JWST/NIRSpec spectroscopy of $z = 7–9$ star-forming galaxies with CEERS: new insight into bright Ly$\alpha$ emitters in ionized bubbles

Mengtao Tang $^\dagger$$^1$*, Daniel P. Stark, $^1$Zuyi Chen $^1$, Charlotte Mason $^{2,3}$, Michael Topping $^1$, Ryan Endsley $^1$$^2$, Peter Sencilina $^5$, Adele Plat $^1$, Ting-Yi Lu $^{2,3}$, Lily Whitler $^1$, Brant Robertson$^6$ and Stéphane Charlot$^7$

$^1$Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA
$^2$Cosmic Dawn Center (DAWN), Copenhagen, Denmark
$^3$Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 København N, Denmark
$^4$Department of Astronomy, University of Texas, Austin, TX 78712, USA
$^5$The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
$^6$Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
$^7$Ecole normale supérieure, CNRS, UMR 7095, Institut d’Astrophysique de Paris, 98 bis bd Arago, F-75014 Paris, France

Accepted 2023 September 7. Received 2023 September 4; in original form 2023 January 17

ABSTRACT

We describe new JWST/NIRSpec observations of galaxies at $z \gtrsim 7$ taken from the CEERS survey. Previous observations of this area have revealed associations of Ly$\alpha$ emitters at redshifts ($z = 7.5, 7.7$, and $8.7$) where the intergalactic medium (IGM) is thought to be mostly neutral, leading to suggestions that these systems are situated in large ionized bubbles. We identify 21 $z \gtrsim 7$ galaxies with robust redshifts in the CEERS data set, including 10 in the Ly$\alpha$ associations. Their spectra are indicative of very highly ionized and metal poor gas, with line ratios ($\text{O}3\alpha = 17.84$ and $\text{Ne}3\alpha = 0.89$, linear scale) and metallicity ($12 + \log (\text{O/H}) = 7.84$) that are rarely seen at lower redshifts. We find that the most distant spectral features are found in the six $z \gtrsim 7$ Ly$\alpha$ emitters in the sample. Each has a hard ionizing spectrum indicating that their visibility is likely enhanced by efficient ionizing photon production. Ly$\alpha$ velocity offsets are found to be very large ($\gtrsim 300$ km s$^{-1}$), likely also contributing to their detectability. We find that Ly$\alpha$ in $z \gtrsim 7$ galaxies is $6–12 \times$ weaker than in lower redshift samples with matched rest-optical spectral properties. If the bubbles around the Ly$\alpha$ emitters are relatively small ($\lesssim 0.5–1$ pMpc), we may expect such significant attenuation of Ly$\alpha$ in these ionized regions. We discuss several other effects that may contribute to weaker Ly$\alpha$ emission at $z \gtrsim 7$. Deep spectroscopy of fainter galaxies in the vicinity of the Ly$\alpha$ emitters will better characterize the physical scale of the ionized bubbles in this field.

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

1 INTRODUCTION

The reionization of intergalactic hydrogen is an important landmark in the early history of structure formation. The details of how and when this process came to pass depends sensitively on the nature of the first energetic objects emitting hydrogen ionizing radiation. Over the past two decades, concerted observational efforts have been devoted to constraining the history of reionization and the nature of early ionizing sources (Dijkstra 2014; Stark 2016; Robertson 2022). The process is thought to be driven by the radiation from both massive stars (e.g. Robertson et al. 2015; Stanway, Eldridge & Becker 2016; Dayal & Ferrara 2018; Finkelstein et al. 2019; Naidu et al. 2020) and active galactic nuclei (AGNs; e.g. Haardt & Salvesterra 2015; Matsuoka et al. 2018; Kulkarni et al. 2019; Dayal et al. 2020). Measurements of the Thomson scattering optical depth faced by cosmic microwave background photons suggest a significant component of free electrons are present by $z \sim 8$ (Planck Collaboration 2016, 2020). Quasar absorption spectra tell us that the intergalactic medium (IGM) is partially neutral at $z \sim 7$ (e.g. Greig et al. 2017; Davies et al. 2018; Wang et al. 2020; Yang et al. 2020a) and mostly ionized by $z \sim 5.5–6$ (e.g. Fan et al. 2006; McGreer, Mesinger & D’Odorico 2015; Yang et al. 2020b).

Observations of Ly$\alpha$ emission from star-forming galaxies provide a complementary picture. While strong Ly$\alpha$ is commonly seen in $z \lesssim 6$ galaxies (e.g. Stark, Ellis & Ouchi 2011; Curtis-Lake et al. 2012; Cassata et al. 2015; De Barros et al. 2017; Jiang et al. 2017), it is very rare at $z \gtrsim 7–8$ (e.g. Fontana et al. 2010; Stark et al. 2010; Treu et al. 2013; Caruana et al. 2014; Pentericci et al. 2014; TIlvi et al. 2014; Hoag et al. 2019; Mason et al. 2019; Jung et al. 2020). The equivalent width (EW) distribution of Ly$\alpha$ shows marked evolution toward weaker line emission between $z \sim 6$ and $z \sim 8$ (e.g. Ono et al. 2012; Schenker et al. 2014; TIlvi et al. 2014; Jung et al. 2018, 2020), as would be expected if the IGM is significantly neutral at $z \gtrsim 7$ (e.g., Mesinger et al. 2015; Zheng et al. 2017; Mason et al. 2018a; Hoag et al. 2019; Mason et al. 2019; Whitler et al. 2020).

As wider-area infrared imaging surveys emerged over the last decade, attention has turned to massive reionization-era galaxies that sit at the bright end of the ultraviolet (UV) luminosity function ($M_{\text{UV}} < -21$). Many of these UV-bright galaxies were found to have

* E-mail: tangmtasua@arizona.edu

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society
Ly\(\alpha\) emission at \(z \approx 7.5\,\text{–}\,9\) (e.g. Ono et al. 2012; Finkelstein et al. 2013; Oesch et al. 2015; Zitrin et al. 2015; Roberts-Borsani et al. 2016; Stark et al. 2017; Larson et al. 2022). Initial studies suggested that Ly\(\alpha\) was fairly ubiquitous in this population (Stark et al. 2017). Subsequent surveys over larger volumes have confirmed that the Ly\(\alpha\) EW distribution in the UV-bright population shows no significant decline over \(6 < z < 8\) (e.g. Endsley et al. 2021b; Jung et al. 2022; Roberts-Borsani et al. 2023b), suggesting that Ly\(\alpha\) photons emitted by massive UV-luminous galaxies are minimally impacted by the neutral IGM at \(z \approx 7\).

The luminosity-dependence of the evolving Ly\(\alpha\) visibility can be understood if the brightest galaxies tend to trace rare ionized regions in the IGM. If the bubbles are large enough, Ly\(\alpha\) will redshift far enough into the damping wing before encountering neutral hydrogen, boosting the transmission of the line through the IGM (e.g. Mason et al. 2018b; Qin et al. 2022). Galaxies in overdensities produce an excess number of ionizing photons, so the largest bubbles are expected to be found in overdense regions (e.g. Barkana & Loeb 2004; Furlanetto, Zaldarriaga & Hernquist 2004; Iliev et al. 2006; Castellano et al. 2016; Garaldi et al. 2022; Kannan et al. 2022; Leonova et al. 2022; Lu et al. 2023). Whether the bright Ly\(\alpha\) emitters at \(z \approx 7\) are tracing such large ionized bubbles is not clear. Several studies have recently reported discovery of excess numbers of bright galaxies in the vicinity of \(z \approx 7\) Ly\(\alpha\) emitters (e.g. Castellano et al. 2018; Tilvi et al. 2020; Endsley & Stark 2022; Jung et al. 2022; Leonova et al. 2022), suggesting possible overdensities in these regions. Some of these bright neighbouring galaxies have also been shown to exhibit strong Ly\(\alpha\) emission (e.g. Jung et al. 2020; Endsley & Stark 2022; Jung et al. 2022), as might be expected if the ionized regions extend over very large volumes.

Much of this progress has been focused on the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (Grogin et al. 2011; Koekemoer et al. 2011) Extended Groth Strip (EGS; Davis et al. 2007) field, with groups of Ly\(\alpha\) emitters confirmed at \(z \approx 7.5, z \approx 7.7,\) and \(z \approx 8.7\). The first Ly\(\alpha\) emitter galaxies were identified as bright dropouts selected based on the presence of intense [O III] + H\(\beta\) emission (Roberts-Borsani et al. 2016). Keck spectroscopic follow-up of this initial sample revealed Ly\(\alpha\) emission at \(z = 7.48\) (Roberts-Borsani et al. 2016; Stark et al. 2017), \(z = 7.73\) (Oesch et al. 2015), and \(z = 8.68\) (Zitrin et al. 2015). Subsequent work has revealed additional Ly\(\alpha\) emitters within 5–10 physical Mpc (pMpc) of each of these galaxies (Tilvi et al. 2020; Jung et al. 2022; Larson et al. 2022), potentially indicating the presence of several extended ionized regions. Deep imaging with the Hubble Space Telescope (HST) suggests that this field may have a significant overdensity of galaxies at these redshifts (Leonova et al. 2022), as would be required to power such large ionized regions.

Previous data do not strongly constrain the size of the ionized bubbles around the \(z \approx 7\) Ly\(\alpha\) emitters in the EGS. The physical scale of ionized regions at a given redshift will depend on the progress of reionization and the nature of the sources driving the process (e.g. Furlanetto et al. 2004; McQuinn et al. 2007; Mason & Gronke 2020; Hutter et al. 2021; Leonova et al. 2022; Smith et al. 2022). If the typical bubble sizes are large at \(z \gtrsim 7.8\), we may expect the entire 5–10 pMpc region spanned by the Ly\(\alpha\) emitters at each redshift to be ionized. But the existing observations are also consistent with there being smaller bubbles \((\lesssim \text{pMpc})\) cented on the individual bright Ly\(\alpha\) emitting galaxies. In this latter case, the visibility of Ly\(\alpha\) may be boosted by other factors. Intense radiation fields can increase the production efficiency of Ly\(\alpha\) photons, enhancing the likelihood of detection. Outflows can result in Ly\(\alpha\) profiles that are redshifted from systemic, boosting the transmission through the IGM (e.g. Stark et al. 2017; Mason et al. 2018b; Endsley et al. 2022). Unfortunately, little is known about the importance of either of these factors in regulating Ly\(\alpha\) detection rates in the reionization era, limiting our ability to interpret \(z \gtrsim 7\) observations in the context of bubble sizes.

Spectroscopy of \(z > 6\) galaxies with JWST (Gardner et al. 2023) has begun providing important insights into the physical properties of systems in the reionization era (e.g. Schaerer et al. 2022b; Sun et al. 2022, 2023; Taylor, Barger & Cowie 2022; Wang et al. 2022; Backhaus et al. 2023; Bunker et al. 2023; Cameron et al. 2023; Carnall et al. 2023; Curtis-Lake et al. 2023; Fujimoto et al. 2023; Hsiao et al. 2023; Jung et al. 2023; Larson et al. 2023; Matthee et al. 2023; Nakajima et al. 2023; Shapley et al. 2023; Tacchella et al. 2023; Trump et al. 2023; Roberts-Borsani et al. 2023a; Sanders et al. 2023a; Saxena et al. 2023b). In this paper, we investigate new JWST/NIRSpec (Ferruit et al. 2022; Jakobsen et al. 2022) observations of galaxies at \(z \gtrsim 7\) in the EGS field taken as part of the Cosmic Evolution Early Release Science (CEERS\(^1\)) program (Finkelstein et al. in preparation, see also Finkelstein et al. 2022, 2023). This JWST data set opens many new avenues of characterizing early galaxies and their surrounding IGM. We investigate the spectroscopic properties of a sample of 21 galaxies at \(z \gtrsim 7\) with robust emission line detections in the CEERS observations. Using this data base, we explore how the galaxies in the EGS with Ly\(\alpha\) detections at \(z \gtrsim 7\) compare to the general population. We investigate the physics regulating Ly\(\alpha\) production and escape, leveraging new constraints on ionizing photon production efficiency, Ly\(\alpha\) velocity offsets, and the Ly\(\alpha\) escape fraction. The observations include 10 galaxies in the three Ly\(\alpha\) associations, improving characterization of the ionized regions surrounding these systems.

The organization of this paper is as follows. In Section 2, we describe the JWST/NIRSpec observations (Section 2.1), spectral energy distributions (SEDs) of galaxies in our spectroscopic sample (Section 2.2), rest-frame optical (Section 2.3), and UV emission line measurements (Section 2.4). Based on CEERS spectroscopic results, we present the physical properties of galaxies in the three groups of Ly\(\alpha\) emitting galaxies at \(z > 7\) in the CEERS field in Section 3. We use the results to discuss the implications for Ly\(\alpha\) visibility in the reionization era in Section 4. Finally, we summarize our conclusions in Section 5. Throughout the paper, we adopt a \(\Lambda\)-dominated, flat universe with \(\Omega_m = 0.7, \Omega_{\Lambda} = 0.3,\) and \(H_0 = 70\) km s\(^{-1}\) Mpc\(^{-1}\). All magnitudes in this paper are quoted in the AB system (Oke & Gunn 1983), and all EWs are quoted in the rest frame.

