01 and 0.08 ± 0.01 (0.15 ± 0.01 and 0.11 ± 0.01) under the assumption of Calz (SMC) dust attenuation law. We model the distribution of $f_{\rm esc,Ly\alpha}$ using an exponential model, as commonly adopted to depict the distributions of Lyα EWs of Lyα-emitting galaxies (e.g., Dijkstra & Westra 2010; Dijkstra & Wyithe 2012; Kerutt et al. 2022). In fact, $f_{\rm esc,Ly\alpha}$ is proportional to the ratio of Lyα and the Hα fluxes, as shown in Equation (6); meanwhile, Lyα EWs are the ratios between Lyα fluxes and the underlying UV continua. Hα fluxes and UV continua both trace the star formation so that $f_{\rm esc,Ly\alpha}$ and Lyα EWs are correlated (e.g., Yang et al. 2016; Sobral & Matthee 2019). The distribution of $f_{\rm esc,Ly\alpha}$ is expressed as follows:

$$P(f_{\rm esc,Ly\alpha}|f_*)=\begin{cases}0, & f_{\rm esc,Ly\alpha}<0 \text{ or } f_{\rm esc,Ly\alpha}>1,\\ \propto e^{-f_{\rm esc,Ly\alpha}/f_*}, & \text{otherwise,}\end{cases} \quad (7)$$

where $\int_0^1 P(f_{\rm esc,Ly\alpha}|f_*)df_{\rm esc,Ly\alpha}=1$, f_* is the characteristic (mean) escape fraction.

For both LAEs and non-LAEs with measured escape fraction $\hat{f}_{\rm esc,Ly\alpha}$ and uncertainty $\Delta\hat{f}_{\rm esc,Ly\alpha}$, we assume the uncertainty is Gaussian, and the probability for an Lyα escape fraction being $\hat{f}_{\rm esc,Ly\alpha}$ should be

$$\begin{aligned}&\mathcal{P}(\hat{f}_{\rm esc,Ly\alpha},\Delta\hat{f}_{\rm esc,Ly\alpha}|f_*)\\&=\int_0^1 P(x|f_*)P_{\mathcal{N}}(x,\hat{f}_{\rm esc,Ly\alpha},\Delta\hat{f}_{\rm esc,Ly\alpha})dx\end{aligned} \quad (8)$$

where $P_{\mathcal{N}}(x,\hat{f}_{\rm esc,Ly\alpha},\Delta\hat{f}_{\rm esc,Ly\alpha})$ is the probability of the escape fraction being measured as $\hat{f}_{\rm esc,Ly\alpha}$ if the true value is x with the uncertainty $\Delta\hat{f}_{\rm esc,Ly\alpha}$:

$$\begin{aligned}&P_{\mathcal{N}}(x,\hat{f}_{\rm esc,Ly\alpha},\Delta\hat{f}_{\rm esc,Ly\alpha})=\frac{1}{(2\pi)^{1/2}\Delta\hat{f}_{\rm esc,Ly\alpha}}\\&\times\exp\left[-\frac{1}{2}\left(\frac{x-\hat{f}_{\rm esc,Ly\alpha}}{\Delta\hat{f}_{\rm esc,Ly\alpha}}\right)^2\right].\end{aligned}$$

The overall likelihood for our observed $f_{\rm esc,Ly\alpha}$ distribution should be a product of the $f_{\rm esc,Ly\alpha}$ likelihoods of all galaxies. With a flat prior of f_* ranging from 0 to 1, we perform Bayesian inference using the Markov Chain Monte Carlo (MCMC[21]) method to model the observed $f_{\rm esc,Ly\alpha}$ distribution. It yields the best-fit $f_*=0.15\pm0.02$ ($0.17^{+0.03}_{-0.02}$) for Calz (SMC) dust law, in agreement with the mean value of $f_{\rm esc,Ly\alpha}$ (0.13 ± 0.01 for Calz, and 0.15 ± 0.01 for SMC). The median $f_{\rm esc,Ly\alpha}$ value predicted by the best-fit exponential distribution is $f_*\ln 2$, i.e., 0.10 ± 0.01 ($0.12^{+0.02}_{-0.01}$) for Calz (SMC) dust law, in line with the median stack of all HAEs (0.090 ± 0.006 for Calz; 0.121 ± 0.009 for SMC).

4.2. The Dependence of $f_{esc,Ly\alpha}$ on Galaxy Properties

Lyα photon transportation can be modulated by the dust and neutral gas content in galaxies, as well as the geometry and kinematics of ISM (Atek et al. 2009; Smith et al. 2019; Maji et al. 2022). We focus on nine physical properties that are most commonly studied: stellar mass ($M_\star$), ionizing efficiency ($\xi_{\rm ion,0}$), redshift (z), the attenuated (observed) and the intrinsic absolute UV magnitudes ($M_{\rm UV,att}$ and $M_{\rm UV,int}$), specific star formation rate (sSFR), sSFR surface density ($\Sigma_{\rm sSFR}$), the observed UV slope ($\beta_{\rm obs}$), and the dust attenuation ($E(B-V)$). The measurements of these physical parameters are described in Section 3.1. To investigate how $f_{\rm esc,Ly\alpha}$ correlates with the properties of HAEs, we split the HAE sample into five equal-sized bins based on each physical parameter, with 32 or 33 HAEs in each bin. We then estimate the overall $f_{\rm esc,Ly\alpha}$ in each bin by stacking the Lyα. The dependence of $f_{\rm esc,Ly\alpha}$ on the nine physical parameters is presented in Figures 6 and 7. To test whether there are monotonic relations between $f_{\rm esc,Ly\alpha}$ and these physical parameters, we perform the nonparametric Spearman's correlation analysis[22] using a Monte Carlo approach following Curran (2014). We include the forced measured $f_{\rm esc,Ly\alpha}$ of non-LAEs and the corresponding uncertainties in the analysis for unbiased statistics. We also test the correlations with the Kendall rank correlation analysis (Isobe et al. 1986), where the 3σ upper limits of non-LAEs' $f_{\rm esc,Ly\alpha}$, instead of their forced measurements, are used as censored data in the analysis. As shown in Figure 6, $f_{\rm esc,Ly\alpha}$ shows a strong dependence on $\beta_{\rm obs}$ and $E(B-V)$, with a Spearman's correlation coefficient (ρ) of -0.37 ± 0.04 (p-value[23] $=0.17^{+2.21}_{-0.16}\times10^{-5}$) and -0.34 ± 0.05 ($p=0.08^{+1.11}_{-0.07}\times10^{-4}$). The Kendall correlation analysis shows correlation coefficients (τ) of -0.34 ± 0.04 with $p\approx10^{-10}$ and -0.37 ± 0.03 with $p\approx4\times10^{-12}$ for $f_{\rm esc,Ly\alpha}$–$\beta_{\rm obs}$ and $f_{\rm esc,Ly\alpha}$–$E(B-V)$ respectively. We quote the median values of ρ, τ, and p, and their uncertainties correspond to the 16 and 84 percentiles. For $M_\star$, $M_{\rm UV,int}$, $M_{\rm UV,att}$, $\xi_{\rm ion}$, sSFR, and $\Sigma_{\rm sSFR}$, we illustrate their correlations

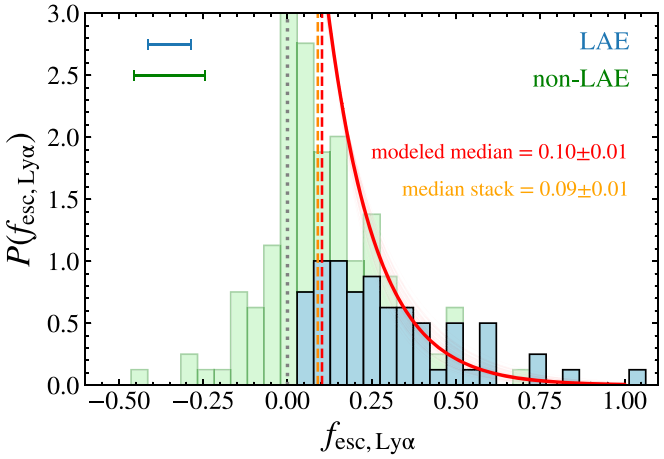

Figure 5. The distribution of $f_{\rm esc,Ly\alpha}$ for LAEs and non-LAEs, assuming the Calz dust attenuation law. For non-LAEs, we show the forced measured $f_{\rm esc,Ly\alpha}$ based on the integrated flux over ±1000 km s^{-1} around the wavelength of Lyα. The mean errorbar of $f_{\rm esc,Ly\alpha}$ of LAEs and non-LAEs are marked as the blue and green horizontal lines. The thick red curve is the best-fit exponential model for the $f_{\rm esc,Ly\alpha}$ distribution; the thin red curves represent 100 exponential models with f_* (Equation (7)) randomly drawn from the MCMC samples. $f_{\rm esc,Ly\alpha}=0$ is labeled as the gray dotted line. The red dashed line denotes the median $f_{\rm esc,Ly\alpha}$ predicted by the best-fit exponential model, and the orange dashed line is the $f_{\rm esc,Ly\alpha}$ measured from the stack of all HAEs.

[21] We use EMCEE (Foreman-Mackey et al. 2013) to perform MCMC: https://emcee.readthedocs.io/en/stable/.

[22] https://github.com/privong/pymccorrelation

[23] p is the p-value indicating the probability of obtaining the current result if the correlation coefficient were zero (no correlation). If p is lower than 0.05, the correlation coefficient is statistically significant.