2 DATA AND ANALYSIS

2.1 NIRSpec observations and reduction

We use the publicly available CEERS JWST/NIRSpec spectra that are centred on the EGS field. The CEERS NIRSpec observations of the EGS field include six pointings using the medium resolution (MR) grating with three grating/filter pairs (G140M/F100LP, G235M/F170LP, and G395M/F290LP), covering 1–5 \(\mu\)m and four additional pointings using the lower resolution prism providing simultaneous coverage of 1–5 \(\mu\)m. Details of the CEERS NIRSpec observations are summarized in the CEERS public Phase 2 PDF;\(^2\) In brief, each grating and prism was observed with three exposures of 13 groups each, using the standard three-shutter slits for MR grating

\footnotetext{1}{https://ceers.github.io/}

\footnotetext{2}{https://www.stsci.edu/jwst/phase2-public/1345.pdf}
observations and three-point nodding. The total exposure time of each grating and prism is 2889 s. The spectral resolution of MR grating is $R \approx 1000$ (corresponding to a velocity resolution $\sigma \approx 130$ km s$^{-1}$) and of prism is $R \approx 100$ ($\sigma \approx 1300$ km s$^{-1}$). Overall, there are 319 galaxies placed on the 6 MR grating pointings (52–55 galaxies on each pointing) and 466 galaxies on the three prism pointings (151–161 galaxies on each pointing). Targets were selected from Lyman break galaxies identified using the HST (Bouwens et al. 2015; Finkelstein et al. 2015; Bouwens et al. 2021) and JWST/NIRCam (Rieke et al. 2023) imaging (Finkelstein et al. 2023; Whiter et al. 2023a; Endsley et al. 2023b), spanning a wide photometric redshift range of $z \approx 0.5–12$. The CEERS collaboration prioritize targets that are at redshifts where key emission lines or continuum will be detected at $1–5 \mu$m that allow the measurement of redshifts and line diagnostics.

The 2D NIRSpec spectra were reduced following the methods described in Shapley et al. (2023). Individual uncalibrated detector exposures were first passed through the JWST calwebb_detector1 pipeline. This step implements masking of all pixels that are saturated, subtraction of signal due to bias, and dark current, and masking of ‘snowballs’ and ‘showers’ resulting from cosmic rays. The resulting images from this step were then corrected for striping. This is done by estimating and subtracting the 1/f noise in each image. The 2D spectrum of each slit on the micro-shutter assembly (MSA; Ferruit et al. 2022) was then cut out. We applied the flat-field correction, photometry calibration, and the wavelength solution using the up-to-date calibration reference data system context (jwst_version $1.27$, pmap). Each slitlet was then rectified, and interpolated onto a common wavelength grid for its grating and filter combination. The calibrated spectra were finally combined, following the three-shutter dither pattern, excluding pixels that had been masked in a previous step of the reduction. The 2D error spectra were calculated as a combination of the variance from Poisson noise, readout noise, flat-fielding, and variance between exposures, all summed in quadrature.

Every reduced spectrum (low and MR) was visually examined by a small group of the co-authors (MT and ZC). We searched for emission lines in all galaxies with spectra, assigning spectroscopic redshifts where lines are detected. In total, we found 21 unique galaxies (16 with MR grating spectra and additional 5 with prism spectra) with robust emission line detections placing them confidently at $z \gtrsim 7$. In Figs 1 and 2, we show the 2D NIRSpec spectra with emission line detections of the 16 galaxies with MR grating observations and the other five with prism observations, respectively. Additional sources are found with detections of single or lower S/N lines that may be at $z \gtrsim 7$, but we limit our analysis here to the 21 galaxies that we can robustly place in our chosen redshift range. Redshift confirmation is typically achieved via detection of the [O III] $\lambda\lambda 4959, 5007$ doublet and one or more of the hydrogen Balmer lines (i.e. Hf, Hg), but we additionally detect Ly$\alpha$ and fainter UV lines (e.g. C III) in several cases. These 21 objects range in redshift between $z = 6.928$ and $z = 8.999$, with a median redshift of $z = 7.545$. The spectroscopic sample includes 10 galaxies in the three associations of Ly$\alpha$ emitters (LAEs) at $z = 7.48$, $z = 7.73$, and $z = 8.68$, which we will discuss in more detail in Section 3. An overview of our $z \gtrsim 7$ spectroscopic catalog is given in Table 1.

The flux calibration was performed to the 2D spectrum of each source using the PHOTOM reference file in the JWST data reduction pipeline. We computed slit loss corrections following the path-loss correction step in the pipeline. Since the galaxies in our sample are not significantly extended, we apply the path-loss correction assuming a point source instead of an uniform extended source. The path-loss correction factor depends on wavelength and is $\sim 1.0–1.4 \times$ for the 21 galaxies in our sample. Nevertheless, the emission line ratios discussed in this paper (described below) do not change significantly (<10 per cent) before and after the path-loss correction. We explored path-loss corrections assuming an uniform source as well and find these have a negligible effect on our analysis based on line ratios. We also explored a different slit loss correction than the path-loss correction procedures applied by JWST pipeline. To do this, we first extracted a postage stamp of each galaxy from the HST F160W image (Skelton et al. 2014). We then smooth the postage stamp and fit the smoothed image with a 2D Gaussian profile. The HST point spread function (PSF) is deconvolved from the profile to obtain the intrinsic morphology of the sources. The deconvolved profile is then convolved with the wavelength-dependent JWST PSF computed using WebPSP$^9$ (Perrin et al. 2014). Using the wavelength-dependent profile, we place a box on each source profile with the same length and width and position angle of the NIRSpec MSA micro-shutter according to the position of the source in the shutter. Finally, we compute the ratio of the light inside the mimic shutter box to the total light of the source profile and derive the slit-loss correction factor as a function of wavelength. We find broadly similar results to the pipeline with emission line ratios that are still stable before and after applying this wavelength-dependent slit-loss correction.

For each source in our spectroscopic catalog, we extract the 1D spectrum from the calibrated 2D spectrum using a boxcar extraction. The extraction aperture is designed to match the emission line profile along the spatial direction and the typical width is $\sim 6$ pixels. Systemic redshifts are derived for each galaxy using the central wavelengths of the strongest rest-frame optical emission lines. Fig. 3 shows 1D profiles of strong rest-frame optical emission lines detected in MR grating spectra. We test the consistency of the wavelength calibration in each of the three bands of the MR gratings by comparing the systemic redshift derived from non-resonant emission lines seen in each setup. The redshifts measured between the three settings are nearly identical ($\Delta z < 0.0005$), indicating that the wavelength calibration should be reliable between different bands. We then derive emission line fluxes for each object. The line fluxes are computed through fitting a Gaussian profile to emission lines with S/N $> 5$. For lines with lower S/N ($< 5$), we calculate the fluxes using direct integration. Four of the galaxies with MR grating spectra are also observed in the low-resolution prism mode. The line fluxes of strong rest-frame optical emission lines are consistent in both spectra within 2$\sigma$ uncertainty. For undetected emission lines, we derive a 3$\sigma$ upper limit of the flux by summing the error spectrum in quadrature over $\sim 700$ km s$^{-1}$. This integration range is chosen to be consistent with the upper bound of line widths found for rest-frame UV and optical emission lines of extreme emission line galaxies (e.g. Maseda et al. 2014; Mainali et al. 2020; Tang, Stark & Ellis 2022).

---

3https://jwst-pipeline.readthedocs.io/


6https://webbpsf.readthedocs.io/
Figure 1. MR ($R \sim 1000$) 2D NIRSpec spectra of the 16 galaxies at $z \gtrsim 7$ at rest-frame optical wavelengths. Blue lines show the rest-frame optical emission lines detected on 2D spectra. The two tentative detections of [O III] $\lambda 4363$ in CEERS-689 and CEERS-698 are shown by blue dotted lines.
among the three bands of the MR gratings. To do this, we choose sources at redshifts where the same strong lines are detected in two bands simultaneously (e.g. H\textbeta detected in both G235M and G395M at \( z \simeq 5.1 \), and H\textbeta detected in both G140M and G235M at \( z \simeq 2.7 \)). Then we measure the flux of the same emission line in each band, and we find that the line fluxes measured in different bands are consistent within the 1\( \sigma \) uncertainty. In this work, we will primarily focus on emission line ratios that are not sensitive to absolute flux calibration but probe the ionizing spectra, the gas excitation conditions, and the gas-phase metallicity. We will consider the following ratios and use the following definitions: O3 (≡ [O III]\( \lambda \lambda 5007/\text{H}\beta \)), O32 (≡ [O III]\( \lambda \lambda 4959, 5007/\text{O II}\lambda \lambda 3726, 3729 \)), Ne3O22 (≡ [Ne III]\( \lambda \lambda 3869/\text{O II}\lambda \lambda 3726, 3729 \)), R23 (≡ ([O III]\( \lambda \lambda 4959, 5007 + \text{[O II]}\lambda \lambda 3726, 3729)/\text{H}\beta \)), and [O III]\( \lambda \lambda 4363/[\text{[O III]}\lambda 5007].

### 2.2 Spectral energy distributions of \( z \gtrsim 7 \) galaxies

We use available broad-band SEDs to investigate the stellar population properties of the spectroscopic sample. The SEDs are also critical for characterization of the [O III]+ H\textbeta EWs based on flux excesses in the 3–5 \( \mu \)m photometry. The imaging-based SED analysis in this paper will focus primarily on objects with NIRCam measurements, given the vastly improved photometric constraints that are achieved relative to what is possible with HST and Spitzer. A subset of the galaxies in our sample (11 of 21) overlap with the region of the EGS that has been observed by NIRCam as part of the CEERS ERS program. These include sources covered in the original NIRCam pointings taken in summer 2022 (CEERS-23, CEERS-24, CEERS-3, CEERS-1027, CEERS-498, CEERS-499, CEERS-44, and CEERS-407) and three additional sources observed in NIRCam imaging undertaken in December 2022 (CEERS-1019, CEERS-1025, and CEERS-1038). The imaging associated with the latter sources has been reduced in the same manner as described in Endsley et al. (2023b) and the reader is directed to that paper for details. For these sources, we compute SEDs using CEERS photometry catalogs presented in Endsley et al. (2023b). The CEERS NIRCam imaging was taken in six broad-band filters (F115W, F150W, F200W, F277W, F356W, and F444W) and one medium-band filter (F410M). The flux of each band was measured in elliptical aperture with a Kron (1980) factor of \( k = 1.2 \) and then was corrected to the total flux by multiplying the ratio of flux measured in \( k = 2.5 \) to \( k = 1.2 \) aperture in F200W. Flux from neighbouring objects that is in the aperture was also subtracted. The derived NIRCam H-band magnitudes (F150W) range between 30.00 and 24.85 in these 11 galaxies, implying absolute UV magnitude (\( M_{\text{UV}} \)) between −16.97 and −22.42. The SEDs are shown in Fig. 4. For these 11 galaxies with NIRCam photometry, colour excesses are readily apparent in the long-wavelength photometry, revealing the characteristic signature of strong rest-frame optical line emission. We will come back to quantify the implied [O III]+ H\textbeta EWs later in the subsection.

For the 10 sources that lack NIRCam photometry, we collate HST and Spitzer photometry. The HST Wide Field Camera 3 (WFC3) IR photometry is computed for each source using the mosaics produced as part of the Complete Hubble Archive for Galaxy Evolution (CHaGE) project (Kokorev et al. 2022). The CHaGE mosaics include the HST/WFC3 data in the EGS field and are matched to the Gaia astrometric frame with a pixel scale of 80 mas pixel\(^{-1}\). Details of the CHaGE data used in CEERS have been described in Chen et al. (2023) and Endsley et al. (2023b). The derived H\textalpha band magnitudes range between 26.90 and 24.91, implying \( M_{\text{UV}} \) between −20.04 and −22.14. At these magnitudes, the Spitzer measurements are mostly not useful given the limited sensitivity

Figure 2. Low resolution (\( R \sim 100 \)) 2D (top of each panel) and 1D (bottom of each panel) NIRSpec prism spectra of the five galaxies at \( z > 7 \) with prism observations. Blue lines mark the position of emission lines detected in the spectrum of each object. The grey shaded regions show the uncertainty of flux.
of the IRAC mosaics across the EGS (typical 5σ limit of $m \approx 25.2$ at 3.6 μm). Only one galaxy in this subset of 10 (CEERS-698) has Spitzer/IRAC flux density measurements with S/N in excess of 5 in both [3.6] and [4.5]. We add this $z = 7.470$ source (one of the original Roberts-Borsani et al. 2016 galaxies) to our sub-sample for SED-based analysis. For the remaining sources, the absence of significant rest-frame optical constraints keeps stellar population modeling from yielding meaningful information. While we do investigate the emission line properties of this subset of nine galaxies, we do not include their imaging-based SED fits in our subsequent analysis.

We fit the SEDs of the 12 sources (11 with JWST/NIRSpec photometry and the one with high S/N HST+Spitzer photometry) with rest-frame UV to optical photometry using the Bayesian galaxy SED modeling and interpreting tool BayEisian Analysis of Galaxy sEDs (BEAGLE, version 0.20.4; Chevallard & Charlot 2016). The BEAGLE

---

<table>
<thead>
<tr>
<th>ID</th>
<th>$z_{\text{spec}}$</th>
<th>RA</th>
<th>Dec.</th>
<th>NIRCam/F150W</th>
<th>WFC3/F160W</th>
<th>$M_{\text{UV}}$</th>
<th>Optical lines*</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEERS-24(^b)</td>
<td>8.999</td>
<td>14:19:35.336</td>
<td>+52:50:37.87</td>
<td>27.93 ± 0.12</td>
<td>...</td>
<td>−19.39</td>
<td>[O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-23(^b)</td>
<td>8.881</td>
<td>14:19:36.301</td>
<td>+52:50:49.19</td>
<td>28.44 ± 0.23</td>
<td>...</td>
<td>−18.87</td>
<td>H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1025(^b)</td>
<td>8.715</td>
<td>14:19:52.211</td>
<td>+52:55:58.63</td>
<td>26.15 ± 0.02</td>
<td>...</td>
<td>−21.12</td>
<td>[O II], [Ne III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1019(^b)</td>
<td>8.678</td>
<td>14:20:08.494</td>
<td>+52:53:26.38</td>
<td>24.85 ± 0.02</td>
<td>...</td>
<td>−22.42</td>
<td>$\gamma$, H$\beta$, [O III]</td>
<td>[1]</td>
</tr>
<tr>
<td>CEERS-1029(^b)</td>
<td>8.610</td>
<td>14:20:52.503</td>
<td>+53:04:11.50</td>
<td>25.63 ± 0.07</td>
<td>...</td>
<td>−21.63</td>
<td>[O II], H$\beta$, [O III]</td>
<td>[2]</td>
</tr>
<tr>
<td>CEERS-1149(^b)</td>
<td>8.175</td>
<td>14:20:21.531</td>
<td>+52:57:58.26</td>
<td>26.58 ± 0.15</td>
<td>...</td>
<td>−20.60</td>
<td>[O II], [Ne III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-3(^c)</td>
<td>8.00</td>
<td>14:19:01.245</td>
<td>+52:59:47.69</td>
<td>28.61 ± 0.28</td>
<td>...</td>
<td>−18.54</td>
<td>H$\gamma$, H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1027(^b)</td>
<td>7.819</td>
<td>14:19:31.919</td>
<td>+52:50:25.50</td>
<td>26.50 ± 0.03</td>
<td>...</td>
<td>−20.60</td>
<td>[Ne III], H$\alpha$, H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1023(^b)</td>
<td>7.776</td>
<td>14:20:45.219</td>
<td>+53:02:01.13</td>
<td>...</td>
<td>26.23 ± 0.16</td>
<td>−20.87</td>
<td>[O II], H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-686(^b)</td>
<td>7.74</td>
<td>14:20:36.207</td>
<td>+52:59:22.42</td>
<td>26.44 ± 0.10</td>
<td>...</td>
<td>−20.66</td>
<td>H$\beta$, [O III]</td>
<td>[3]</td>
</tr>
<tr>
<td>CEERS-689(^b)</td>
<td>7.545</td>
<td>14:19:59.773</td>
<td>+52:56:31.12</td>
<td>24.91 ± 0.18</td>
<td>...</td>
<td>−22.14</td>
<td>[O II], [Ne III], H$\gamma$, [O III]</td>
<td>[1]</td>
</tr>
<tr>
<td>CEERS-698(^b)</td>
<td>7.470</td>
<td>14:20:12.076</td>
<td>+53:00:26.79</td>
<td>25.18 ± 0.04</td>
<td>...</td>
<td>−20.86</td>
<td>[O II], [Ne III], H$\epsilon$, [O III]</td>
<td>[4], [5]</td>
</tr>
<tr>
<td>CEERS-1163(^b)</td>
<td>7.448</td>
<td>14:19:57.712</td>
<td>+52:58:19.16</td>
<td>26.80 ± 0.17</td>
<td>...</td>
<td>−20.24</td>
<td>[O II], H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1038(^b)</td>
<td>7.194</td>
<td>14:20:09.527</td>
<td>+52:54:05.75</td>
<td>27.73 ± 0.13</td>
<td>...</td>
<td>−19.25</td>
<td>H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-498(^b)</td>
<td>7.18</td>
<td>14:19:15.131</td>
<td>+52:50:03.30</td>
<td>26.76 ± 0.03</td>
<td>...</td>
<td>−20.21</td>
<td>H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-499(^b)</td>
<td>7.168</td>
<td>14:19:15.121</td>
<td>+52:50:03.01</td>
<td>30.00 ± 0.54</td>
<td>...</td>
<td>−16.97</td>
<td>H$\gamma$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-44(^c)</td>
<td>7.10</td>
<td>14:20:00.268</td>
<td>+53:00:40.57</td>
<td>27.59 ± 0.06</td>
<td>...</td>
<td>−19.37</td>
<td>H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-407(^b)</td>
<td>7.028</td>
<td>14:19:21.436</td>
<td>+52:52:57.23</td>
<td>27.99 ± 0.09</td>
<td>...</td>
<td>−18.95</td>
<td>[O II], H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1102(^c)</td>
<td>7.00</td>
<td>14:20:21.851</td>
<td>+52:57:15.43</td>
<td>26.90 ± 0.22</td>
<td>...</td>
<td>−20.04</td>
<td>H$\beta$, [O III], H$\epsilon$</td>
<td></td>
</tr>
<tr>
<td>CEERS-717(^b)</td>
<td>6.932</td>
<td>14:20:19.537</td>
<td>+52:58:19.85</td>
<td>25.28 ± 0.06</td>
<td>...</td>
<td>−21.64</td>
<td>[O II], H$\gamma$, H$\beta$, [O III]</td>
<td></td>
</tr>
<tr>
<td>CEERS-1143(^b)</td>
<td>6.928</td>
<td>14:20:18.482</td>
<td>+52:58:10.22</td>
<td>26.88 ± 0.18</td>
<td>...</td>
<td>−20.04</td>
<td>[O II], [Ne III], H$\gamma$, H$\beta$, [O III]</td>
<td></td>
</tr>
</tbody>
</table>

* Emission line optical emission lines detections. [O II], [Ne III], and [O III] refer to [O II]$\lambda$3727, [Ne III]$\lambda$3869, and [O III]$\lambda$4959, respectively.

b Emission lines detected in NIRSpec MR grating spectra.

c Emission lines detected in NIRSpec prism spectra.

Figure 3. MR ($R \sim 1000$) NIRSpec 1D spectra of rest-frame optical emission lines of $z \geq 7$ galaxies with more than three detected lines ([O III], Hβ, and [O II] or [Ne III] or Hγ; 11 objects). The two tentative detections of [O III] λ4363 in CEERS-689 and CEERS-698 are shown by blue dotted lines. The spectra have been shifted to the rest frame. The grey shaded regions show the flux uncertainty. Spectra of the remaining five galaxies with three and fewer rest-frame optical line detections are shown in Fig. A1.
set up and SED fitting procedures are mainly described in Endsley et al. (2023b) and are briefly summarized in the following. BEAGLE utilizes a combination of the latest version of Bruzual & Charlot (2003) stellar population synthesis models and the Gutkin, Charlot & Bruzual (2016) photoionization models of star-forming galaxies with the CLOUDY code (Ferland et al. 2013). We assume a constant star formation history (CSFH) and allow the galaxy age to vary between 1 Myr and the age of the Universe at the given spectroscopic redshift with a log-uniform prior. We adopt a Chabrier (2003) initial mass function (IMF) with a stellar mass range of 0.1–300 M⊙ and allow the metallicity to vary in the range $-2.2 \leq \log (Z/Z_\odot) \leq -0.3$ ($Z_\odot = 0.01524$; Caffau et al. 2011). The interstellar metallicity is set to equal to the stellar metallicity with dust-to-metal mass ratio fixed to $\xi_d = 0.3$ for ease of comparison with previous SED studies (although we note we will take a different approach when modelling the NIRSpec data). We put an upper limit on metallicity of $0.5 Z_\odot$ to avoid uncommon solutions of near-solar metallicity in reionization-era galaxies. The ionization parameter (the ratio of the number density of ionizing photons to the number density of hydrogen atoms in the H II regions; Penston et al. 1990) $U$ is adjusted in the range $-4.0 \leq \log U \leq -1.0$. We adopt log-uniform priors for $Z$ and $U$. We assume the Small Magellanic Cloud (SMC) extinction curve (Pei 1992) to account for the dust attenuation and the V-band optical depth $\tau_V$ is allowed to vary between 0.001 and 5 with a log-uniform prior. Finally, we adopt the prescription of Inoue et al. (2014) to include the absorption of IGM. When fitting the SEDs, we remove fluxes in filters that lie blueward of Ly$\alpha$ to avoid introducing the uncertain flux contribution from Lyman series emission and absorption. The best-fitting model with observed SED of each source is presented in Fig. 4.
The SED fitting results allow us to examine the slit-loss correction adopted on NIRSpec spectra. Here we compare the [O III] λ5007 fluxes (which is the brightest emission line with the highest S/N at rest-frame optical in our sample) inferred from the best-fitting SED models and the slit-loss corrected [O III] λ5007 fluxes measured by NIRSpec. We find that both fluxes are consistent within 1σ confidence interval, indicating that the slit-loss correction adopted in this work is reliable.

The physical properties of the models that reproduce the observed SEDs are shown in Table 2. The SEDs suggest minimal dust reddening, with V-band optical depths ranging between τ_V of 0.003 and 0.12, with a median of 0.03. The inferred CSFH ages are very young, ranging between 1.5 and 27 Myr, with a median (5 Myr) that is 2 – 14 times lower than the typical age at these redshifts (e.g. Tacchella et al. 2022; Endsley et al. 2023a; Whiter et al. 2023a; Endslsey et al. 2023b; Whiter et al. 2023b). The young ages reflect SEDs dominated by a recent burst or upturn in star formation history (SFH), with the light in the rest-frame UV and optical powered by a very young stellar population (though AGN can also contribute; see Larson et al. 2023). At these young ages, the rest-frame optical continuum is weak and emission lines are strong, leading to significant colour excesses from very large EW rest-frame optical emission lines. The rest-frame [O III] + Hβ EWs implied by the SEDs range between 879 and 3276 Å, with a median of 1994 Å. These [O III] + Hβ EWs reveal that our CEERS spectroscopic sample lies at the more extreme part of the ϵ ≳ 7–9 population, with values that correspond to the upper 43 per cent of the [O III] + Hβ EW distribution (Endslsey et al. 2023b). The Lyα emitters (EW_Lyα > 9 Å) in our sample show even larger [O III] + Hβ EWs = 1786–3276 Å, corresponding to the upper 11 per cent of the [O III] + Hβ EW distribution. This suggests that our CEERS spectroscopic sample is not representative of the ϵ ≳ 7–9 population, a result that is not surprising given that selection on [O III] is necessary for redshift confirmation. This bias must be taken into account when interpreting the spectroscopic properties of the sample.

The derived stellar masses range between 6.2 × 10^6 and 4.9 × 10^8 M_⊙ in our fiducial CSFH BEAGLE models. These masses correspond to the very young stellar population that dominates the SED. As has been discussed in the recent literature (e.g. Roberts-Borsani, Ellis & Laporte 2020; Laporte et al. 2021; Tacchella et al. 2022; Whiter et al. 2023a, b), older populations can easily be hidden underneath the light of the young stellar population, increasing the stellar mass by over an order of magnitude. To determine upper bounds on the stellar mass in these galaxies, we fit the SEDs using models that incorporate non-parametric SFHs. We follow an approach that is very similar to that described in Whiter et al. (2023a). In brief, we fit the SEDs with PROSPECTOR (Leja et al. 2019; Johnson et al. 2021), which is based on the Flexible Stellar Population Synthesis code (Conroy, Gunn & White 2009; Conroy & Gunn 2010) and the nebular emission models of Byler et al. (2017). We adopt a Chabrier (2003) IMF with a mass range of 0.1–300 M_⊙.