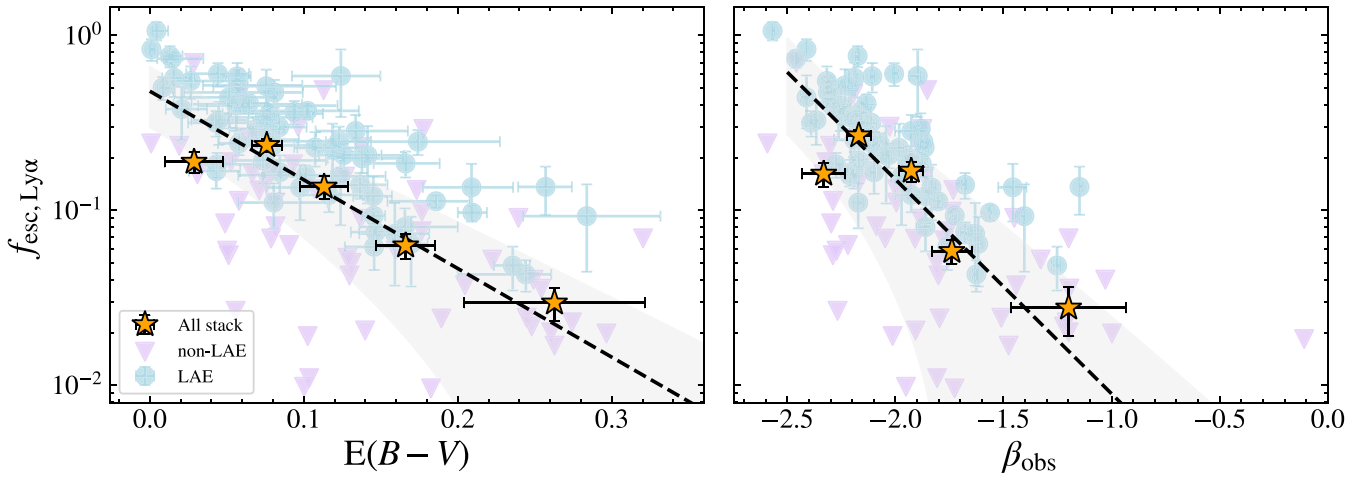

Figure 6. $f_{\rm esc,Ly\alpha}$ varying with the observed UV slope $\beta_{\rm obs}$ and $E(B-V)$, under the assumption of Calz dust attenuation law. The light blue circles denote LAE individuals, and the light purple triangles are non-LAEs with forced $f_{\rm esc,Ly\alpha}$ measured from the integrated Lyα flux within $\pm$1000 km $^{-1}$. For illustration purposes, we do not show the uncertainties of $f_{\rm esc,Ly\alpha}$ of non-LAEs. The orange stars are the overall $f_{\rm esc,Ly\alpha}$ of all HAEs derived from the Lyα stack results. The dashed lines represent the best-fit relations between $f_{\rm esc,Ly\alpha}$ and $E(B-V)$ or $\beta_{\rm obs}$ based on the stack results as described in Equation (9). We show their 1σ uncertainties as shaded regions.

with respect to $f_{\rm esc,Ly\alpha}$ Figure 7, and label the correlation coefficients of Spearman's and Kendall's tests as well as their corresponding p-values. We find that $f_{\rm esc,Ly\alpha}$ also weakly correlates with $M_\star$ and $M_{\rm UV,int}$, with p of both the two correlation tests much smaller than 0.05. For $M_{\rm UV,att}$, the correlation coefficients, ρ and τ, are both close to zero with large p-values, so $f_{\rm esc,Ly\alpha}$ is almost not related to $M_{\rm UV,att}$. In the cases of $\xi_{\rm ion}$, sSFR, and $\Sigma_{\rm sSFR}$, the p-values of Spearman's correlation coefficients can be comparable to or exceed 0.05 within 1σ ranges. The large uncertainties in p are a result of the large uncertainties of the forced measured $f_{\rm esc,Ly\alpha}$ of non-LAEs. On the other hand, their Kendall's correlation analysis reveals τ close to zero with large p-values, indicating no correlations for these quantities.

In conclusion, our results demonstrate that $f_{\rm esc,Ly\alpha}$ has a strong dependence on $\beta_{\rm obs}$ and $E(B-V)$. $f_{\rm esc,Ly\alpha}$ is also monotonically related to $M_\star$ and $M_{\rm UV,int}$, but no correlations can be confirmed between $f_{\rm esc,Ly\alpha}$ and $M_{\rm UV,att}$, $\xi_{\rm ion}$, logsSFR and $\Sigma_{\rm sSFR}$.

In fact, $E(B-V)$ and $\beta_{\rm obs}$ have long been thought to be the most crucial factors to trace $f_{\rm esc,Ly\alpha}$ from low redshifts (e.g., Atek et al. 2009; Yang et al. 2016) to cosmic noon (e.g., Hayes et al. 2010; Sobral & Matthee 2019). We present their relationship with respect to $f_{\rm esc,Ly\alpha}$, for the first time, at redshift $z>5$. As shown in Figure 6, $f_{\rm esc,Ly\alpha}$ appears remarkably anticorrelated with $E(B-V)$ and $\beta_{\rm obs}$. We fit the relation using the orthogonal distance regression method[24] with the stack $f_{\rm esc,Ly\alpha}$ in Figure 6, assuming

$$f_{\rm esc,Ly\alpha} = \begin{cases} f_{\rm Ly\alpha,0} \cdot 10^{-0.4 k_{\rm Ly\alpha} E(B-V)}, \\ f'_{\rm Ly\alpha,0} \cdot 10^{-\alpha_{\rm Ly\alpha}(\beta_{\rm obs}+2.5)}. \end{cases} \quad (9)$$

For Calz (SMC) law, we obtain $f_{\rm Ly\alpha,0} = 0.48 \pm 0.18$ (0.46 ± 0.19), $k_{\rm Ly\alpha} = 12.68 \pm 4.07$ (15.21 ± 5.73). For the $f_{\rm esc,Ly\alpha}$–$\beta_{\rm obs}$ relation, we obtain the best-fit $f'_{\rm Ly\alpha,0} = 0.62 \pm 0.34$ (0.54 ± 0.19), and $\alpha_{\rm Ly\alpha} = 1.23 \pm 0.48$ (0.71 ± 0.21). This indicates that dust-free galaxies at $z = 5$–6, on average, exhibit $f_{\rm esc,Ly\alpha}$ of 0.48 ± 0.18 (0.46 ± 0.19); galaxies with blue UV slopes $\beta \sim -2.5$ have $f_{\rm esc,Ly\alpha}$ around 0.62 ± 0.34 (0.54 ± 0.19).

Among the 165 HAEs, we note that two LAEs (HAE-60916 and HAE-87611) exhibit double-peaked Lyα profiles (Figure D1; see Appendix D for more details). They have blue UV slopes, $\beta_{\rm obs} \approx -2.107 \pm 0.007$ and -1.871 ± 0.023, and high Lyα escape fractions, $f_{\rm esc,Ly\alpha} \approx 0.58 \pm 0.11$ and 0.25 ± 0.04. The predicted $f_{\rm esc,Ly\alpha}$ based on their $\beta_{\rm obs}$ using Equation (9) is 0.20 ± 0.14 and 0.10 ± 0.09 respectively. The observed $f_{\rm esc,Ly\alpha}$ of the two double-peaked LAEs is $\sim 2\sigma$ higher than the predicted values. The presence of blue Lyα peaks is likely to signal a boosted escape of Lyα photons compared to normal star-forming galaxies, consistent with previous findings on double-peaked LAEs at the cosmic noon (Furtak et al. 2022; Marques-Chaves et al. 2022). Studies on statistical samples with high S/N Lyα profiles are expected in the future using deep spectroscopy.

4.3. Redshift Evolution of $f_{esc,Ly\alpha}$

We study the evolution of $f_{\rm esc,Ly\alpha}$ as a function of redshift as shown in Figure 8. We split the HAE sample into two redshift bins, $z = 5.24$ ($z = 4$, 9–5.5) and $z = 5.84$ ($z = 5.5$–6.2), and then measured the stacked Lyα respectively. The observed $f_{\rm esc,Ly\alpha}$ is 0.104 ± 0.007 at $z = 5.24$ and becomes 0.086 ± 0.010 at $z = 5.84$. We compare our measurements with the $f_{\rm esc,Ly\alpha}$ measurements in the literature. The observed $f_{\rm esc,Ly\alpha}$ values are only about 56% and 37% of the estimates in Hayes et al. (2011) at $z = 5.4$ and $z = 5.8$ respectively. On the other hand, our measurements are generally consistent with the best-fit function in Konno et al. (2016; green dashed line in Figure 8), although they are still lower than their measurement at this redshift range (green circle in Figure 8). We note that, because Hα was not accessible at $z = 5$–6 until the launch of JWST, all previous measurements rely on UV LFs to estimate the SFR density and Hα luminosity density at these redshifts, while our results present a direct measurement for the first time. The offset between Hayes et al. (2011) and Konno et al. (2016) is due to the differences of the Lyα and UV luminosity limits for deriving the Lyα and UV luminosity densities from their LFs: Hayes et al. (2011) integrated Lyα LFs from $L_{\rm Ly\alpha} = 10^{41.3}$ erg s^{-1} to

[24] https://docs.scipy.org/doc/scipy/reference/odr.html

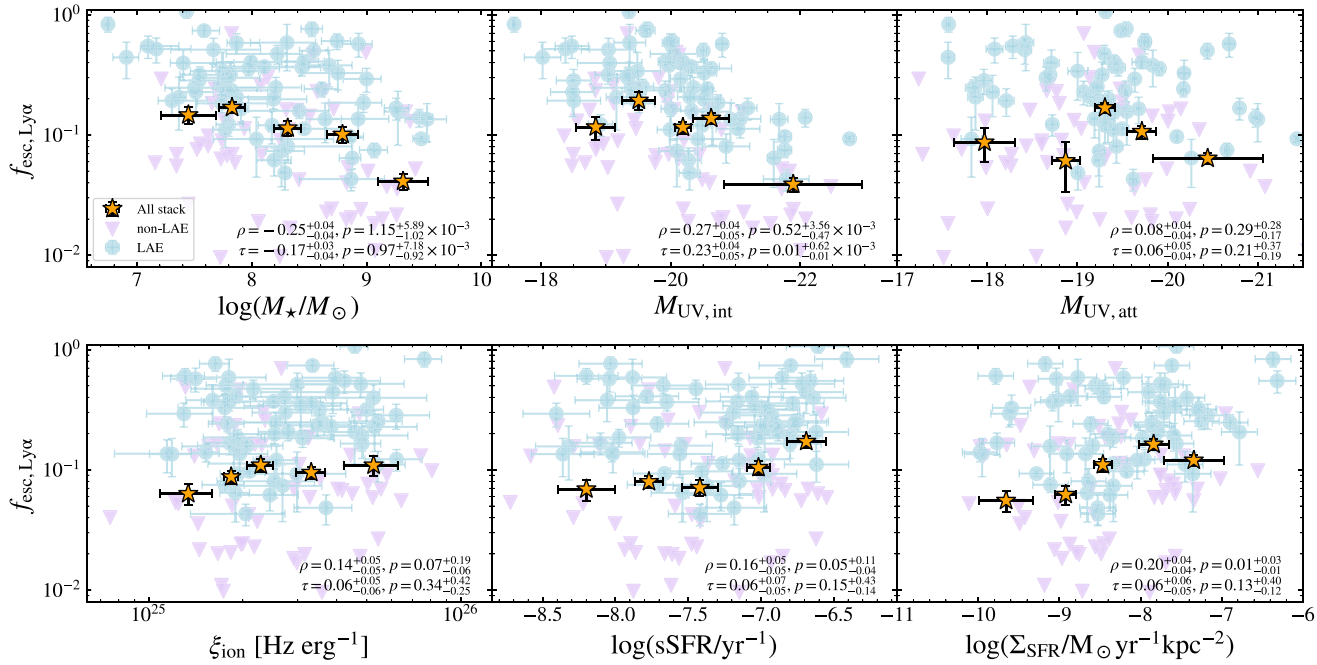

Figure 7. $f_{\rm esc,Ly\alpha}$ varying with different physical parameters of HAEs assuming the Calz dust attenuation law. The light blue dots denote the $f_{\rm esc,Ly\alpha}$ measured for individual LAE; the light purple triangles are the forced measured $f_{\rm esc,Ly\alpha}$ for non-LAEs. The uncertainties of $f_{\rm esc,Ly\alpha}$ of non-LAEs are not shown for illustration purposes. The orange stars represent the stack $f_{\rm esc,Ly\alpha}$. We label the correlation coefficients of the Spearman's correlation analysis (ρ) and Kendall's correlation analysis (τ) for each property, as well as the corresponding p-values. The Spearman's correlation analysis is based on the measured values and uncertainties of all individual LAEs and non-LAEs, while in the Kendall's tests we adopt the 3σ upper limits of $f_{\rm esc,Ly\alpha}$ for non-LAEs.