In this work, we choose 9 Å as the EW threshold of Lyα emitters because it equals to the 5σ Lyα EW limit that can be detected in a UV-bright galaxy (M_V < −21, or equivalently NICMaf/F150W < 26) at z = 8 with CEERS NIRSpec observations.
a SMC dust attenuation law (Pei 1992) with a log-uniform prior of V-band optical depth $\tau_V$, and the Inoue et al. (2014) IGM attenuation model, which are similar to the CSFH SED fitting with BEAGLE. Here, we consider models with non-parametric SFH in PROSPECTOR, which are piecewise constant functions in time. We adopt eight age bins spanning from the time of observation to the lookback time corresponding to a formation redshift $z_{\text{form}}$. We allow $z_{\text{form}}$ to vary between 15 and 30. The two most recent time bins of the SFH are fixed to 0–3 Myr and 3–10 Myr, and the remaining six bins are spaced evenly in logarithmic lookback time. For non-parametric SFH prior, we use the built-in ‘continuity’ prior in PROSPECTOR and also a ‘bursty’ version of the continuity prior (Tacchella et al. 2022; Whitlet al. 2023b). The continuity prior allows a smoothly evolving star formation rate (SFR) over time, while the bursty continuity prior allows more sharp changes in SFR between time bins. The results reveal stellar masses assuming continuity prior (bursty continuity prior) that range between $4.6 \times 10^7$ and $4.9 \times 10^7 M_\odot$ ($1.0 \times 10^7–5.4 \times 10^7 M_\odot$), roughly $\sim$ (2) times larger than the CSFH models. We list these values in Table 2. While our results do not depend sensitively on the stellar masses, we will consider the range between the BEAGLE CSFH and PROSPECTOR non-parametric SFH models (with continuity prior) as the allowed range for each source.

The observed SEDs also constrain the hydrogen ionizing photon production efficiency ($\xi_{\text{ion}}$), equal to the ratio of the hydrogen ionizing photon production rate ($N_{\text{ion}}$) and the far-UV continuum luminosity at rest-frame 1500 Å ($L_{1500}$). Because ionizing photons are reprocessed into recombination lines, $\xi_{\text{ion}}$ also sets the production efficiency of Ly$\alpha$ photons. Galaxies with larger $\xi_{\text{ion}}$ will have larger Ly$\alpha$ luminosities for fixed SFR, potentially boosting their visibility. In this paper, we define $\xi_{\text{ion}}$ as the hydrogen ionizing photon production rate per unit intrinsic $L_{1500}$, the observed UV luminosity at rest-frame 1500 Å (including nebular and stellar continuum) corrected for dust attenuation (see Chevallard et al. 2018 and Tang et al. 2019 for definitions of various $\xi_{\text{ion}}$). This is the most commonly used definition of $\xi_{\text{ion}}$ in literature. The derived values of the ionizing photon production efficiency range between log($\xi_{\text{ion}}$/erg^{-1} Hz^{-1}) $=25.61$ and 26.00. At lower redshifts, $\xi_{\text{ion}}$ is known to correlate with the derived [O III]+H$\beta$ EW (Chevallard et al. 2018; Tang et al. 2019; Onodera et al. 2020). In Fig. 5, we plot the inferred ionizing photon production efficiencies versus [O III]+H$\beta$ EW, overlaying the $z \geq 7$ results on the $z \simeq 2$ extreme emission line galaxy (EELG) sample from Tang et al. (2019). While the $z \geq 7$ galaxies are consistent with the lower redshift relation, they are situated in a region of the diagram that is rarely seen in lower redshifts samples of EELGs. These results suggest that the $z \geq 7$ Ly$\alpha$ emitters are among the most efficient ionizing agents known in normal star-forming galaxies. This result is not surprising given the very young stellar populations which dominate the SEDs of these systems. At sufficiently young ages under a CSFH, the B star population has yet to build up its contribution to the UV continuum luminosity, leading to very large values of $\xi_{\text{ion}}$. We will discuss implications of the ionizing production for Ly$\alpha$ visibility in Section 4.

In the end of this section, we explore the impact of non-zero ionizing photon escape fraction ($f_{\text{esc}}$) to the physical parameters derived from BEAGLE SED fitting. For galaxies to be the primary ionizing agents that are responsible for cosmic reionization, the ionizing photon escape fraction is required to be $f_{\text{esc}} \simeq 0.1–0.2$ (e.g. Robertson et al. 2015; Finkelstein et al. 2019; Naidu et al. 2020). We fit the SEDs of the 12 galaxies with JWST/NIRCam or high S/N HST/Spitzer photometry with CSFH BEAGLE models (assuming ionization-bounded nebula) while varying $f_{\text{esc}}$ as a free parameter. The parameter spaces of other physical quantities (e.g. stellar mass, age, dust attenuation, ionization parameter, and metallicity) remain the same and we allow $f_{\text{esc}}$ to vary in the range 0–0.9, assuming a uniform prior. The SED fitting results show $f_{\text{esc}} = 0.08–0.6$ with a median value of $f_{\text{esc}} = 0.2$, but the uncertainty of each fitted $f_{\text{esc}}$ is large (with a typical 1σ confidence interval of ±0.2). Other physical parameters remain consistent with the values derived from BEAGLE models with $f_{\text{esc}} = 0$ except for stellar age. The stellar ages inferred from non-zero $f_{\text{esc}}$ models are on average 0.2 dex younger than those inferred from models with $f_{\text{esc}} = 0$. This is because when some of the ionizing photons produced in galaxies escape directly into the IGM (and thus are not reprocessed into nebular emission), it requires younger stellar populations to produce more ionizing photons to reproduce the same nebular emission line strength.

2.3 Rest-frame optical emission lines in $z \geq 7$ galaxies

In this section, we characterize the rest-frame optical line ratios and discuss what they tell us about the ionized gas properties of our spectroscopic sample. The reader is also directed to a variety of other papers on the rest-frame optical lines in the reionization era (e.g. Arrabal Haro et al. 2023; Bunker et al. 2023; Cameron et al. 2023; Fujimoto et al. 2023; Hisao et al. 2023; Jung et al. 2023; Larson et al. 2023; Mascia et al. 2023; Nakajima et al. 2023; Sanders et al. 2023a; Saxena et al. 2023a; Sanders et al. 2023b; Saxena et al. 2023b). Our primary interest is in the nature of the galaxies in the CEERS spectroscopic sample, with a particular focus on those with Ly$\alpha$ emission (which we will come back to discuss in Sections 3 and 4). A detailed examination of the ionized gas properties of the general population at $z \geq 7$ is beyond the scope of the current study and will eventually require deeper data that is less biased to strong [O III] emitters. In what follows, we first consider the gas properties
implied by photoionization model fits. We then quantify the restframe optical line ratios of our sample, comparing to galaxies at lower redshifts.

A summary of the CEERS rest-frame optical emission line measurements is given in Table 3. We detect the [O III]λ4959, 5007 doublet in all 21 sources and Hβ in 20 sources in our sample. In addition 11 (7) galaxies, we detect the [O II] (Ne III) lines, allowing a range of ionization and excitation-sensitive line ratio diagnostics to be derived. We detect Hβ (Hβ) in 9 (3) galaxies, allowing constraints on Balmer line ratios. In a further two sources (CEERS-1019 and CEERS-1027), we detect the auroral [O III]λ4363 (S/N = 3–4), with tentative [O III]λ4363 features (S/N = 2) seen in additional two galaxies (CEERS-689 and CEERS-698). When combined with the other strong lines, the auroral lines allow constraints on the gas-phase oxygen abundance via the direct method (e.g. Izotov et al. 2006; Sanders et al. 2020). In two Lyα emitting galaxies (CEERS-1027 and CEERS-698), we also detect broad emission components. The intrinsic full-widths at half maximum (FWHM) after subtracting the spectral resolution measured by the instrument team (R ~ 1000, corresponding to FWHM = 300 km s⁻¹; Jakobsen et al. 2022) for the broad [O III]λ5007 lines of these two galaxies (Fig. 6) are 375 and 529 km s⁻¹, which comprise 21 per cent and 37 per cent of the total [O III]λ5007 fluxes. Such broad [O III] emission may indicate highly ionized gas outflows driven by stellar feedback associated with very young stellar populations (see Table 2), which is similar as seen in a young (~3 Myr), Lyman continuum emitting galaxy at lower redshift (Mainali et al. 2022). Fast outflows may also scatter Lyα photons redwards as they backscatter off outflowing gas (e.g. Verhamme, Schaerer & Maselli 2006; Steidel et al. 2010), as we will discuss in Section 2.4.

We first must consider the dust attenuation corrections for the emission line ratios. In Section 2.2, we demonstrated that the SEDs of our sample are typically best fit with little to no dust reddening.

Table 3. Rest-frame optical emission line ratios of z ≥ 7 galaxies identified in CEERS NIRSpec spectra. Line ratios constraining ionizing spectra (O32 and Ne3O2), oxygen abundance and excitation (O3 and R23), dust attenuation (Hα/Hβ), and electron temperature ([O III]λ4363/[O III]λ5007) are provided. For galaxies without significant [O II] or [Ne III] measurements, 3σ upper limits are given for O32, R23, or Ne3O2 ratios. In the last row, we provide the line ratios measured from the composite spectrum of the 16 MR grating sources at z > 7.

<table>
<thead>
<tr>
<th>ID</th>
<th>zspec</th>
<th>O3</th>
<th>O32</th>
<th>R23</th>
<th>Ne3O2</th>
<th>Hα/Hβ</th>
<th>[O III]λ4363/[O III]λ5007</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEERS-24</td>
<td>8.999</td>
<td>3.80 ± 1.95</td>
<td>&gt;3.98</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-23</td>
<td>8.881</td>
<td>4.78 ± 1.68</td>
<td>8.12 ± 3.15</td>
<td>7.09 ± 2.50</td>
<td>0.25 ± 0.26</td>
<td>0.668 ± 0.355</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1025</td>
<td>8.678</td>
<td>6.46 ± 1.01</td>
<td>18.06 ± 5.99</td>
<td>8.98 ± 1.42</td>
<td>1.08 ± 0.48</td>
<td>0.467 ± 0.123</td>
<td>0.025 ± 0.007</td>
</tr>
<tr>
<td>CEERS-1019</td>
<td>8.610</td>
<td>5.25 ± 2.42</td>
<td>4.81 ± 1.25</td>
<td>8.02 ± 3.71</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1149</td>
<td>8.175</td>
<td>6.99 ± 1.35</td>
<td>18.77 ± 6.06</td>
<td>9.81 ± 1.92</td>
<td>0.77 ± 0.38</td>
<td>0.497 ± 0.179</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-27</td>
<td>7.819</td>
<td>5.84 ± 2.32</td>
<td>&gt;3.68</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1023</td>
<td>7.776</td>
<td>4.59 ± 2.33</td>
<td>2.80 ± 1.25</td>
<td>7.97 ± 4.12</td>
<td>&lt;0.82</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-686</td>
<td>7.74</td>
<td>5.25 ± 0.41</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-689</td>
<td>7.545</td>
<td>6.00 ± 0.99</td>
<td>10.13 ± 2.67</td>
<td>8.84 ± 1.48</td>
<td>0.71 ± 0.46</td>
<td>0.333 ± 0.174</td>
<td>0.059 ± 0.032</td>
</tr>
<tr>
<td>CEERS-698</td>
<td>7.470</td>
<td>6.41 ± 0.60</td>
<td>14.72 ± 4.40</td>
<td>9.03 ± 0.88</td>
<td>0.87 ± 0.49</td>
<td>0.358 ± 0.085</td>
<td>0.014 ± 0.008</td>
</tr>
<tr>
<td>CEERS-1163</td>
<td>7.448</td>
<td>3.67 ± 2.91</td>
<td>&gt;4.20</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1038</td>
<td>7.194</td>
<td>3.79 ± 1.18</td>
<td>&gt;4.43</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-498</td>
<td>7.18</td>
<td>3.37 ± 0.30</td>
<td>&gt;23.45</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-499</td>
<td>7.168</td>
<td>3.66 ± 3.38</td>
<td>&gt;4.33</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-44</td>
<td>7.10</td>
<td>4.52 ± 0.50</td>
<td>&gt;18.93</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-407</td>
<td>7.028</td>
<td>6.55 ± 1.96</td>
<td>5.96 ± 3.91</td>
<td>10.00 ± 3.13</td>
<td>&lt;0.59</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1102</td>
<td>7.00</td>
<td>4.09 ± 0.58</td>
<td>2.12 ± 0.26</td>
<td>8.01 ± 1.16</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-717</td>
<td>6.932</td>
<td>6.07 ± 1.97</td>
<td>5.58 ± 2.76</td>
<td>9.19 ± 3.06</td>
<td>&lt;0.30</td>
<td>0.935 ± 0.421</td>
<td>...</td>
</tr>
<tr>
<td>CEERS-1143</td>
<td>6.928</td>
<td>9.68 ± 1.42</td>
<td>14.46 ± 5.70</td>
<td>14.25 ± 2.15</td>
<td>0.36 ± 0.36</td>
<td>0.390 ± 0.286</td>
<td>...</td>
</tr>
<tr>
<td>Composite</td>
<td>7.7</td>
<td>6.82 ± 0.26</td>
<td>17.84 ± 1.71</td>
<td>9.42 ± 0.36</td>
<td>0.89 ± 0.13</td>
<td>0.463 ± 0.038</td>
<td>0.022 ± 0.005</td>
</tr>
</tbody>
</table>