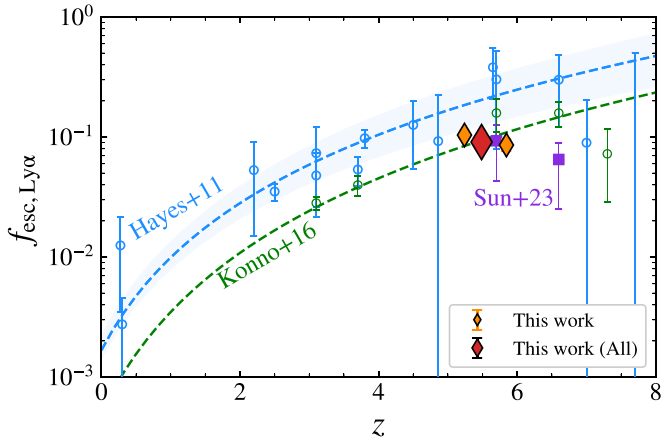

Figure 8. The redshift evolution of global $f_{\rm esc,Ly\alpha}$ from $z=0$ to $z=8$. Our stacked measurement of $f_{\rm esc,Ly\alpha}$ for all HAE samples is shown as a red diamond based on the assumption of Calz dust law. The orange and white diamonds are $f_{\rm esc,Ly\alpha}$ measured by stacking the HAE sample in two redshift bins ($z=5.24$, 5.84). $f_{\rm esc,Ly\alpha}$ measurements compiled by Hayes et al. (2011) based on Lyα and UV ($z>2.2$) or Hα luminosity densities ($z<2.2$) are shown in blue circles, and the best-fit relation for their compilation is shown as the blue dashed line. The green circles denote the measurements in Konno et al. (2016), also using Lyα and UV LFs, and their best-fit relation is shown as the green dashed line. The purple squares are the $f_{\rm esc,Ly\alpha}$ derived from Lyα and Hα luminosity densities at $z>5$ in Sun et al. (2023) with HAEs observed by NIRCam/WFSS.

$L_{\rm Ly\alpha}=10^{42.9}$ erg s^{-1} and UV LFs down to -17.6, while Konno et al. (2016) integrated LyαLFs from $L_{\rm Ly\alpha}=10^{41.7}$ erg s^{-1} to $L_{\rm Ly\alpha}=10^{44.4}$ erg s^{-1} and UV LFs down to -17.0 mag. In addition to this, the measurements of faint-end LFs and the conversion between UV and Hα SFR could lead to large systematic uncertainties in these UV-based $f_{\rm esc,Ly\alpha}$ estimates (Konno et al. 2016; Sun et al. 2023). The observed $f_{\rm esc,Ly\alpha}$ are in good agreement with Sun et al. (2023), where $f_{\rm esc,Ly\alpha}$ is derived from Hα luminosity densities of HAEs also captured by JWST/NIRCam WFSS. Our results, combined with those in Sun et al. (2023), suggest a tentative declining trend for the observed $f_{\rm esc,Ly\alpha}$ at $z\gtrsim5$. The average $E(B-V)$ of our HAE sample is about 0.14 at $z=5.24$ and 0.09 at $z=5.84$, implying an intrinsically higher $f_{\rm esc,Ly\alpha}$ at the higher redshift due to their less dusty content. The decline of $f_{\rm esc,Ly\alpha}$ might be a result of the increasing neutral fraction of IGM as z increases. To completely understand the redshift evolution of $f_{\rm esc,Ly\alpha}$, direct observations on Lyα and Hα emission lines across wide redshift spans are required in the future.

5. Discussion

In this section, we discuss the physical origins of the correlations as outlined in Section 4, connect the escape of Lyα and LyC photons in Section 5.1, and further qualitatively constrain ionizing photon budget for the reionziation in Section 5.2.

5.1. Implications for Lyα and LyC Escape

As illustrated in Section 4.2, $f_{\rm esc,Ly\alpha}$ correlates with $M_\star$ and $M_{\rm UV,int}$, albeit to a lesser extent compared to the strong dependence of $f_{\rm esc,Ly\alpha}$ on $\beta_{\rm obs}$ and $E(B-V)$. However, $M_\star$ and $M_{\rm UV,int}$ may be highly coupled with $\beta_{\rm obs}$ and $E(B-V)$. $M_{\rm UV,int}$ are the products of observed $M_{\rm UV,att}$ corrected by the effect of dust attenuation, so it is highly dependent on the dust attenuation parameter adopted. We explore the coupling of $f_{\rm esc,Ly\alpha}$–$M_\star$ and $f_{\rm esc,Ly\alpha}$–$M_{\rm UV,int}$ relations with $\beta_{\rm obs}$ and $E(B-V)$ in Figure 9. Comparing the expected $f_{\rm esc,Ly\alpha}$ by Equation (9) based on the median $\beta_{\rm obs}$ and $E(B-V)$ in each $M_\star$ and $M_{\rm UV,int}$ bin, we find that the variation of $f_{\rm esc,Ly\alpha}$ across $M_\star$ and $M_{\rm UV,int}$ is consistent with the prediction of Equation(9). It implies that both the $f_{\rm esc,Ly\alpha}$–$M_\star$ and $f_{\rm esc,Ly\alpha}$–$M_{\rm UV,int}$ correlations are modulated by

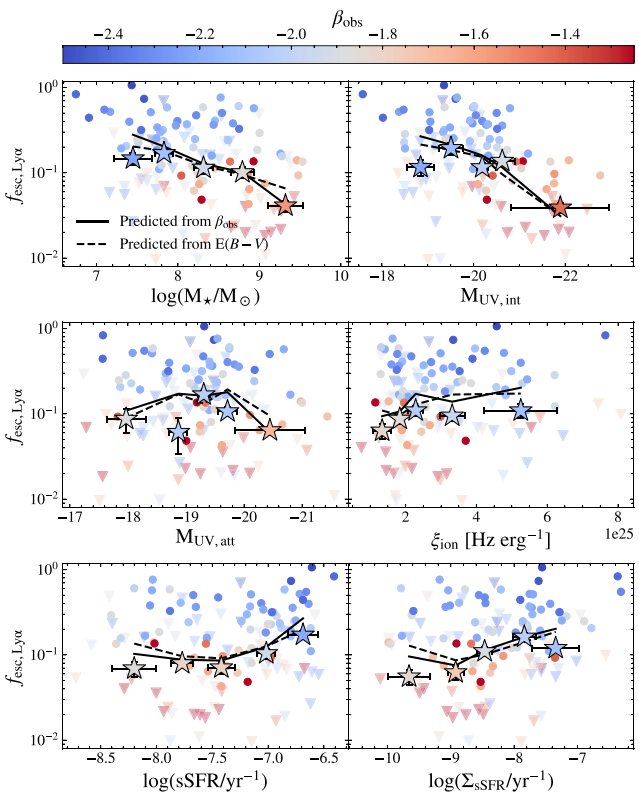

Figure 9. The coupling of β_{obs} and $E(B-V)$ on the dependence of $f_{esc,Ly\alpha}$ on $M_\star$, $M_{UV,att}$, $M_{UV,int}$, ξ_{ion}, sSFR, and Σ_{sSFR}. We show the $f_{esc,Ly\alpha}$ of individual LAEs as circles and forced measured $f_{esc,Ly\alpha}$ of non-LAEs as triangles, color-coded by their β_{obs}. The stack $f_{esc,Ly\alpha}$ are marked as stars color-coded by the median β_{obs} in each stack bin. The solid and dashed lines are predicted $f_{esc,Ly\alpha}$ from Equation (9) based on the median β_{obs} and $E(B-V)$ in each stack bin, respectively.

the underlying coupling among β_{obs}, $E(B-V)$ and $M_\star$, $M_{UV,int}$. Likewise, we examine the variations of $f_{esc,Ly\alpha}$ versus $M_{UV,att}$, sSFR, Σ_{sSFR}, and ξ_{ion}, and find that the trends are determined by the variations of median β_{obs} and $E(B-V)$ in each stack bin. As illustrated in Figure 9, we can reproduce the trend in $f_{esc,Ly\alpha}$ versus these properties by applying the median values of β_{obs} and $E(B-V)$ to Equation (9). To disentangle their impacts on $f_{esc,Ly\alpha}$, it is necessary to track the evolution of $f_{esc,Ly\alpha}$ within a fixed $M_\star$ or $M_{UV,int}$ bin with varying β_{obs} and $E(B-V)$. This requires a large sample with direct and accurate Lyα and Hα detection, which cannot be achieved with our current sample.