Figure 6. [O III]λ5007 velocity profiles of CEERS-1027 (top panel) and CEERS-698 (bottom panel). Broad emission line components are detected for both objects. The observed line flux (uncertainty) is shown by black lines (grey shaded regions). Each line profile is fitted by a narrow and a broad Gaussian profile, which are shown by cyan dashed lines. The blue lines present the sum of the narrow and the broad Gaussian profiles.
of the stellar continuum, with median UV slope $\beta = -2.4$ and median $\tau_V = 0.03$ in the BEAGLE models. Only 1 of 21 sources (CEERS-1023) shows evidence of significant reddening in its rest-frame UV photometry with UV slope $\beta = -0.9$. The blue UV slopes we find in the majority of the sample are consistent with expectations for $z \approx 7$–9 galaxies (e.g. McLure et al. 2011; Dunlop et al. 2012; Finkelstei et al. 2012; Rogers, McLure & Dunlop 2013; Bouwens et al. 2014; Jiang et al. 2020; Bhatawdekar & Conselice 2021; Topping et al. 2022). Measurement of the hydrogen Balmer line ratios (i.e. $H_\gamma$/$H_\beta$) enables derivation of the attenuation facing the emission lines. Assuming the case B recombination and $T_{eb} = 10^4$ K gas, we expect the intrinsic $H_\gamma$/$H_\beta$ ratio to be 0.468 (Osterbrock & Ferland 2006). The presence of dust will preferentially attenuate $H_\gamma$ relative to $H_\beta$, decreasing the flux ratio. At $z \approx 2$, galaxies with similar [O III] EWs to those in our spectroscopic sample have negligible attenuation, with $H\alpha$/$H_\beta$ ratios nearly identical to the case B value (Tang et al. 2019). The $H_\gamma$/$H_\beta$ line ratios in individual galaxies in our sample are highly uncertain (median error = 017) and thus not surprisingly span a significant range (0.33–0.93). To more robustly assess the typical level of attenuation in our sample, we create a composite spectrum by stacking the NIRSpec MR grating spectra. We first shift individual spectra to the rest-frame using the systemic redshifts measured from [O III]$\lambda 5007$. Each spectrum is then interpolated to a common rest-frame wavelength scale of 1 Å and normalized by individual [O III]$\lambda 5007$ luminosities. Finally, the spectra are stacked using inverse variance weighted luminosities in each wavelength bin (i.e. weighted by $1/\sigma^2$ where $\sigma$ is the standard deviation of individual spectra).

We show the resulting $z \approx 7$–9 composite spectrum in Fig. 7. We measure $H_\gamma$/$H_\beta = 0.463 \pm 0.038$ in the composite, consistent with the case B recombination value and hence implying negligible attenuation ($E(B-V) = 0.02$; assuming the Cardelli, Clayton & Mathis 1989 curve). We also consider a stack of the 3 $z > 7$ Ly$\alpha$ emitting galaxies (CEERS-1019, CEERS-1027, and CEERS-698). In this composite, we measure $H_\gamma$/$H_\beta = 0.458 \pm 0.046$, also consistent with negligible attenuation ($E(B-V) = 0.04$). Based on these results, we will not apply dust corrections to our fiducial line ratio measurements. However, we will comment on the influence that such corrections may have on our results, considering in particular CEERS-1023, which shows modest reddening of its stellar continuum as mentioned above.

As a first exploration of the gas properties of our sample, we use BEAGLE (Chevallard & Charlot 2016) to fit available rest-frame optical emission lines including [O II] $\lambda 3727$, H$\gamma$; [O III] $\lambda 4363$, H$\beta$; [O III] $\lambda 4959$, and [O III] $\lambda 5007$. We use a similar BEAGLE SFH model setup that described in Section 2.2 and in Plat et al. (in preparation), allowing the ionization parameter $U$ and the interstellar metallicity $Z_{ISM}$ to vary in the range $-4 < \log U < -1$ and $-2.2 < \log (Z/Z_\odot) < 0.24$ with log-uniform priors, and the dust-to-metal ratio $\xi_d$ to vary in $0.1 < \xi_d < 0.5$ with uniform prior. The stellar population age is allowed to vary between 1 and 10 Myr with a log-uniform prior because the hydrogen-ionizing spectrum (hereafter ‘ionizing spectrum’) reaches steady state after 10 Myr of CSFH. In modelling the emission line ratios, we must take into account observations that have demonstrated that high-redshift galaxies are typically alpha-enhanced (e.g. Steidel et al. 2016; Strom et al. 2017, 2018; Shapley et al. 2019; Sanders et al. 2020; Topping et al. 2020; Cullen et al. 2021; Runco et al. 2021; Strom et al. 2022), as expected given the delay in iron production (relative to oxygen and other alpha elements). Given that it is the iron abundance in massive star atmospheres that regulates the emergent ionizing spectrum, whereas the oxygen acts as an important coolant in the $T \approx 10^4$ K ionized gas, this implies that the stellar population metallicity is likely to be lower than the oxygen abundance in the H II regions. Motivated by these results, we consider BEAGLE photoionization models where the stellar metallicity (tracing the iron abundance [Fe/H]) is set to 1/5 of the interstellar metallicity (tracing the oxygen abundance [O/H]). This corresponds to the theoretical limit for type II supernovae (Nomoto et al. 2006). We have also considered BEAGLE models where the stellar metallicity is matched to the interstellar metallicity, and we confirm that our main results are not strongly dependent on these assumptions. Finally, motivated by the Balmer line ratios in the composite spectra described above, we assume the emission lines are not attenuated, setting $\tau_V = 0$ in our fits. However, we will also explore how sensitive our results are to this assumption.

We summarize the BEAGLE model results in Table 4. The observed rest-frame optical flux ratios are generally well-reproduced by the models. As an example, we show the consistency between the observed O32 indices and the preferred model O32 values in the

\footnote{The comparison between metallicities and elemental abundances is based on a solar abundance pattern, i.e. the stellar metallicity relative to solar metallicity, $Z/Z_\odot$, is lower than the oxygen abundance relative to the solar oxygen abundance, $[O/H]/[O/H]_\odot$.}
left-hand panel of Fig. 8. We plot the model-preferred ionization parameters and gas-phase oxygen abundances in the middle and right-hand panels of Fig. 8, respectively. The spectroscopic sample is characterized by high ionization parameters (median log $U = -2.11$) and metal poor gas (median $12 + \log (O/H) = 7.84$). If we consider the eight galaxies with the largest O32 ratios in the sample (O32 > 10), we find models prefer similarly metal poor gas (median $12 + \log (O/H) = 7.84$) and even higher ionization parameters (log $U = -1.80$). Similar results are seen when we consider the 4 $z \geq 7$ galaxies with strong Lyα emission (EW $> 9 \AA$). Here, we find log $U = 2.14$ to $-1.31$ and 12 $+ \log (O/H) = 7.58$ to 7.85. By fitting the rest-frame optical emission lines measured from the composite MR grating spectrum, we also find a high ionization parameter (log $U = -1.76$) and metal-poor gas (12 $+ \log (O/H) = 7.74$). It is clear from this analysis that while the entire $z \geq 7$ sample is quite extreme in its nature, the Lyα emitting galaxies stand out as having among the most metal poor-galaxy and largest ionization parameters (see red symbols in Fig. 8). These model results are not strongly dependent on our assumption of zero attenuation. If we instead allow $A_V$ to take on a slightly larger value that is still consistent with the blue UV slopes in the SEDs (i.e. $A_V = 0.1$), we find that the gas-phase metallicities and ionization parameters change by less than 0.1 dex.

The gas-phase metallicities of our sources are consistent with those found in other studies of the $z \geq 7$ population with similar stellar masses (e.g. Cameron et al. 2023; Curti et al. 2023; Nakajima et al. 2023; Sanders et al. 2023b), and they are well below what is typically seen in UV-selected galaxies at $z \approx 2$ (12 $+ \log (O/H) = 8.1$ – 8.5; Steidel et al. 2016; Sanders et al. 2021). Taken at face value, these results suggest that very metal poor gas is common in $z \geq 7$ galaxies dominated by recent upturns in star formation (and/or combined with AGN contributions; e.g. Larson et al. 2023) that power strong [O III]$+\lambda 4959, 5007$ emission.

The gas-phase oxygen abundance can also be constrained with the direct method in the two galaxies with robust detections of the [O III]$\lambda 4363$ auroral emission line (CEERS-1019, CEERS-1027). We follow the procedures in Izotov et al. (2006) to derive the direct $T_e$ oxygen abundance. We use the observed [O III]$\lambda 4363$/[O III]$\lambda 5007$ ratios to compute the O$^+$ zone electron temperature ($T_e$[O III]) with PYTHON package PyNeb. We assume an electron density of $n_e = 250$ cm$^{-3}$ (Sanders et al. 2016), and we have confirmed that the derived electron temperature and oxygen abundance change negligibly over $n_e = 100$–1000 cm$^{-3}$. Since we do not have auroral [O II]$\lambda\lambda 7320, 7330$ detections, we derive the O$^+$ zone electron temperature $T_e$(O II) using the relation from Campbell, Terlevich & Melnick (1986) and Garnett (1992): $T_e$(O II) = 0.7x $T_e$(O III) + 3000 K. Then the O$^+$ abundance and O$^+$ abundance are derived using $T_e$ and $n_e$ with [O III]$\lambda\lambda 4959, 5007$/H$\beta$ and [O III]$\lambda\lambda 3726, 3797$/H$\beta$ ratios. The final oxygen abundance is $O/H = O^+$/H$^+$ + O$^+$/$H$. The results indicate very metal poor H II regions, with values consistent with the photoionization modeling described above. For CEERS-1019, the direct method suggests 12 $+ \log (O/H) = 7.73$ (0.10 $Z_\odot$), where the solar metallicity corresponds to a gas-phase oxygen abundance 12 $+ \log (O/H) = 8.71$; Gutkin et al. 2016), whereas the BEAGLE models suggest 12 $+ \log (O/H) = 7.77$. For CEERS-1027, the direct method indicates 12 $+ \log (O/H) = 7.74$ (0.11 $Z_\odot$), which is again similar to the preferred BEAGLE model value (12 $+ \log (O/H) = 7.58$) within 1σ confidence interval. The oxygen abundances of these two galaxies are also consistent with the values derived from direct-$T_e$ method reported in Sanders et al. (2023b).

---

9http://research.iac.es/proyecto/PyNeb/
We also gain insight into the nature of our sample through comparison of line ratios to other JWST measurements at $z \gtrsim 7$ and those at lower redshifts. The O32 index is sensitive to the ionization state of the nebular gas, providing a probe of the ionization parameter. The Ne3O2 index is nearly equivalent as O32 since both neon and oxygen are alpha elements and Ne$^{++}$ and O$^{++}$ have similar ionization potentials (e.g. Pérez-Montero et al. 2007; Levesque & Richardson 2014; Wittok et al. 2021). Therefore, Ne3O2 is also sensitive to the ionization conditions of the H II regions (e.g. Levesque & Richardson 2014; Strom et al. 2018), with the benefit of being slightly less dependent on the attenuation correction owing to the proximity of the two emission lines in wavelength. In Fig. 9, we plot the O32 and Ne3O2 ratios of the $z \gtrsim 7$ sample in CEERS. The O32 values range between 2.12 and $\sim 30.22$. We find an average value of $O32 = 17.84$ from the composite spectrum (Fig. 7), consistent with the average O32 found in $z \gtrsim 7$ galaxies with JWST in literature (O32 $\sim 10–20$; e.g. Cameron et al. 2023; Mascia et al. 2023; Sanders et al. 2023a; Saxena et al. 2023b. This is about 4 x larger than values seen at $z \sim 3–5$ (O32 $\sim 3–5$; e.g. Christensen et al. 2012; Troncoso et al. 2014; Shapley et al. 2023) and about an order of magnitude larger than typical values seen at $z \sim 2$ (O32 $\sim 1–2$; e.g. Sanders et al. 2016; Steidel et al. 2016; Sanders et al. 2021), suggesting that the $z \gtrsim 7$ spectroscopic sample is comprised of galaxies with extreme ionization conditions, consistent with the high ionization parameters preferred by BEAGLE. The Ne3O2 ratios suggest a similar picture, with an average value (Ne3O2 = 0.89) measured from composite spectrum (Fig. 7) that is consistent with $z \gtrsim 7$ measurements in literature (Ne3O2 $\sim 1$; e.g. Cameron et al. 2023; Trump et al. 2023) and well above the average at $z \sim 2$ (Ne3O2 = 0.10–0.15; Steidel et al. 2016; Sanders et al. 2021) and at $z \sim 3–5$ (Ne3O2 = 0.2–0.4; Christensen et al. 2012; Troncoso et al. 2014; Shapley et al. 2017; Wittok et al. 2021; Shapley et al. 2023). These conclusions are unchanged if we apply a modest attenuation correction ($r_v = 0.1$) to the emission lines. The Ne3O2 ratios do not change at all, O32 is moderately reduced (average O32 = 14.31) but still well above the typical values at $z \sim 2$.