In fact, the dependencies on β_{obs} and $E(B-V)$ are also found in $f_{esc,LyC}$ of low-redshift LyC leaking galaxies. With a compilation of LyC leakers at $z\sim0.3$, Chisholm et al. (2022) found an analytical relation between $f_{esc,LyC}$ and β_{obs}: $f_{esc,LyC} = (1.3\pm0.6)\times10^{-4}\times10^{(-1.22\pm0.1)\beta_{obs}}$. The exponential index (−1.22) is in good agreement with that of our best-fit $f_{esc,Ly\alpha}$–β_{obs} relation (-1.23) in Equation (9). Note that we only compare the Chisholm et al. (2022) $f_{esc,LyC}$–β_{obs} relation with our Calz dust law-based $f_{esc,Ly\alpha}$–β_{obs} relation, because the β_{obs} in Chisholm et al. (2022) are measured with the modeled stellar continuum assuming the Reddy et al. (2016) attenuation law similar in shape to the Calz law. The modeled β_{obs} is sensitive to the effect of dust attenuation. Furthermore, Chisholm et al. (2022) models the SED of local LyC leakers using STARBURST99 (Leitherer et al. 1999), while we perform SED modeling using BEAGLE, which is based on Bruzual & Charlot (2003); the different stellar continuum models also introduce systematics. Nevertheless, the similarity in correlation trends in $f_{esc,Ly\alpha}$–β_{obs} and $f_{esc,LyC}$–β_{obs} is consistent with a connection between the escape processes of Lyα and LyC photons. Their escape processes are sensitive to the opacity and geometry of neutral HI gas in both the vicinity of star-forming regions and the circumgalactic medium (Dijkstra et al. 2016; Byrohl & Gronke 2020; Kimm et al. 2022; Maji et al. 2022). Recently, Begley et al. (2023) investigated the LyC–Lyα connection at $z\approx4$–5 using $f_{esc,LyC}$ predicted by the EWs of low-ionization, far-UV absorption lines in star-forming galaxies (Garilli et al. 2021) and IRAC-measured Hα. They derived a monotonic relation, $f_{esc,LyC}\simeq0.15^{+0.06}_{-0.04}f_{esc,Ly\alpha}$. We scale our measured $f_{esc,Ly\alpha}$ and $f_{esc,Ly\alpha}$–β_{obs} relation by 0.15 in Figure 10. The scaled $f_{esc,Ly\alpha}$–β_{obs} relation is well consistent with the trend of LyC leakers within 1σ. Likewise, we estimate $f_{esc,Ly\alpha}\xi_{ion}$ as a function of β_{obs} using the median stack $f_{esc,Ly\alpha}$ and the median values ofξ_{ion} in each stack bin, and then scale $f_{esc,Ly\alpha}\xi_{ion}$ by 0.15 as a proxy to $f_{esc,LyC}\xi_{ion}$, the production rate of ionizing photons that can escape to and ionize the IGM. As shown in Figure 10, the scaled $f_{esc,Ly\alpha}\xi_{ion}$–$\beta_{obs}$ relation is also in good agreement with the $f_{esc,LyC}\xi_{ion}$–β_{obs} relation in Chisholm et al. (2022). It implies that $f_{esc,Ly\alpha}\xi_{ion}$ can be used to qualitatively estimate the total emitted ionizing emissivity that impacts the IGM (see Section 5.2 for more details). Despite systematics, such as bias in sample selection and the slight difference in Lyα transmission at $z=4$–5 and $z=5$–6, the consistency of the scaled $f_{esc,Ly\alpha}$ and $f_{esc,LyC}$ suggests that $f_{esc,Ly\alpha}$ is likely to be proportional to $f_{esc,LyC}$. With this scaling relation, the overall $f_{esc,LyC}$ for our HAEs are $\sim$0.015 ($f_{esc,Ly\alpha}\approx$0.1) and thus negligible for the estimation of ξ_{ion} in Equation (4).

The link between Lyα and LyC escape is also in line with the correlations we observe between $f_{esc,Ly\alpha}$ and galaxy properties. If the physical conditions that favor LyC photon leakage also facilitate the escape of Lyα, observables that can indicate $f_{esc,LyC}$ are expected to show strong correlations with $f_{esc,Ly\alpha}$. From cosmological radiation hydrodynamics simulations, Choustikov et al. (2023) concludes that a good diagnostic of $f_{esc,LyC}$ should simultaneously encapsulate at least two of the following three conditions: (1) high sSFR for the production of ionizing photons, (2) the age of stellar population and timescale for stellar feedback, and (3) a proxy for neutral gas content and the ionization state of ISM. In this context, β_{obs} and $E(B-V)$ serve as strong indicators for the escape of LyC and Lyα photons. $E(B-V)$ directly tracks the dust content along the line of sight and may also trace neutral hydrogen (e.g., Katz et al. 2022); the production and destruction of dust are also related to supernovae activities (e.g., Priestley et al. 2021). Galaxies with bluer β_{obs} are biased toward having higher sSFRs and younger stellar populations; β_{obs} is also very sensitive to dust attenuation. $E(B-V)$ and β_{obs} satisfy at least two of the three criteria and thus are good probes of $f_{esc,Ly\alpha}$. The three physical conditions discussed above also help us to understand the relatively poor dependence of $f_{esc,Ly\alpha}$ on sSFR, ξ_{ion}, Σ_{sSFR}, and $M_{UV,att}$. sSFR or ξ_{ion} represents just one facet of the three necessary conditions so that a high sSFR or ξ_{ion} only could not guarantee a high $f_{esc,Ly\alpha}$. Σ_{sSFR} alone cannot be used to reveal feedback processes or the state of the ISM along the sightline. $M_{UV,att}$ is a complex $M_\star$, dust, SFR, stellar population, etc., so the $f_{esc,Ly\alpha}$–$M_{UV,att}$ may suffer from large scatters introduced by the coupling of $M_{UV,att}$ with respect to these parameters. Also, $M_{UV,att}$ does not trace the

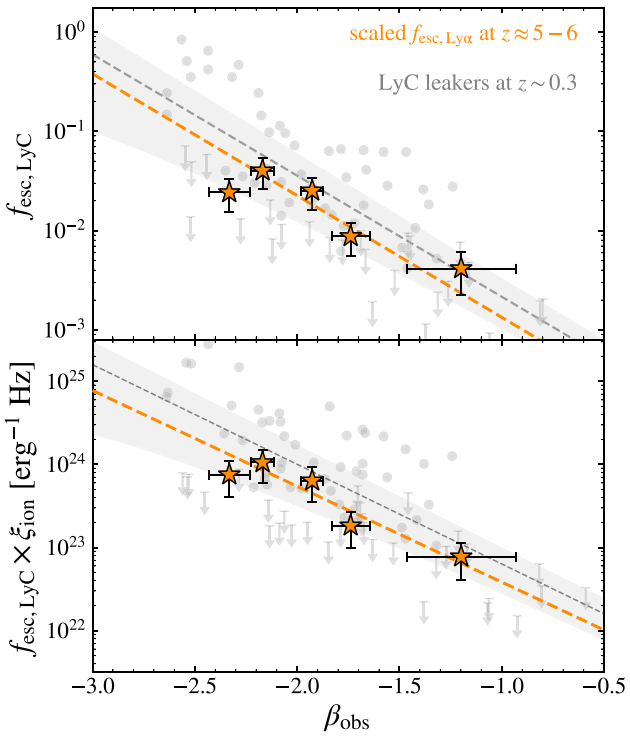

Figure 10. The relation between $\beta_{\rm obs}$ and $f_{\rm esc,LyC}$, $f_{\rm esc,LyC}\xi_{\rm ion}$. The gray circles represent LyC leakers at $z\sim0.3$ from Chisholm et al. (2022), where $\beta_{\rm obs}$ are measured as observed stellar continuum slope at 1550 Å. The gray downward arrows refer to the 1σ upper limits of $f_{\rm esc,LyC}$ for galaxies without LyC detected, i.e., LyC flux is not significantly detected given the observed backgrounds. The gray dashed line and the shaded region are the analytical relations of $f_{\rm esc,LyC}$ and $f_{\rm esc,LyC}\xi_{\rm ion}$ with respect to $\beta_{\rm obs}$, and the 1σ uncertainties of the relations. The orange stars are our stack $f_{\rm esc,Ly\alpha}$ measurements assuming the Calz dust law, scaled by a factor of 0.15 according to Begley et al. (2023), and the orange dashed lines are the best-fit $f_{\rm esc,Ly\alpha}$–$\beta_{\rm obs}$ and $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$–$\beta_{\rm obs}$ relation as described in Equation (9), also scaled by 0.15.

ISM state along the sightline. Although Choustikov et al. (2023) found $M_{\rm UV,att}$ can be a weak tracer of sSFR and dust attenuation at fixed $M_\star$, in which case $M_{\rm UV,att}$ satisfies two out of three conditions while the stellar age would still introduce some scatters, our sample size is not large enough to disentangle on the $M_{\rm UV,att}$ dependence of $f_{\rm esc,Ly\alpha}$ at a fixed $M_\star$.

5.2. Implications for Cosmic Reionization

There is an increasing body of evidences showing that dwarf galaxies are the main drivers of reionization (Atek et al. 2024; Mascia et al. 2023; Simmonds et al. 2023). However, the relative importance of massive and low-mass galaxies in driving reionization remains uncertain (Robertson et al. 2015; Finkelstein et al. 2019; Naidu et al. 2020). It is necessary to quantify the relative contribution of galaxies with different UV magnitude or masses to the total ionizing emissivity, i.e., the number of ionizing photons emitted per unit time and comoving volume. $\dot{n}_{\rm ion}$ contributed by a specific population can be cast as a product of the total observed UV luminosity density $\rho_{\rm UV,obs}$, the production efficiency $\xi_{\rm ion}$, and the escape fraction of ionizing photons $f_{\rm esc,LyC}$:

$$\dot{n}_{\rm ion} = \rho_{\rm UV,\,obs}\times\xi_{\rm ion}\times f^{\rm rel}_{\rm esc,LyC} \quad (10)$$

where $f^{\rm rel}_{\rm esc,LyC}$ is the escape fraction of LyC photons relative to the UV photons, defined as $f^{\rm rel}_{\rm esc,LyC}=f_{\rm esc,LyC}/f_{\rm UV}=f_{\rm esc,LyC}\,10^{0.4A_{\rm UV}}$ (Bouwens et al. 2016; Steidel et al. 2018; Bian & Fan 2020). We adopt $A_{\rm UV}$ from the best-fit SED models at the rest-frame 1500 Å. We emphasize that the definitions of ionizing efficiency and escape fraction of ionizing photons may vary among literature (Bouwens et al. 2016; Steidel et al. 2018; Tang et al. 2019; Simmonds et al. 2023), which may cause confusion. In this section, we follow the definition in Bouwens et al. (2016). We define $\xi_{\rm ion}$ as the ionizing production rate per unit *intrinsic* UV luminosity as described in Section 3.1. We adopt the relative escape fraction, $f^{\rm rel}_{\rm esc,LyC}$, so that the product $f^{\rm rel}_{\rm esc,LyC}\xi_{\rm ion}$ represents the production of ionizing photons that escape to and ionize the IGM per *observed* UV luminosity density. Likewise, we define the relative escape fraction of Lyα photons as $f^{\rm rel}_{\rm esc,Ly\alpha}=f_{\rm esc,Ly\alpha}/f_{\rm UV}=f_{\rm esc,Ly\alpha}\,10^{0.4A_{\rm UV}}$.

If $f_{\rm esc,Ly\alpha}$ is proportional to $f_{\rm esc,LyC}$, as discussed in Section 5.1, $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ is proportional to $f^{\rm rel}_{\rm esc,LyC}\xi_{\rm ion}$. We investigate the dependence of $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ on $M_{\rm UV,att}$ and $M_\star$ in Figure 11. We fit $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ as a function of $M_{\rm UV,att}$ and $M_\star$, yielding the following:

$$\log f^{\rm rel}_{\rm esc,Ly\alpha}\,\xi_{\rm ion}=(0.23\pm0.10)\times(M_{\rm UV,att}+19.5)+(24.84\pm0.08) \quad (11)$$

$$\log f^{\rm rel}_{\rm esc,Ly\alpha}\,\xi_{\rm ion}=(-0.33\pm0.04)\times\log(M_\star/10^{8.5}M_\odot)+(24.92\pm0.03). \quad (12)$$

To provide a reference for future works that may define the ionizing efficiency and escape fraction differently, we also present the analytical relations of $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ with respect to $M_{\rm UV,att}$ and $M_\star$ in Appendix C, which can be proportional to the ionizing emissivity per unit *intrinsic* UV luminosity $f_{\rm esc,LyC}\xi_{\rm ion}$.