Both O32 and Ne3O2 are known to increase with the rest-frame optical emission line EWs (e.g. Tang et al. 2019; Sanders et al. 2020). In Fig. 10, we plot the O32 ratios of our sample versus the [O III]+H$\beta$ EW for the 12 galaxies where we have robust SED constraints and O32 measurements. We overlay the $z \gtrsim 7$ results on the $z \sim 2$ relation derived from spectroscopic follow-up of EELGs (Tang et al. 2019). At $z \sim 2$, it is seen that O32 increases from 2 at [O III]+H$\beta = 300$ Å to O32 = 20 at [O III]+H$\beta = 3000$ Å, reflecting the tendency for ionization parameter to be larger in galaxies with very young stellar populations (see discussion in Tang et al. 2019 and Plat et al. in preparation). The $z \gtrsim 7$ measurements are consistent with the $z \sim 2$ relation, albeit mostly sampling the large O32 ratios associated
with the highest [O III]+Hβ EWs (>1000 Å). In the context of O32 – [O III]+Hβ EW relation, the presence of large O32 ratios in our sample largely reflects the young stellar population ages (and associated large [O III]+Hβ EWs) of the sources in the spectroscopic sample.

Based on Fig. 10, the redshift evolution in the distribution of O32 (and other ionization-sensitive line ratios) will depend significantly on the evolution in the [O III]+Hβ EW distribution. Recent work has shown that the [O III]+Hβ EW of the overall population increases between z ≃ 2 and z ≃ 7, going from typical values of 250 Å in relatively massive galaxies at z ≃ 2 (e.g. Reddy et al. 2018; Boyett et al. 2022) to 780 Å at z ≃ 7 (e.g. Endsley et al. 2021a, 2023b). This trend will result in larger O32 ratios being more common at z ≃ 7, undoubtedly contributing somewhat to the larger O32 ratios seen in our CEERS spectroscopic sample. Recent ALMA observations have also revealed that the ionization-sensitive far-infrared [O III]88 μm/[C II]158 μm ratios are ubiquitously high in z > 6 galaxies, with values larger than in lower-redshift galaxies (e.g. Carniani et al. 2020; Harikane et al. 2020). It is likely that [O III]88 μm/[C II]158 μm also correlates with [O III]+Hβ EWs (e.g. Wittek et al. 2022). But the bias in our spectroscopic sample towards larger [O III]+Hβ EWs (as discussed in Section 2.2) further shifts our O32 distribution to extreme values. If z > 7 galaxies follow a similar relation between O32 and [O III]+Hβ EW as is seen at z ≃ 2 (as appears to be the case in Fig. 10), then we would expect representative z > 7 galaxy samples to show O32 extending 2–3 × lower than the values seen in the CEERS sample. Deeper rest-frame optical spectra will be required to obtain representative line ratio distributions for the full population of z > 7 galaxies.

We also quantify the [O III]λ5007/Hβ flux ratio (O3) and the R23 index (([O II]λ3727+O3)/Hβ) in Table 3. We measure an average O3 value of 6.8 from the composite spectrum of our CEERS z > 7 NIRSpec sample (Fig. 7). This is consistent with the average O3 ratios measured in z > 7 galaxies in literature (O3 ≃ 5–7; e.g. Cameron et al. 2023; Nakajima et al. 2023; Trump et al. 2023; Sanders et al. 2023a). It is only marginally greater than the O3 seen in typical z < 2 galaxies (4.3 in KBSS, 3.4 – 4.5 in MOSDEF; Steidel et al. 2016; Sanders et al. 2021), where the range quoted in the MOSDEF sample corresponds to line ratios for stars with stellar masses between 10^9 and 10^10 M⊙. The extreme ionization conditions of the z > 7 galaxies contribute to the larger O3 values, but this is countered by the lower oxygen abundance of the z > 7 galaxies in our sample. The net effect is that O3 ratios are only slightly larger than those at z ≃ 2. In Fig. 11, we plot the R23 indices of the z > 7. We find R23 values that range between 6 and 14, with an average of 9.4 measured from the composite spectrum (Fig. 7). These are comparable to the R23 values seen in z ≃ 2 star-forming galaxies (R23 = 8.5 in KBSS and R23 = 8.9–9.1 for M⋆ = 10^8–10^9 M⊙ galaxies in MOSDEF; Steidel et al. 2016; Sanders et al. 2021) and other z > 7 galaxies (R23 ≃ 7–10; e.g. Cameron et al. 2023; Mascia et al. 2023; Sanders et al. 2023a; Saxena et al. 2023b). The origin of the similar R23 indices (in spite of very different gas-phase properties) is analogous to the explanation for O3. The lower metallicity of the z > 7 galaxies increases the excitation (boosting collisionally excited lines relative to Hβ, but the effect this has on R23 is countered by the reduced oxygen abundance. Thus in the R23 versus O3 plane (Fig. 11), we primarily see the z > 7 galaxies shift towards larger O3 at fixed R23.

In summary, the CEERS NIRSpc observations have revealed a population of galaxies at z ≃ 7 that is very different than that at z ≃ 2, with significantly more metal poor gas and larger ionization parameters. The differences are most clear in the ionization-sensitive line ratios. The O32 values in our spectroscopic sample are often
over an order of magnitude larger than those typical at \( z \approx 2 \). We find similar line ratios in samples of galaxies at lower redshifts with comparable [O III] + H\( \beta \) EWs as those in our sample. These include the most extreme tail of the EELG population at \( z \approx 2 \) (e.g. Tang et al. 2019; Du et al. 2020) and a subset of galaxies in the LzLCS survey (Flury et al. 2022). We will come back to directly compare the Ly\( \alpha \) properties of these samples to the \( z \gtrsim 7 \) galaxies in Section 4.

### 2.4 Rest-frame UV emission lines in \( z \gtrsim 7 \) galaxies

The emission lines in the rest-frame UV of \( z \gtrsim 7 \) galaxies also provide insights into the Ly\( \alpha \) emission in the reionization era and the ionizing sources. NIRSpec data not only probe the Ly\( \alpha \) EWs, but they also constrain the velocity profile and transmission of Ly\( \alpha \), both valuable diagnostics for assessment of the impact of the IGM on the line. The fainter UV metal emission lines (i.e. C III], C IV) and He II provide information on the stars, gas, and AGN activity in the reionization-era population, allowing investigation of whether \( z \gtrsim 7 \) sources with Ly\( \alpha \) are different from the general population. We describe the analysis associated with each of these measurements in this section.

We detect Ly\( \alpha \) emission in six \( z \gtrsim 7 \) galaxies in the CEERS spectroscopic sample (see Table 5). Two of the six (CEERS-1027, Fig. 12, CEERS-44, Fig. 2) are newly confirmed as Ly\( \alpha \) emitters in these observations. Both are strong Ly\( \alpha \) emitters for \( z \gtrsim 7 \) (e.g. Endsley et al. 2021b), with rest-frame EWs of 77.6 \( \AA \) (CEERS-44) and 20.4 \( \AA \) (CEERS-1027). The other four are previously confirmed Ly\( \alpha \) emitters, including CEERS-1019 (Zitrin et al. 2015), CEERS-1029 (Larson et al. 2022; tentative Ly\( \alpha \) detection found in NIRSpec spectrum), CEERS-686 (Jung et al. 2022), and CEERS-698 (Roberts-Borsani et al. 2016; Stark et al. 2017). We notice that the Ly\( \alpha \) emission lines of CEERS-1019, CEERS-1029, and CEERS-698 have lower S/N (\( \gtrsim 3-4 \); Fig. A2). To test whether the emission features of these three objects are due to noise fluctuations, we follow the test in Chen et al. (in preparation) that randomly place 1000 circular apertures with radius =3 pixels (consistent with the aperture we calculate Ly\( \alpha \) flux) near the Ly\( \alpha \) emission feature of each object and measure the aperture fluxes. For CEERS-1019 and CEERS-698, we find that no random aperture shows flux comparable to the emission feature. This indicates that the emission features of these two objects are more likely to be Ly\( \alpha \) than random fluctuations. For CEERS-1029, there are a few random apertures with fluxes approaching the significance of the emission feature. Because the emission feature of CEERS-1029 lies at the expected position of Ly\( \alpha \) (after taking into account the difference between air and vacuum wavelengths) based on the Ly\( \alpha \) redshift reported in (Larson et al. 2022), we recognize it as a tentative Ly\( \alpha \) emission. The low S/N Ly\( \alpha \) detections of these three objects demonstrate that we are pushing the limits of NIRSpec spectra in CEERS observations. Future, deep spectroscopy will help to improve these measurements.

### Figure 12. Ly\( \alpha \) emission line of the newly detected Ly\( \alpha \) emitter CEERS-1027. The Ly\( \alpha \) spectrum has been shifted to the rest-frame. The grey shaded regions show the uncertainty of flux. Medium-resolution Ly\( \alpha \) spectra of the four objects with Ly\( \alpha \) detections (CEERS-1019, CEERS-1149, CEERS-1027, and CEERS-698) are shown in Fig. A2.
Each of CEERS-1019, CEERS-1029, CEERS-686, and CEERS-698 show UV continuum emission in the CEERS spectra, allowing us to compute Lya EWs. Here we compute Lya EW using the average continuum flux between rest-frame 1225 and 1255 Å, which minimizes the contribution from nearby absorption features (e.g. Kornei et al. 2010; Stark et al. 2010). The derived values range between 4.2 (CEERS-1029) and 41.9 Å (CEERS-686). In three of these four galaxies, the Lya EWs are consistent (within 1σ uncertainties) with the ground-based measurements from MOSFIRE. In the case of CEERS-1019, the offset is just outside of the 1σ confidence interval. The ground-based Lya EW measurement is $28^{+14}_{-11}$ Å, with the large uncertainty due to the presence of an OH sky line near the Lya feature. The NIRSpec observations suggest the EW is much lower (9.8 ± 2.4 Å). The NIRSpec micro-shutter is centred on the two brightest clumps of the galaxy, as revealed in NIRCam imaging. It is possible that the MOSFIRE observations may have picked up additional Lya emission associated with a fainter clump outside of the NIRSpec micro-shutter. Given the impact of the sky line on the ground-based Lya measurement, we will adopt the NIRSpec value as fiducial in our analysis, but we will comment on how our results would change if the MOSFIRE value is correct.

The transmission of Lya through a partially neutral IGM depends on the velocity profile of the line as it emerges from the galaxy. Lya is typically observed with a positive velocity offset relative to the systemic redshift of its host galaxy, owing largely to resonant scattering effects (whereby the redshifted component mostly has its origin in line photons that are backscattered from outflowing gas on the far side of the galaxy). The larger the velocity offset ($\Delta V_{\text{sys}}$), the further Lya photons will be shifted into the damping wing profile before they encounter neutral hydrogen. If velocity offsets in galaxies at $z \gtrsim 7$ are large enough, the transmission of Lya will be less impacted by the neutral IGM (e.g. Dijkstra, Mesinger & Wyithe 2011; Stark et al. 2017; Mason et al. 2018b; Hashimoto et al. 2019; Endsley et al. 2022).

Observational efforts to quantify the velocity offsets of reionization-era galaxies have long been stunted by the challenge of measuring systemic redshifts with non-resonant emission lines. A small number of $z \gtrsim 6$ $\Delta V_{\text{sys}}$ measurements have been made via detection of the faint C III] emission line (Stark et al. 2015, 2017; Hutchison et al. 2019). Recent studies with ALMA have made further progress via detections of [C II]158μm and [O III]8μm fine structure lines (e.g. Willott et al. 2015; Pentericci et al. 2016; Carniani et al. 2017; Hashimoto et al. 2019; Endsley et al. 2022). JWST spectroscopy has the potential to make velocity offset work much easier in the reionization era given the ease of efficiently detecting the strong rest-frame optical emission lines in large samples.

CEERS spectroscopy provides MR ($R \sim 1000$) rest-frame optical emission line detections for four sources with Lya emission, allowing us to estimate the Lya velocity offsets. We calculate systemic redshifts ($z_{\text{sys}}$) by simultaneously fitting the centroids of Hβ, [O III]λ4959, and [O III]λ5007 emission lines, and we calculate Lya redshifts using the wavelength of the peak of Lya emission in the NIRSpec spectra. The derived Lya velocity offsets are on average large for these four objects, with 323 km s$^{-1}$ for CEERS-1027, 458 km s$^{-1}$ for CEERS-1019, 545 km s$^{-1}$ for CEERS-686, and 1938 km s$^{-1}$ for CEERS-1029. The Lya velocity offsets of CEERS-1027, CEERS-1019, and CEERS-686 are consistent with those measured for $z > 6$ galaxies with similar M$_{UV}$ (Fig. 13), while the Lya velocity offset of CEERS-1029 is much larger, and we will discuss this object further in Section 3. If the broad [O III]λ5007 emission detected in CEERS-1027 and CEERS-686 is driven by outflows, the Lya velocity offsets of these two galaxies are found to be similar to two times of the outflow velocity measured from [O III]λ5007 line width ($2v_{\text{out}} = 318$ km s$^{-1}$ for CEERS-1027 and $=449$ km s$^{-1}$ for CEERS-698, where $v_{\text{out}} = \text{FWHM}_{\text{broad}}/2.355$). This is consistent with the predictions of the ‘shell-model’ of Lya radiation transfer for Lya emission backscattering off an expanding shell (Verhamme et al. 2006). We note that due to the M of NIRSpec grating, the uncertainty of Lya line centre measurement and hence the velocity offset is large ($\Delta V_{\text{sys}} \approx 0.005$) for three sources except CEERS-1027, which has the brightest Lya emission with the highest S/N (=7) among these four sources.