The contribution to the total ionizing photon budget from galaxies fainter than $M_{\rm UV}$ can be expressed as follows:

$$\dot{n}_{\rm ion}(>M_{\rm UV})\propto\int_{M_{\rm UV}}^{M_{\rm lim}} f^{\rm rel}_{\rm esc,Ly\alpha}\,\xi_{\rm ion}(M)\times\Phi_{\rm UV}(M)L_{\rm UV}(M)\,dM \quad (13)$$

where $M_{\rm lim}$ is the faint-end limit to compute the luminosity density by integrating over the UV LFs. We adopt the UV LF with a Schechter functional form and redshift-dependent Schechter parameters in Bouwens et al. (2022). At $z=5.5$, we have the characteristic luminosity $M^*=-21.02\pm0.04$, the normalization $\Phi^*=0.56\pm0.04$, and the faint-end slope $\alpha=-1.90\pm0.02$. Although there are no direct observations on $f_{\rm esc,Ly\alpha}$ of faint galaxies ($M_{\rm UV,att}>-17$ mag), we can give a qualitative constraint on the contribution of UV-faint galaxies by extrapolating the $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$–$M_{\rm UV,att}$ analytical relation to the faint-end. With $M_{\rm lim}=-13$ (Bouwens et al. 2016) and Equation (11), we find that galaxies fainter than −16 mag will contribute about 72% of the total ionizing budget, i.e., the accumulative emissivity fraction $\frac{\dot{n}_{\rm ion}(>-16\ {\rm mag})}{\dot{n}_{\rm ion,\,tot}}\approx0.76$ where $\dot{n}_{\rm ion,\,tot}$ is estimated by integrating from −24 mag to $M_{\rm lim}$. We present the accumulative emissivity fraction as a function of $M_{\rm UV,att}$ in Figure 12. The fraction becomes 94% if $M_{\rm lim}=-10$.

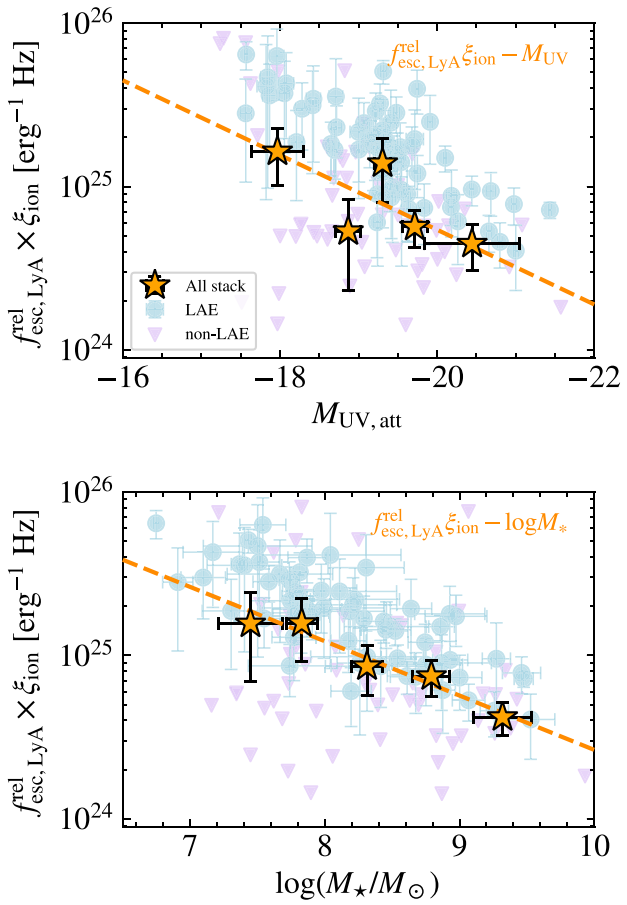

Figure 11. $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ as a function of $M_{\rm UV,att}$, and $M_\star$ assuming the Calz dust law. The blue circles indicate $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ for each individual LAEs, and purple triangles are $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ of non-LAEs with forced measured $f^{\rm rel}_{\rm esc,Ly\alpha}$. The fringed orange stars indicate the stack $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ for all HAEs. We adopt the median stack $f_{\rm esc,Ly\alpha}$ values and the median $\xi_{\rm ion}$ and $A_{\rm UV}$ of HAEs in each stack bin. We show the best-fit relations of the stack $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ of all HAEs with respect to $M_{\rm UV,att}$, $M_\star$ as the orange lines.

Despite large uncertainties due to the lack of observations of $f_{\rm esc,Ly\alpha}$ on the faint-end, our result suggests that UV-faint galaxies are the main contributor to the total ionizing emissivity, at least at $z=5$–6. If the $f_{\rm esc,Ly\alpha}$–$f_{\rm esc,LyC}$ connection and $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ dependence on $M_{\rm UV,att}$ hold at the EoR, UV-faint galaxies will provide the majority of ionizing photons to reionize the Universe. These faint young galaxies may have an elevated ionizing photon production efficiency (Maseda et al. 2020). Further observations on $f_{\rm esc,Ly\alpha}$ of faint galaxies ($M_{\rm UV,att}>-17$ mag) are required to quantitatively constrain the ionizing emissivity from them. Moreover, to reveal the role of galaxies with different stellar masses during the reionization, it is necessary to measure the stellar mass functions at high redshifts, and observe Lyα emission of low-mass galaxies that extend Equation (12) to the lower-mass end.

6. Summary

We investigate the escape of Lyα with a large sample of 165 HAEs at $z\approx4.9$–6.3 in the GOODS-S field, which lies in the overlap region of the JWST FRESCO program and the public MUSE surveys. The JWST/NIRCam WFSS observations provided by FRESCO allow us to construct an unbiased

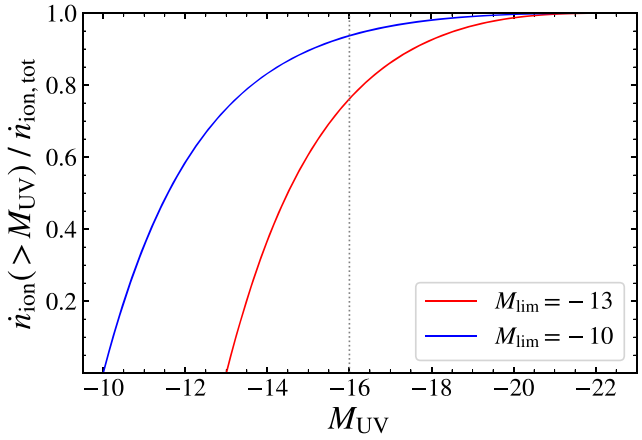

Figure 12. The accumulative emissivity of galaxies at $z=5.5$. Here, $\frac{\dot{n}_{\rm ion}(>M_{\rm UV})}{\dot{n}_{\rm ion,\,tot}}$ refers to the fraction of ionizing budget contributed by galaxies fainter than a specific $M_{\rm UV,att}$, as in Equation (13). We set the faint-end limit $M_{\rm lim}=-13$ and -10 for $\dot{n}_{\rm ion,tot}$, shown as the red and blue lines respectively. We mark $M_{\rm UV,att}=-16$ as the vertical dotted line for reference.

HAE sample with precise measurements of their Hα emission lines. We search for their Lyα emission in the 3D data cubes of VLT/MUSE. We measure the Lyα escape fraction ($f_{\rm esc,Ly\alpha}$) of individual galaxies and perform a stacking analysis for the overall populations. Our main conclusions are as follows:

1. The overall $f_{\rm esc,Ly\alpha}$ of HAEs at $z\approx5$–6 is 0.090 ± 0.006 (0.121 ± 0.009) under the assumption of Calz (SMC) dust attenuation law. The distribution of $f_{\rm esc,Ly\alpha}$ can be described using an exponential model with a median escape fraction of 0.10 ± 0.01 ($0.12^{+0.02}_{-0.01}$) for the Calz (SMC) dust law.
2. $f_{\rm esc,Ly\alpha}$ exhibits strong anticorrelations with the dust attenuation $E(B-V)$ and the observed UV slope ($\beta_{\rm obs}$), indicative of the crucial role that dust plays in modulating the escape of Lyα photons. Adopting the Calz (SMC) dust law, our best-fit relations between $f_{\rm esc,Ly\alpha}$ and $E(B-V)$ and $\beta_{\rm obs}$ predict $f_{\rm esc,Ly\alpha}\approx0.48$ (0.46) for dust-free galaxies and $f_{\rm esc,Ly\alpha}\approx0.62$ (0.54) for young galaxies with very blue UV slopes $\beta_{\rm obs}\approx-2.5$.
3. We find that $f_{\rm esc,Ly\alpha}$ also correlates with the stellar mass ($M_\star$) and the intrinsic UV magnitude ($M_{\rm UV,int}$). The correlations are regulated by the underlying coupling between $M_\star$, $M_{\rm UV,int}$ and $\beta_{\rm obs}$, $E(B-V)$. We do not find strong correlations between $f_{\rm esc,Ly\alpha}$ and the ionizing efficiency ($\xi_{\rm ion}$), the observed UV magnitude ($M_{\rm UV,att}$), the specific star formation rate (sSFR), and the sSFR surface density ($\Sigma_{\rm sSFR}$). The variations of $f_{\rm esc,Ly\alpha}$ are also regulated by the scatterings of $\beta_{\rm obs}$ and $E(B-V)$ coupled with these parameters.
4. We find a tentative decrease in $f_{\rm esc,Ly\alpha}$ as z grows at $z\gtrsim5$ combined with measurements in Sun et al. (2023). It implies an increasing neutral fraction of IGM at this epoch. At $z=5.2$ and $z=5.8$, our measured $f_{\rm esc,Ly\alpha}$ is only 56% and 37% of the previous estimates based on Lyα and UV LFs (Hayes et al. 2011). The difference could arise from the uncertainties in estimating the Lyα LFs and the conversion between UV and Hα luminosities in previous studies without direct Hα constraints.
5. The $f_{\rm esc,Ly\alpha}$–$\beta_{\rm obs}$ and $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$–$\beta_{\rm obs}$ relations have log-space slopes very close to those of the $f_{\rm esc,LyC}$–$\beta_{\rm obs}$ and

$f_{\rm esc,LyC}\xi_{\rm ion}$–$\beta_{\rm obs}$ relations measured from LyC leakers at $z\sim0.3$ (Chisholm et al. 2022). The scaling factor is consistent with the monotonic $f_{\rm esc,Ly\alpha}$–$f_{\rm esc,LyC}$ connection at $z=4$–5 (Begley et al. 2023), implying that $f_{\rm esc,Ly\alpha}$ can be proportional to $f_{\rm esc,LyC}$. This may indicate that the escapes of Lyα and LyC are regulated by similar physical processes. It is also in line with the weak dependence of $f_{\rm esc,Ly\alpha}$ on $M_{\rm UV,att}$, $\xi_{\rm ion}$, sSFR, and $\Sigma_{\rm sSFR}$. The four parameters are not enough to depict the physical conditions that favor LyC and Lyα escape, as predicted by simulations (Choustikov et al. 2023), so they are all weak tracers of LyC or Lyα in isolation.

6. We define the relative escape fraction of Lyα photons $f^{\rm rel}_{\rm esc,Ly\alpha}$ as $f_{\rm esc,Ly\alpha}$ divided by the escape fraction of UV photons. We fit the analytical relations of $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$ with respect to $M_{\rm UV,att}$ and $M_\star$, which is proportional to the ionizing emissivity per unit observed UV luminosity density. By extrapolating the $f^{\rm rel}_{\rm esc,Ly\alpha}\xi_{\rm ion}$–$M_{\rm UV,att}$ relation to the faint-end, we calculate the fraction of ionizing emissivity from UV-faint galaxies ($M_{\rm UV,att}>-16$ mag). Qualitatively, we infer that galaxies with UV fainter than -16 mag may contribute >70% of the ionizing budget at $z=5.5$. If the dependence of $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ on $M_{\rm UV,att}$ holds until the EoR, UV-faint galaxies are predicted to be the main contributor to the reionization. Note that our HAE samples only reach $M_{\rm UV,att}\sim-17$ mag, so direct constraints on $f_{\rm esc,Ly\alpha}$ at the fainter end are required to better quantify the ionizing budget from faint galaxies.

JWST has shown promising prospects to unveil the mysteries of the EoR. Deeper and wider spectroscopic surveys in the future would reach the fainter end with lower stellar masses, and further probe the escape process of Lyα and LyC photons in newborn infant galaxies.

Acknowledgments

We thank P. A. Oesch for very helpful discussions. X.L., Z.C., Y.W., Z.L., and M.L. are supported by the National Key R&D Program of China (grant No. 2023YFA1605600), the National Science Foundation of China (grant No. 12073014), and the science research grants from the China Manned Space Project with No. CMS-CSST2021-A05. F.S. acknowledges support from the NRAO Student Observing Support (SOS) award SOSPA7-022 and the contract to the University of Arizona, NAS5-02105.

This work is based on observations made with the VLT/MUSE, NASA/ESA Hubble Space Telescope, and NASA/ESA/CSA James Webb Space Telescope. MUSE observations used in this work are taken by the MUSE-Wide Survey and the MUSE Hubble UDF Survey as parts of the MUSE Consortium. HST and JWST 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 and NAS 5-26555 for HST. The JWST observations are associated with program GO-1895. The HST observations are taken from Hubble Legacy Field Hubble Data Release. The authors are sincerely grateful to the FRESCO team for generously developing this program with a zero-exclusive-access period.

The authors also acknowledge Zechang Sun for maintaining the high-performance computing platform for the high-redshift research group in Department of Astronomy, Tsinghua University.

Facility: JWST/NIRCam, ALMA, VLT/MUSE.

Software: Grizli (v1.8.2 Brammer 2019), JWST calibration pipeline (v1.6.2 Bushouse et al. 2022), BEAGLE (Chevallard & Charlot 2016), EAZY (Brammer et al. 2008), EMCEE (Foreman-Mackey et al. 2013), SciPy (Virtanen et al. 2020), ZAP (Soto et al. 2016).

Data Availability

The HAE catalog and stack results analyzed in this work are available on Zenodo: doi:10.5281/zenodo.10802630.

Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via the FRESCO High Level Science Product (Oesch & Magee 2023).

Appendix A
Sample Selection

We present the distribution of the magnitude versus line flux and the photometric criteria for our HAE selection in Figure A1 as described in Section 2.2.1. We select mock galaxies from

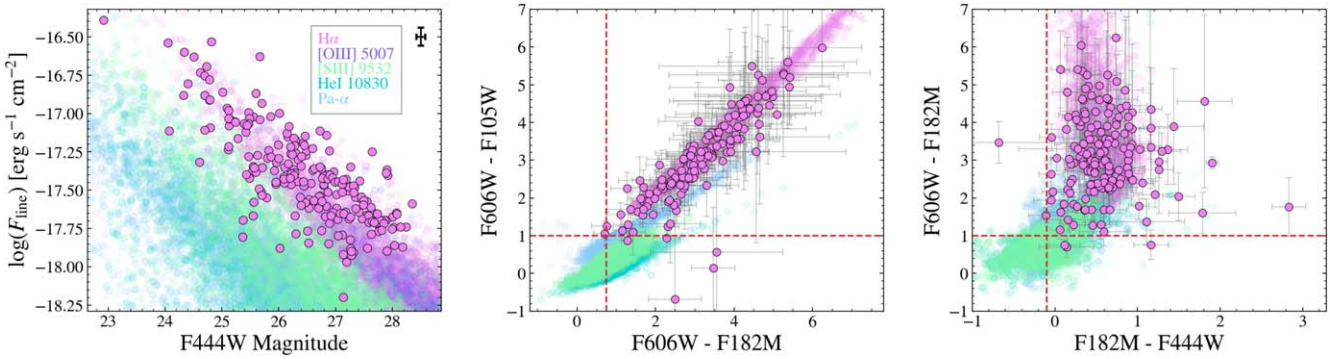

Figure A1. The flux distribution of selected HAEs (left panel) and color–color diagrams showing our photometric selection criteria (middle and right panel). The magenta dots represent the Hα flux and magnitude of our HAE sample. In the left panel, we label the typical uncertainty in line flux and magnitude. Overlying colored dots are emitters from JAGUAR mocks, residing in redshift windows that their specific emission lines could be captured by F444W grism: Hα at $z\approx4.9$–6.7, [O III] λ5007 at $z\approx6.9$–9.0, He I at $z\approx2.5$–3.6, Paα λ10830 at $z\approx1.05$–1.7, and [S III] λ9532 at $z\approx3.0$–4.3. Our selection criteria matches well with mock HAEs and can effectively rule out most low-redshift objects.

JAGUAR (Williams et al. 2018), which have certain emission features captured by NIRCam/WFSS F444W grism, and model their lines with Gaussian profiles on the mock spectra. To mimic the observational limits, we cut the mock galaxy sample by a magnitude limit of 29 for F444W and flux limit of $10^{-18.3}$ erg s^{-1} cm^{-2} for emission lines. As illustrated in Figure A1, the colors of our sample are generally consistent with those mock HAEs and could be distinguished effectively from emitters at lower redshifts. In the left panel of Figure A1, these HAEs stay at a tightly linear region in the magnitude−Hα flux diagram, possibly suggesting main-sequence galaxies at this redshift range. They perfectly overlap with the mock HAEs, and could be distinguished easily from emitters at lower redshifts. The main possible contamination is [O III] emitters at higher redshifts, most of which could be successfully picked out by the procedures in Section 2.2.2.

Appendix B
$f_{\rm esc,Ly\alpha}$ under the Assumption of SMC Dust Attenuation Law

As discussed in Section 3.1, the SMC dust attenuation law is broadly adopted in high-redshift studies. It is suggested by previous submillimeter observations (e.g., Capak et al. 2015) for typical $z \sim 5$–6 galaxies. Conventionally, Calz and SMC attenuation laws are both applied and complement each other (e.g., Bouwens et al. 2016; Shivaei et al. 2018). Figure B1 illustrates $f_{\rm esc,Ly\alpha}$ of individual galaxies and from median-stack measurements based on the SMC dust law.

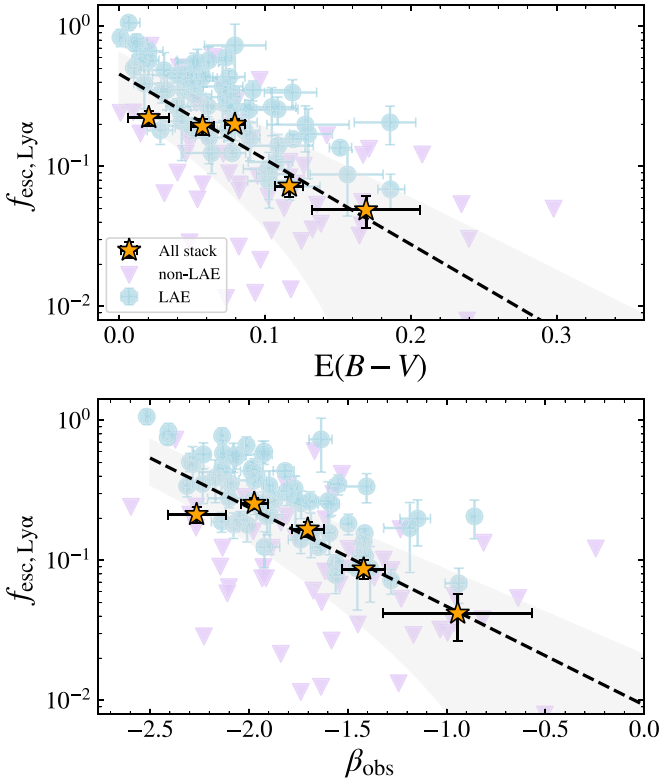

Figure B1. Same as Figure 6 while assuming the SMC dust attenuation law. The dashed lines and gray-shaded regions represent the best-fit analytical relation as described in Equation (9).