To explore how the large Lya velocity offsets of $z > 7$ Ly$\alpha$ emitting galaxies impact the IGM transmission, we estimate the damping wing optical depth of Ly$\alpha$ as a function of velocity offset (Miralda-Escudé 1998) using the methods described in Endsley et al. (2022). We consider these galaxies are in ionized regions, assuming the IGM inside the ionized region is completely ionized and outside is completely neutral. For CEERS-1027 ($\Delta V_{\text{sys}} = 332$ km s$^{-1}$), CEERS-1019 ($\Delta V_{\text{sys}} = 458$ km s$^{-1}$), and CEERS-686 ($\Delta V_{\text{sys}} = 545$ km s$^{-1}$), if they are in relatively large ionized regions ($R = 1$ pMpc) the Ly$\alpha$ transmission through the neutral IGM can reach up to 50 per cent at their velocity offsets. Even in very small ionized regions ($R = 0.1$ pMpc) their Ly$\alpha$ transmission is still significant ($\sim 15–20$ per cent) at velocity offset $\Delta V_{\text{sys}} \approx 300–500$ km s$^{-1}$, and their Ly$\alpha$ visibility can still be boosted together with their enhanced ionizing photon (and hence Ly$\alpha$) production efficiencies (Fig. 5). For CEERS-1029 ($\Delta V_{\text{sys}} = 1938$ km s$^{-1}$), the Ly$\alpha$ transmission at its offset
velocity offset is very large (70 per cent) even in a small ionized region with $R = 0.1$ pMpc. This will contribute significantly to its Lyα visibility. However, we note that although the IGM transmission can be significant at the measured Lyα velocity offsets of CEERS $z > 7$ Lyα emitting galaxies, in Section 4 we will discuss how the observed Lyα EWs may be lower than expected. This could imply that most of the Lyα flux was emitted at bluer wavelengths and attenuated by the IGM damping wing, thus only emission from the reddest part of the intrinsic Lyα profile is visible.

We also use the CEERS spectra to calculate Lyα escape fractions ($f_{\text{esc, Ly}\alpha}$). We use the new Hβ detections to estimate the Lyα flux expected in our spectra in absence of resonance scattering and dust absorption. For case B, recombination and $T_e = 10^4$ K gas, the intrinsic Lyα/Hβ ratio is expected to be 24.9 (Osterbrock & Ferland 2006). The Lyα emitters in our sample have observed ratios that range between 0.87 and 8.44. The inferred $f_{\text{esc, Ly}\alpha}$ of the six galaxies with Lyα detections are 0.035–0.339 with a median of 0.073. The low implied Lyα escape fractions are consistent with expectations for typical UV-bright galaxies at $z \approx 2$ (e.g. Hayes et al. 2011; Ciardullo et al. 2014; Matthee et al. 2016; Sobral et al. 2017) and also the Lyα emitting galaxies at $z \approx 6$ (Ning et al. 2023), but we will show in Section 4 that they are significantly lower than is typically found in lower redshift galaxies with similar rest-frame optical line spectra.

After Lyα, the next strongest emission lines in the rest-frame UV are fainter collisionally excited features from highly ionized metal species (i.e. C III, C IV, and O III)). These rest-frame UV metal lines are high energy transitions, requiring hard ionizing flux and large electron temperature (and hence metal poor gas) for significant excitation. They appear most prominently in UV spectra powered by very young stellar populations (under CSFH) that have yet to build up significant far-UV continuum luminosities from B stars (e.g. Erb et al. 2010; Stark et al. 2014; Rigby et al. 2015; Senchyna et al. 2017; Berg et al. 2018; Mainali et al. 2020; Tang et al. 2021a). Previous work has revealed a handful of intense C III and O III) emitters in $z \gtrsim 6$ galaxies (e.g. Stark et al. 2015; Laporte et al. 2017; Mainali et al. 2017; Stark et al. 2017; Hutchison et al. 2019; Jiang et al. 2021; Topping et al. 2021), with rest-frame EWs ($\approx 10–20$ Å) that are more than an order of magnitude larger than what is common at $z \approx 1–3$ (e.g. Shapley et al. 2003; Steidel et al. 2016; Du et al. 2017; Maseda et al. 2017).

One of the most intense C III) emitters at $z \gtrsim 7$ is EGS-zs8-1 (EW =22 Å; Stark et al. 2017), the brightest galaxy in the $z = 7.7$ Lyα association in the EGS field. While this galaxy was not targeted with CEERS spectroscopy, the data set includes observations of many similarly intense optical line emitters, allowing us to determine if such strong C III) is common at $z \gtrsim 7$. We detect C III) emission (S/N =2–4) in three galaxies in the NIRSpec observations (CEERS-1019, 1049, and 1027) with respect to the systemic redshift, respectively. We use the peak of entire Lyα feature to compute the velocity offset relative to the $z_{\text{sys}}$, which is 1938 km s$^{-1}$ (i.e. equivalent to the velocity offset of the first component).

Table 6. EWs of rest-frame UV emission lines detected in CEERS NIRSpec $z > 7$ galaxies.

<table>
<thead>
<tr>
<th>ID</th>
<th>$z_{\text{spec}}$</th>
<th>EW(C IV,1549) (Å)</th>
<th>EW(He II, III) (Å)</th>
<th>EW(O III,1661) (Å)</th>
<th>EW(C IV,1908) (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEERS-1019</td>
<td>8.678</td>
<td>...</td>
<td>...</td>
<td>8.0 ± 4.9</td>
<td>10.9 ± 6.0</td>
</tr>
<tr>
<td>CEERS-1149</td>
<td>8.175</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>16.0 ± 4.4</td>
</tr>
<tr>
<td>CEERS-1027</td>
<td>7.819</td>
<td>5.7 ± 2.1</td>
<td>3.9 ± 1.6</td>
<td>...</td>
<td>12.9 ± 4.5</td>
</tr>
</tbody>
</table>

Figure 14. C III) EW as a function of O32 for CEERS NIRSpec galaxies at $z > 7$. LAEs (EW(C IV) > 9 Å) and non-LAEs at $z > 7$ are shown by red stars and blue circles, respectively. As a comparison, we overplot literature data of galaxies at $z \approx 1$–3 (open cyan symbols; Pettini et al. 2000; Quider et al. 2009; Erb et al. 2010; Christensen et al. 2012; Rigby et al. 2021; Quider et al. 2009; Pettini et al. 2000; Mainali et al. 2020; Tang et al. 2021a). Previous work has revealed a handful of intense C III and O III) emitters in $z \gtrsim 6$ galaxies (e.g. Stark et al. 2015; Laporte et al. 2017; Mainali et al. 2017; Stark et al. 2017; Hutchison et al. 2019; Jiang et al. 2021; Topping et al. 2021), with rest-frame EWs ($\approx 10–20$ Å) that are more than an order of magnitude larger than what is common at $z \approx 1–3$ (e.g. Shapley et al. 2003; Steidel et al. 2016; Du et al. 2017; Maseda et al. 2017).

One of the most intense C III) emitters at $z \gtrsim 7$ is EGS-zs8-1 (EW =22 Å; Stark et al. 2017), the brightest galaxy in the $z = 7.7$ Lyα association in the EGS field. While this galaxy was not targeted with CEERS spectroscopy, the data set includes observations of many similarly intense optical line emitters, allowing us to determine if such strong C III) is common at $z \gtrsim 7$. We detect C III) emission (S/N =2–4) in three galaxies in the NIRSpec observations (CEERS-1019, 1049, and 1027) with respect to the systemic redshift, respectively. We use the peak of entire Lyα feature to compute the velocity offset relative to the $z_{\text{sys}}$, which is 1938 km s$^{-1}$ (i.e. equivalent to the velocity offset of the first component).
Figure 15. UV diagnostics and the position of $z \geq 7$ CEERS galaxies. The left-hand panel shows the O III] $\lambda$1666/He II $\lambda$1640 diagnostic and the position of CEERS-1027 (the cyan star). We overplot the line ratio predictions from photoionization models for AGN (magenta squares; Feltre, Charlton & Gutkin 2016; Mignoli et al. 2019) and star-forming galaxy (orange stars; Gutkin et al. 2016), and the line ratios of extreme metal-poor galaxies at low redshift (green symbols; Berg et al. 2016; Senchyna et al. 2017, 2019; Berg et al. 2019a). The UV line ratios of CEERS-1027 are consistent with local metal-poor star-forming galaxies and fall in the region where massive stars are likely to be the dominant ionizing source. The right-hand panel shows the C III] $\lambda\lambda$1907, 1909/He II $\lambda$1640 versus C III] EW diagram. The black dashed lines separate the regions dominated by AGN (left), composite of AGN and star formation (middle), and star formation (right) predicted by nebular emission models calculated in Hirschmann et al. (2019). Models are calculated using prescriptions for star-forming galaxies (Gutkin et al. 2016), AGN (Feltre et al. 2016), and quiescent galaxies (Hirschmann et al. 2017). The three galaxies at $z \geq 7$ with C III] detections (CEERS-1019, CEERS-1149, and CEERS-1027) fall in the composite region. We overplot the diagnostics of star-forming galaxies at $z \sim 2$ (open cyan symbols; Stark et al. 2014; Du et al. 2020; Tang et al. 2021a), local star-forming galaxies (grey symbols; Berg et al. 2016; Senchyna et al. 2017, 2019; Berg et al. 2019a), and $z \sim 2$–3 AGN (open red hexagon) from Hainline et al. (2011).

temperature is larger in metal-poor gas, which results in higher O$^{++}$ to O$^{+}$ ratios as O$^{+}$ is the most efficient coolant (e.g. Kewley, Nicholls & Sutherland 2019). Notably, the three LAEs with Ly$\alpha$ EW $> 9$ Å are situated in the upper right of the C III] EW versus O32 diagram, again underlying their extreme nature relative to the galaxy population seen at $z \sim$ 7–9. We verify that the BEAGLE model fits described in Section 2.3 can reproduce the C III] EWs. When we add C III] to the optical lines in our BEAGLE model fits, we find that the inferred metallicities remain similarly low, whereas the ionization parameters increase modestly (median change of 0.3 dex).

One galaxy in our sample (CEERS-1027 at $z = 7.819$) additionally shows faint detections of C IV (EW $= 5.7$ Å) and He II (EW $= 3.9$ Å) emission (Table 6; see Fig. A3 for spectra). The C IV and He II line widths are narrow ($\sigma = 65$ km s$^{-1}$) and consistent with the widths of the other emission lines, so we interpret these as nebular in nature (i.e. not broad stellar features). Powering both lines requires a hard ionizing spectrum, with significant ionizing flux in excess of 48 eV (C IV) and 54 eV (He II). At lower redshifts, it has been shown that both are often associated with metal poor galaxies (e.g. Stark et al. 2014; Berg et al. 2016; Senchyna et al. 2017; Berg et al. 2018; Senchyna et al. 2019; Berg et al. 2019a; Tang et al. 2021a; Saxena et al. 2022; Senchyna et al. 2022; Schaerer et al. 2022a), although the precise origin of the line emission may require contributions from additional ionizing sources (e.g. X-ray binaries, or AGN; e.g. Hainline et al. 2011; Volonteri et al. 2017; Plat et al. 2019; Berg et al. 2019b; Saxena et al. 2020a, b; Berg et al. 2021; Olivier et al. 2022). The optical lines indicate that CEERS-1027 is extremely metal poor ($12 + \log (O/H) = 7.58$), consistent with the values that are usually associated with C IV and He II emitters. To investigate the origin of the line emission in CEERS-1027, we compare the UV line ratios to expectations from photoionization models in the left-hand panel of Fig. 15. The C IV/He II ratio ($\sim 1.4$) versus O III] $\lambda$1666/He II ratio ($< 1.0$) is consistent with metal poor galaxies seen locally. The line ratios fall in the region of the diagram where massive stars are likely to be the dominant ionizing source, although the data do not rule out mixtures of AGN and massive stars, or contributions of energetic photons from X-ray binaries or intermediate mass black holes.

For the other two C III] emitting galaxies in our sample (CEERS-1019 at $z = 8.678$ and CEERS-1149 at $z = 8.175$), we can investigate their ionizing sources in the C III] EW—C III]/He II diagram (e.g. Nakajima et al. 2018; Hirschmann et al. 2019) together with CEERS-1027. In the right-hand panel of Fig. 15, we show the C III] EW versus C III]/He II of the three CEERS galaxies with intense C III] emission. Based on the photoionization models studied in Hirschmann et al. (2019), these three galaxies fall in the region where the line emission is likely dominated by a combination of massive stars and AGN. CEERS-1019 and CEERS-1149 present similar C III] EWs and C III]/He II ratios as star-forming galaxies at $z \sim$ 0–3 (e.g. Stark et al. 2014; Berg et al. 2016; Senchyna et al. 2017, 2019; Berg et al. 2019a; Du et al. 2020; Tang et al. 2021a), but again the data do not rule out a significant contribution of ionizing photons from AGN (e.g. for CEERS-1019; Larson et al. 2023).