Appendix C
The Analytical Relations of $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ with Respect to $\beta_{\rm obs}$, $M_{\rm UV,att}$, and $M_\star$

The definitions of ionizing efficiency, $\xi_{\rm ion}$, may vary among the literature, from the ionizing photon production rate per unit *intrinsic* UV luminosity (e.g., Bouwens et al. 2016; Chisholm et al. 2022) to the ionizing photon production rate per unit *observed* UV luminosity (e.g., Ning et al. 2023; Simmonds et al. 2023). If $\xi_{\rm ion}$ is defined as the production rate per unit *observed* UV luminosity, the absolute escape fraction, $f_{\rm esc,LyC}$, should be used in Equation (13) to calculate the ionizing emissivity per *observed* UV luminosity density $n_{\rm ion}/\rho_{\rm UV,obs}$. To provide a reference to future works that may adopt different definitions from this work, we also fit the analytical relations of $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ with respect to $M_{\rm UV,att}$ and $M_\star$, as supplements to Equations (11) and (12). We also present $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ as a function of $\beta_{\rm obs}$, which is proportional to $f_{\rm esc,LyC}\xi_{\rm ion}$ as shown in Figure 10. As shown in Figure C1, these yield

$$\log f_{\rm esc,Ly\alpha}\,\xi_{\rm ion} = (-1.15 \pm 0.35) \times (\beta_{\rm obs} + 2.0) + (24.56 \pm 0.09), \quad \text{(C1)}$$

$$\log f_{\rm esc,Ly\alpha}\,\xi_{\rm ion} = (0.20 \pm 0.09) \times (M_{\rm UV,att} + 19.5) + (24.35 \pm 0.07), \quad \text{(C2)}$$

$$\log f_{\rm esc,Ly\alpha}\,\xi_{\rm ion} = (-0.57 \pm 0.07) \times \log(M_\star/10^{8.5}M_\odot) + (24.63 \pm 0.05). \quad \text{(C3)}$$

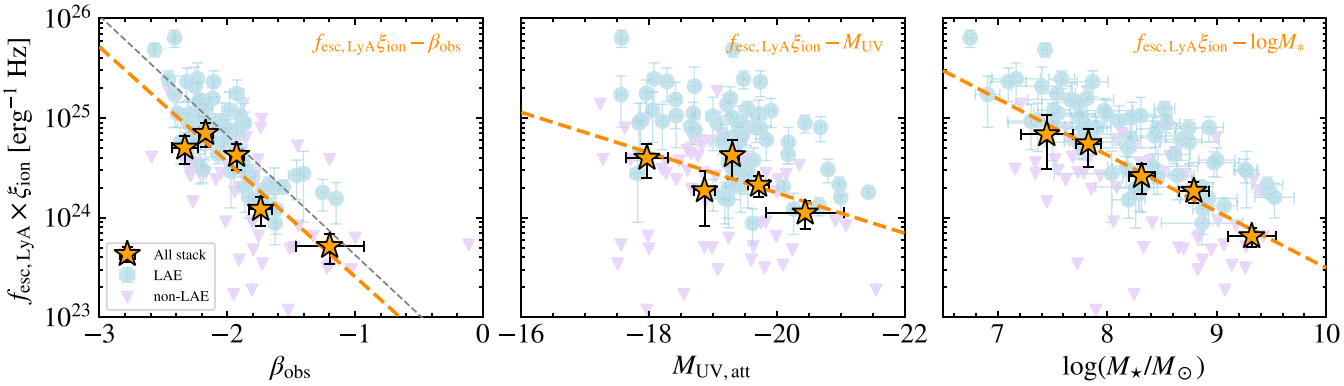

Figure C1. $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ as a function of $\beta_{\rm obs}$, $M_{\rm UV,att}$, and $M_\star$ assuming the Calz dust law. The blue circles indicate $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ for each individual LAEs, and purple triangles are $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ of non-LAEs with forced measured $f_{\rm esc,Ly\alpha}$. The fringed orange stars indicate the stack $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ for all HAEs. We adopt the stack $f_{\rm esc,Ly\alpha}$ values and the median $\xi_{\rm ion}$ of HAEs in each stack bin. We show the best-fit relations between the stack $f_{\rm esc,Ly\alpha}\xi_{\rm ion}$ of all HAEs and $\beta_{\rm obs}$, $M_{\rm UV,att}$, $M_\star$ as the orange lines. In the left panel, we show the scaled $f_{\rm esc,LyC}\xi_{\rm ion}$–$\beta_{\rm obs}$ relations, which is the $f_{\rm esc,LyC}\xi_{\rm ion}$–$\beta_{\rm obs}$ relation in Chisholm et al. (2022) divided by 0.15, as the gray dashed line.

Appendix D
Double-peaked LAEs

As discussed in Section 4.2, among the 165 HAEs, we find two LAEs presenting double-peaked Lyα profiles, shown in Figure D1. Their $f_{\rm esc,Ly\alpha}$ (0.58 ± 0.11 for HAE-60916 and 0.25 ± 0.04 for HAE-87611) are 2σ higher than the predicted values based on their $\beta_{\rm obs}$ using Equation (9) (predicted 0.20 ± 0.14 for HAE-60916 and 0.10 ± 0.09 for HAE-87611). The blue Lyα peaks of the two sources are both weaker than the red peaks. Considering the strong attenuation of the IGM, the intrinsic emissivity of the blue peaks can be much stronger and possibly stronger than the red peaks. The presence of blue Lyα peak at $z > 5$ might be a signal of enhanced escape of Lyα photons compared to normal star-forming galaxies. This is consistent with the scenarios of double-peaked LAEs with strong blue peaks at cosmic noons (Furtak et al. 2022; Marques-Chaves et al. 2022). They generally present very blue $\beta_{\rm obs}$, high-ionization state (e.g., high [O III]/O II]) and potential high Lyα or LyC escape fraction, and might be experiencing an extreme star formation episode triggered by gas inflows.

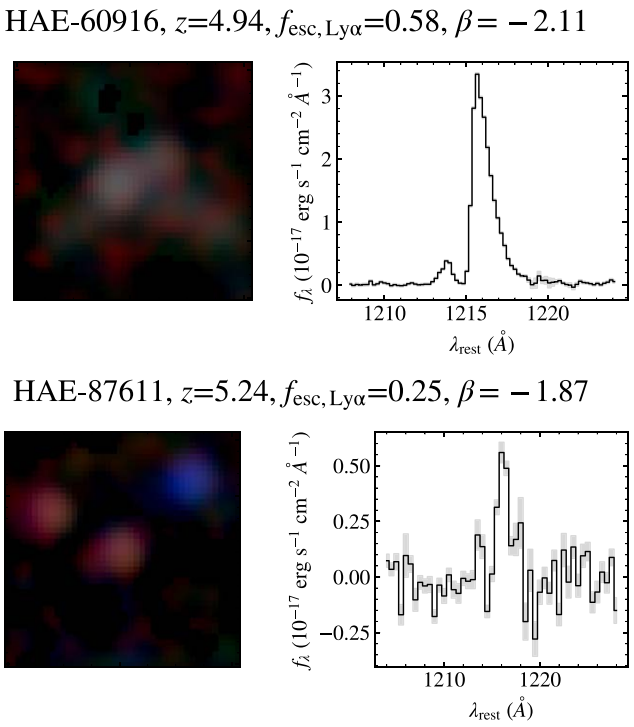

Figure D1. The two double-peaked Lyα-emitting galaxies in our final HAE sample. For each source, we show the $1\farcs25 \times 1\farcs25$ red, green, and blue thumbnail in the right panel, composed of JWST F182M, F210M, and F444W. The left panels show the rest-frame Lyα spectra extracted from the MUSE cubes.

Appendix E
Possible Systematics for Lyα Flux Measurement

In this section, we discuss possible systematics and the uncertainties it may introduce during the measurement of Lyα flux. We emphasize that the systematics we mention in the following can only introduce negligible uncertainties in the final $f_{\rm esc,Ly\alpha}$ estimates, proved by our well-tried experiments.

E.1. Contamination Rate

We compare our HAE catalog with the public JADES-Deep catalog, which contains not only photometric redshifts for each source based on JADES' deep imaging but also spectroscopical confirmation for some of the sources combined with FRESCO's WFSS data, identified by the JADES team independently (Sun 2024, in preparation). There are 80 HAEs in our sample covered by JADES-Deep, among which $\gtrsim$91% have photo-z in the range of 4.9–6.7, predicting Hα in F444W. Further, $\gtrsim$70% of them are reported to have Hα emission in JADES's catalog. These values are simply lower limits because we have not considered the mismatch due to different segmentation strategies during the catalog matching. Indeed, we do check the sources without JADES confirmed spec-z, and find most of them are (1) multicomponents and suffering mismatch issues, (2) missed by the JADES team as they only report Hα emission with S/Ns $\geqslant$5 (chat with Sun in private), while our detection threshold is S/Ns $\geqslant$4. To sum up, our HAE catalog is reliable with high purity.

E.2. The Accuracy of Lyα Measurement

As described in Section 3.2, Kerutt et al. (2022) and Bacon et al. (2023) release their LAE catalog utilizing MUSE-Wide and MUSE-Deep survey data respectively. There are 39 isolated HAEs recorded in the two literature catalogs. Different extraction and flux measurement methods are adopted, including blind line searching algorithm LSDCAT (Herenz & Wisotzki 2017) and HST-prior extraction ORIGIN (Mary et al. 2020). In this work, instead, we extract the spectra from the exposure-time-weighted coadded data cube using circular and boxy apertures, and measure the line flux by fitting skewed Gaussian profiles. For $\rm EW_{Ly\alpha}$, Kerutt et al. (2022) takes a fixed UV slope, $\beta = -1.97$, for their entire LAE sample and extrapolates the measured flux at the effective wavelength of the HST bands for the continua at the Lyα wavelength, while Bacon et al. (2023) adopts the values provided by PYPLATEFIT.[25] Their methods are totally different from ours, which yields the continuum flux density for Lyα lines by extrapolating the best-fit SED models. We compare the reported Lyα properties of the 39 HAEs from the two catalogs with our own measurements, yielding a median flux ratio of 1.04 ± 0.38 and EW ratio of 0.85 ± 0.18. Although the scattering is large, especially at the faint Lyα end, our measurement is consistent with the literature within 1σ.

E.3. Stack of Median-filtered Data Cubes with Different Depths

The MUSE data cubes used in this work have very different depths, ranging from 1h for MUSE-Wide to $\gtrsim$140h for MXDF. Therefore, the Lyα S/N for each individual HAE could vary dramatically. Theoretically, their median-stack is expected not to be biased toward those with highest S/Ns, i.e., those covered by MUSE-Deep. Another concern is the median filtering we use to remove the continuum and nearby sources. It could possibly reduce the Lyα flux by suppressing the wing of the emission lines, even though the optimized window size is chosen. We perform a test to check if these effects would significantly change our result.

Our goal is to check whether the median stack of median-filtered spectra can recover the intrinsic flux. To start with, we

[25] https://github.com/musevlt/pyplatefit

generate a series of mock 1D spectra in a spectral sampling of 0.5 Å, with random S/Ns and continua but uniform flux. 70% spectra have low S/Ns uniformly distributed from 0.1 to 3, 15% have intermediate S/Ns from three to five, and the remaining have the highest S/Ns from five to 15. We then remove the continua using a median filter of 101 pixels, the same as used in the real MUSE data, and stack all the mock spectra by taking the median value at each wavelength grid. We perform experiments for 500 times using 20, 30, 100, 200 spectra. The median values of the recovered flux for each setting reach 96%–99% of the initially set value, suggesting that a $\lesssim$5% underestimate is possible. For a $f_{\rm esc,Ly\alpha} \sim 0.1$, this systematics would simply result in an error of $\lesssim$0.005, smaller than the uncertainty as we report in this work.

ORCID iDs

Xiaojing Lin https://orcid.org/0000-0001-6052-4234
Zheng Cai https://orcid.org/0000-0001-8467-6478
Yunjing Wu https://orcid.org/0000-0003-0111-8249
Zihao Li https://orcid.org/0000-0001-5951-459X
Fengwu Sun https://orcid.org/0000-0002-4622-6617
Xiaohui Fan https://orcid.org/0000-0003-3310-0131
Zuyi Chen https://orcid.org/0000-0002-2178-5471
Mingyu Li https://orcid.org/0000-0001-6251-649X
Fuyan Bian https://orcid.org/0000-0002-1620-0897
Yuanhang Ning https://orcid.org/0000-0001-9442-1217
Linhua Jiang https://orcid.org/0000-0003-4176-6486
Gustavo Bruzual https://orcid.org/0000-0002-6971-5755
Stephane Charlot https://orcid.org/0000-0003-3458-2275
Jacopo Chevallard https://orcid.org/0000-0002-7636-0534

References

Aniano, G., Draine, B. T., Gordon, K. D., & Sandstrom, K. 2011, PASP, 123, 1218
Atek, H., Kunth, D., Schaerer, D., et al. 2009, A&A, 506, L1
Atek, H., Labbé, I., Furtak, L. J., et al. 2024, Nature
Bacon, R., Accardo, M., Adjali, L., et al. 2010, Proc. SPIE, 7735, 773508
Bacon, R., Brinchmann, J., Conseil, S., et al. 2023, A&A, 670, A4
Begley, R., Cullen, F., McLure, R. J., et al. 2024, MNRAS, 527, 4040
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bian, F., & Fan, X. 2020, MNRAS, 493, L65
Bouwens, R. J., Illingworth, G., Ellis, R. S., Oesch, P., & Stefanon, M. 2022, ApJ, 940, 55
Bouwens, R. J., Smit, R., Labbé, I., et al. 2016, ApJ, 831, 176
Brammer, G. 2019, Grizli: Grism redshift and line analysis software, Astrophysics Source Code Library, ascl:1905.001
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2022, spacetelescope/jwst: JWST v1.6.2, Zenodo, doi:10.5281/zenodo.6984366
Byrohl, C., & Gronke, M. 2020, A&A, 642, L16
Cai, Z., Cantalupo, S., Prochaska, J. X., et al. 2019, ApJS, 245, 23
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Capak, P. L., Carilli, C., Jones, G., et al. 2015, Natur, 522, 455
Chabrier, G. 2003, PASP, 115, 763
Chen, Z., Stark, D. P., Mason, C., et al. 2024, MNRAS, 528, 7052
Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415
Chisholm, J., Gazagnes, S., Schaerer, D., et al. 2018, A&A, 616, A30
Chisholm, J., Saldana-Lopez, A., Flury, S., et al. 2022, MNRAS, 517, 5104
Choustikov, N., Katz, H., Saxena, A., et al. 2024, MNRAS, 529, 3751
Curran, P. A. 2014, arXiv:1411.3816
Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605, A70
Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, ApJ, 828, 71
Dijkstra, M., & Westra, E. 2010, MNRAS, 401, 2343
Dijkstra, M., & Wyithe, J. S. B. 2012, MNRAS, 419, 3181
Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, arXiv:2306.02465
Endsley, R., Stark, D. P., Whitler, L., et al. 2023, MNRAS, 524, 2312
Faisst, A. L., Capak, P. L., Emami, N., Tacchella, S., & Larson, K. L. 2019, ApJ, 884, 133
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, ApJ, 879, 36
Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022a, ApJ, 930, 126
Flury, S. R., Jaskot, A. E., Ferguson, H. C., et al. 2022b, ApJS, 260, 1
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Furtak, L. J., Plat, A., Zitrin, A., et al. 2022, MNRAS, 516, 1373
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, A&A, 674, A1
Garilli, B., McLure, R., Pentericci, L., et al. 2021, A&A, 647, A150
Gazagnes, S., Chisholm, J., Schaerer, D., Verhamme, A., & Izotov, Y. 2020, A&A, 639, A85
Greene, T. P., Chu, L., Egami, E., et al. 2016, Proc. SPIE, 9904, 99040E
Gutkin, J., Charlot, S., & Bruzual, G. 2016, MNRAS, 462, 1757
Hayes, M., Östlin, G., Schaerer, D., et al. 2010, Natur, 464, 562
Hayes, M., Schaerer, D., Östlin, G., et al. 2011, ApJ, 730, 8
Helton, J. M., Sun, F., Woodrum, C., et al. 2024, ApJ, 962, 124
Herenz, E. C., & Wisotzki, L. 2017, A&A, 602, A111
Isobe, T., Feigelson, E. D., & Nelson, P. I. 1986, ApJ, 306, 490
Jiang, L., Ning, Y., Fan, X., et al. 2022, NatAs, 6, 850
Jin, X., Yang, J., Fan, X., et al. 2023, ApJ, 942, 59
Kakiichi, K., Ellis, R. S., Laporte, N., et al. 2018, MNRAS, 479, 43
Kashino, D., Lilly, S. J., Matthee, J., et al. 2023, ApJ, 950, 66
Katz, H., Garel, T., Rosdahl, J., et al. 2022, MNRAS, 515, 4265
Katz, H., Ďurovčíková, D., Kimm, T., et al. 2020, MNRAS, 498, 164
Kennicutt, R. C. J. 1998, ARA&A, 36, 189
Kerutt, J., Wisotzki, L., Verhamme, A., et al. 2022, A&A, 659, A183
Kim, K. J., Bayliss, M. B., Rigby, J. R., et al. 2023, ApJ, 955, L17
Kimm, T., Bieri, R., Geen, S., et al. 2022, ApJS, 259, 21
Konno, A., Ouchi, M., Nakajima, K., et al. 2016, ApJ, 823, 20
Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, MNRAS, 485, L24
Kusakabe, H., Verhamme, A., Blaizot, J., et al. 2022, A&A, 660, A44
Lam, D., Bouwens, R. J., Labbé, I., et al. 2019, A&A, 627, A164
Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, ApJ, 823, 64
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Lin, X., Cai, Z., Zou, S., et al. 2023, ApJL, 944, L59
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maji, M., Verhamme, A., Rosdahl, J., et al. 2022, A&A, 663, A66
Marchi, F., Pentericci, L., Guaita, L., et al. 2018, A&A, 614, A11
Marques-Chaves, R., Schaerer, D., Álvarez-Márquez, J., et al. 2022, MNRAS, 517, 2972
Mary, D., Bacon, R., Conseil, S., Piqueras, L., & Schutz, A. 2020, A&A, 635, A194
Mascia, S., Pentericci, L., Calabrò, A., et al. 2024, A&A, 685, A3
Maseda, M. V., Bacon, R., Lam, D., et al. 2020, MNRAS, 493, 5120
Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, ApJ, 869, 150
Naidu, R. P., Matthee, J., Oesch, P. A., et al. 2022, MNRAS, 510, 4582
Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, ApJ, 892, 109
Ning, Y., Cai, Z., Jiang, L., et al. 2023, ApJL, 944, L1
Oesch, P., & Magee, D. 2023, The JWST FRESCO Survey, STScI/MAST, doi:10.17909/GDYC-7G80
Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, MNRAS, 525, 2864
Osterbrock, D. E. 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei (Melville, NY: Univ. Science Books)
Pahl, A. J., Shapley, A., Steidel, C. C., Chen, Y., & Reddy, N. A. 2021, MNRAS, 505, 2447
Pei, Y. C. 1992, ApJ, 395, 130
Priestley, F. D., Chawner, H., Matsuura, M., et al. 2021, MNRAS, 500, 2543
Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259
Reddy, N. A., Oesch, P. A., Bouwens, R. J., et al. 2018, ApJ, 853, 56
Reddy, N. A., Steidel, C. C., Pettini, M., & Bogosavljević, M. 2016, ApJ, 828, 107
Rieke, M. J., Kelly, D. M., Misselt, K., et al. 2023, PASP, 135, 028001
Robertson, B. E. 2022, ARA&A, 60, 121
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Saldana-Lopez, A., Schaerer, D., Chisholm, J., et al. 2022, A&A, 663, A59
Saldana-Lopez, A., Schaerer, D., Chisholm, J., et al. 2023, MNRAS, 522, 6295
Saxena, A., Bunker, A. J., Jones, G. C., et al. 2024, A&A, 684, A84
Saxena, A., Pentericci, L., Ellis, R. S., et al. 2022, MNRAS, 511, 120
Schaerer, D., Marques-Chaves, R., Barrufet, L., et al. 2022, A&A, 665, L4
Scoville, N., Faisst, A., Capak, P., et al. 2015, ApJ, 800, 108
Shapley, A. E., Sanders, R. L., Reddy, N. A., Topping, M. W., & Brammer, G. B. 2023, ApJ, 954, 157
Shivaei, I., Reddy, N. A., Siana, B., et al. 2018, ApJ, 855, 42

Simmonds, C., Tacchella, S., Maseda, M., et al. 2023, MNRAS, 523, 5468
Smith, A., Ma, X., Bromm, V., et al. 2019, MNRAS, 484, 39
Sobral, D., & Matthee, J. 2019, A&A, 623, A157
Soto, K. T., Lilly, S. J., Bacon, R., Richard, J., & Conseil, S. 2016, MNRAS, 458, 3210
Stefanon, M., Bouwens, R. J., Illingworth, G. D., et al. 2022, ApJ, 935, 94
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, ApJ, 869, 123
Sun, F., Egami, E., Pirzkal, N., et al. 2022, ApJL, 936, L8
Sun, F., Egami, E., Pirzkal, N., et al. 2023, ApJ, 953, 53
Tang, M., Stark, D. P., Chevallard, J., & Charlot, S. 2019, MNRAS, 489, 2572
Theios, R. L., Steidel, C. C., Strom, A. L., et al. 2019, ApJ, 871, 128
Urrutia, T., Wisotzki, L., Kerutt, J., et al. 2019, A&A, 624, A141
Vanzella, E., Castellano, M., Bergamini, P., et al. 2022, A&A, 659, A2
Vanzella, E., Giavalisco, M., Inoue, A. K., et al. 2010, ApJ, 725, 1011
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Weilbacher, P. M., Palsa, R., Streicher, O., et al. 2020, A&A, 641, A28
Whitaker, K. E., Ashas, M., Illingworth, G., et al. 2019, ApJS, 244, 16
Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, ApJS, 236, 33
Wisotzki, L., Bacon, R., Blaizot, J., et al. 2016, A&A, 587, A98
Xu, X., Henry, A., Heckman, T., et al. 2022, ApJ, 933, 202
Yang, H., Malhotra, S., Gronke, M., et al. 2016, ApJ, 820, 130
Zhang, S., Cai, Z., Xu, D., et al. 2023, Sci, 380, 494
Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2021, ApJ, 923, 223