3 REIONIZATION-ERA LYMAN-ALPHA EMITTERS

The EGS field contains three groups of Ly$\alpha$ emitting galaxies at $z > 7$ (see Fig. 16). The visibility of Ly$\alpha$ in these associations has led to the suggestion that they trace rare ionized bubbles in the mostly neutral IGM. However, the Ly$\alpha$ detections may alternatively reflect properties internal to the galaxies (i.e. efficient Ly$\alpha$ production). Prior to JWST, it was challenging to study these galaxies in any detail. Since redshift confirmation was mostly limited to those systems with strong Ly$\alpha$ emission, it was impossible to identify the extent to which Ly$\alpha$
visibility was enhanced in surrounding galaxies. In this section, we describe what CEERS spectroscopy has revealed about the galaxies with and without Lyα in the three associations.

### 3.1 z = 7.5 association of Lyα emitters

The CEERS data set includes observations of CEERS-698, the brightest galaxy ($H_{160} = 25.2$, $M_{UV} = -21.9$) in the $z \approx 7.5$ association and two additional galaxies (CEERS-689 and CEERS-1163) with redshifts that place them within 3.1 pMpc of CEERS-698. The latter two systems do not show Lyα emission. Below we summarize the properties of each of these galaxies.

#### 3.1.1 CEERS-698

CEERS-698 was first identified in Roberts-Borsani et al. (2016) as a UV-bright dropout (EGS-zs8-2) with an IRAC [4.5] excess suggesting large [O III]+Hβ EW ($\approx 2564\ A$; Table 2), a young light-weighted age (3.8 Myr), and very efficient ionizing photon production ($\langle [\text{H}\alpha/\text{Hz}\rangle = 25.8$). The source was spectroscopically confirmed with Keck/MOSFIRE via detection of Lyα (EW = 9.3 Å) at $z_{\text{Ly}\alpha} = 7.477$ in Stark et al. (2017), following tentative identification in Roberts-Borsani et al. (2016). CEERS/NIRSpec observations detect a suite of rest-frame optical lines in CEERS-698, revealing line ratios that reflect the extreme ionizing nature of the SED. The large O32 (14.72) and Ne3O2 (0.87) ratios are rarely seen in normal star-forming galaxies at lower redshifts, pointing to gas with a very high ionization parameter ($\log U = -1.85$). Photoionization models suggest metal poor gas ($12 + \log (O/H) = 7.81$), consistent with the direct value implied by the tentative detection of [O III]λ4363 (see Table 4).

Lyα is detected with NIRSpec with similar flux and EW (9.5 Å) to that measured previously with Keck. Using the systemic redshift derived from the rest-frame optical lines ($z_{\text{sys}} = 7.470$), we are able to compute a Lyα velocity offset of 545 ± 184 km s$^{-1}$ (Fig. 13). The NIRSpec detections of Lyα and Hβ indicate that the transmission of Lyα is very low (0.045 ± 0.015), suggesting the majority of Lyα is scattered out of the NIRSpec micro-shutter. While such low Lyα escape fractions are common in $z \approx 2$ galaxies (e.g. Hayes et al. 2011; Ciardullo et al. 2014; Matthee et al. 2016; Sobral et al. 2017), we will show in Section 4 that lower redshift galaxies with properties similar to CEERS-698 often transmit a much larger fraction of their Lyα.

#### 3.1.2 CEERS-689

This is a multicomponent galaxy with integrated H-band magnitude ($H_{160} = 24.9$, $M_{UV} = -22.1$). The NIRSpec microshutter covers one of the clumps with $H_{160} = 26.2$. The strong rest-frame optical lines are detected ([O II], [Ne III], Hγ, Hβ, [O III]) indicating a systemic redshift of $z_{\text{sys}} = 7.545$. This places this galaxy 3.1 pMpc away from CEERS-698. The ionization-sensitive rest-frame optical line ratios (O32 = 10.13 and Ne3O2 = 0.71) point to a high ionization parameter, which is also recovered in the photoionization model fits ($\log U = -2.02$). The gas-phase metallicity implied by the photoionization model fits is low ($12 + \log (O/H) = 7.84$), consistent with the direct method value suggested by the tentative detection of [O III]λ4363 (see Table 4). In spite of the intense ionizing nature of this source, we detect no Lyα with NIRSpec, suggesting relatively weak Lyα (EW <26.1 Å) and correspondingly low transmission of Lyα through the micro-shutter (<0.055), with both corresponding to 3σ limits.

We note that this system was recently discussed in Jung et al. (2022) with an ID of z8_32350. A faint emission feature (S/N = 4.6) in Keck spectroscopy yielded tentative Lyα confirmation of this galaxy at a slightly higher redshift ($z = 7.759$). Since there are multiple clumps in this system, we have confirmed that the NIRSpec micro-shutter centroid is very close (0.06 arcsec) to the...
coordinates reported in Jung et al. (2022). Given the confidence in the NIRSpec redshift (from numerous rest-frame optical lines), there is a disagreement with the redshift reported in Jung et al. (2022).

3.1.3 CEERS-1163

This fainter galaxy ($H_{160} = 26.8$, $M_{UV} = -20.2$) is detected in [O III] and Hβ, indicating a systemic redshift of $z = 7.448$ that places this system 1.2 pMpc away from CEERS-698 and 3.6 pMpc from CEERS-689. No Lyα is seen in the NIRSpec spectrum. Comparing the Hβ flux and the Lyα upper limit, we find a 3σ upper limit on the Lyα escape fraction of $<0.145$. Deeper spectroscopy is required to provide a more stringent constraint on whether this faint galaxy may have enhanced Lyα emission.

3.2 $z = 7.7$ association of Lyα emitters

EGS-zs8-1 was the first galaxy identified in the $z = 7.7$ association of Lyα emitters in the EGS. The source was spectroscopically confirmed at $z_{sys} = 7.730$ by Oesch et al. (2015) following photometric selection as a bright ($H_{160} = 25.0$) galaxy with an IRAC excess in Roberts-Borsani et al. (2016). Detection of the [C III], C III] doublet (Stark et al. 2017) enabled confirmation of the Lyα velocity offset (340 km s$^{-1}$). Subsequent work by Tilvi et al. (2020) and Jung et al. (2022) has identified 4 more galaxies at $z = 7.63$−7.83 with confident (>3σ) Lyα detections. The CEERS observations provide rest-frame UV to optical spectra of CEERS-686, one of the Lyα emitters from Jung et al. (2022; z8,13573). CEERS spectra also present two new confirmations of galaxies at similar redshifts ($z = 7.82$ and $z = 7.76$), one of which shows Lyα (CEERS-1027) and one of which does not (CEERS-1023). We summarize the properties of these three galaxies below.

3.2.1 CEERS-686

This is one of the fainter galaxies in the CEERS $z \gtrsim 7$ spectroscopic sample ($H_{160} = 26.4$, $M_{UV} = -20.7$). Jung et al. (2022) present Keck/MOSFIRE detection of strong Lyα emission at $z_{sys} = 7.7482$, revealing one of the largest Lyα EWs (69.1$^{+29.8}_{-20.9}$ Å) known at $z > 7$. The CEERS prism spectrum detects the [O III] doublet and Hβ, confirming a systemic redshift of $z_{sys} = 7.74$ and placing this galaxy only 0.6 pMpc away from EGS-zs8-1. We also detect Lyα (and underlying UV continuum) in the CEERS spectrum, with line flux similar to that measured with MOSFIRE. The implied Lyα EW (41.9 ± 1.6 Å) is somewhat lower than the Keck measurement, but still consistent within the 1σ uncertainties. The Lyα/Hβ ratio in the NIRSpec spectrum indicates that a large fraction of the Lyα emission is transmitted through the NIRSpec microshutter (0.325 ± 0.028). This confirms that the large Lyα EW is not simply due to efficient ionizing photon production, but the transmission is also enhanced. Such strong Lyα emitters tend to have a large fraction of their Lyα profile escaping near the systemic redshift (e.g. Hashimoto et al. 2013; Erb et al. 2014; Shibuya et al. 2014; Tang et al. 2021b), but owing to the low resolution of the prism spectrum, we are unable to extract a meaningful Lyα velocity offset for this galaxy.

3.2.2 CEERS-1027

This moderately faint ($H_{160} = 26.5$, $M_{UV} = -20.6$) galaxy is covered in the NIRCam footprint, with an SED that indicates a young stellar population (3.4 Myr), efficient ionizing photon production (log $[\xi_{ion}/\text{erg}^{-1}\text{ Hz}] = 25.8$), and very strong nebular emission ([O III]+Hβ EW =276 Å). The NIRSpec spectrum reveals numerous emission lines, confirming the systemic redshift ($z_{sys} = 7.819$) and placing this galaxy 5.2 pMpc from EGS-zs8-1. The line ratios of CEERS-1027 are the most extreme in the sample, with both O32 (>30.22) and NeO2 (>1.80) pointing to gas with very extreme ionization conditions. The BEAGLE models reproduce the lines with very low metallicity (12 + log (O/H) = 7.58) and high ionization parameter (log U = −1.31). Tentative detection of [O III],4363 reveals a similarly low metallicity (12 + log (O/H) = 7.74) using the direct method. The UV spectrum reveals reasonably strong Lyα (EW =20.4 Å) and lower S/N detections of C IV (EW =5.7 Å), He II (EW =3.9 Å), and C III] (EW =12.9 Å). As discussed in Section 2.4, the UV line ratios are consistent with powering by low metallicity stars. The Lyα emerges with a velocity offset of 323 ± 18 km s$^{-1}$ with respect to the systemic redshift (Fig. 13). In spite of the large Lyα EW, we infer a relatively low escape fraction for Lyα (0.085 ± 0.018), suggesting the vast majority of the line emission is not making its way through the NIRSpec micro-shutter.

3.2.3 CEERS-1023

This galaxy is reasonably faint ($H_{160} = 26.2$, $M_{UV} = -20.9$) and red (β = −0.9), suggesting significant attenuation. We detect [O III], Hβ, and [O II] with NIRSpec, revealing a systemic redshift ($z_{sys} = 7.776$) that places this galaxy 2.0 pMpc away from EGS-zs8-1. The observed value of O32 (2.80) is among the lowest in the sample, suggesting less extreme ionization conditions. If we were to apply a reddening correction consistent with the UV slope, the O32 ratio would be even lower (i.e. O32 = 2.31 for E(B − V) = 0.16). We do not detect Lyα, with the 3σ upper limit implying a Lyα escape fraction of <0.135. While deeper spectroscopy will help put more robust constraints on the presence of Lyα, the non-detection is not surprising given the very red colours seen in HST.

3.3 $z = 8.7$ association of Lyα emitters

The EGS contains two very bright galaxies with Lyα emission at $z \approx 8.7$ (Zitrin et al. 2015; Larson et al. 2022). These are among the highest redshift galaxies with robust Lyα detections. HST WFC3/IR imaging suggests that there may be an overdensity in the region (Leonova et al. 2022), potentially contributing the formation of an early ionized structure that is aiding in the visibility of Lyα. The CEERS observations provide spectra of both $z \approx 8.7$ Lyα emitters (CEERS-1019 and CEERS-1029), while also confirming two additional galaxies (CEERS-23, CEERS-1025) at the same redshift. We describe the properties of the four galaxies below.

3.3.1 CEERS-1019

This bright ($H_{160} = 24.9$, $M_{UV} = -22.4$) galaxy was identified in Roberts-Borsani et al. (2016) and spectroscopically confirmed with Lyα emission at $z_{sys} = 8.683$ in Zitrin et al. (2015). NIRCam imaging reveals it is a multicomponent galaxy with three separate clumps, as is common for bright galaxies at these redshifts (e.g. Chen et al. 2023). The CEERS NIRSpec observations reveal numerous strong rest-frame optical lines, confirming a systemic redshift of $z_{sys} = 8.678$. The NIRCam SED indicates the galaxy is dominated by a very young stellar population (4.9 Myr) with powerful nebular emission ([O III]+Hβ EW =2597 Å) and extremely efficient ionizing photon production (log $[\xi_{ion}/\text{erg}^{-1}\text{ Hz}] = 26.0$), though AGN can
also contribute (Larson et al. 2023). A 4.6σ detection of N VI,1243 has been reported in Mainali et al. (2018), suggesting a hard ionizing spectrum which might be powered by AGN or fast radiative shocks. In NIRSpec spectra neither N VI,1238 or N VI,1243 is detected with 3σ upper limit of EW of $<$4.8 Å, though the NIRSpec upper limit does not rule out the detection of N VI,1243. Both O32 (18.06) and Ne3O2 (1.08) point to very high ionization gas. To reproduce the line ratios, the photoionization models require both low metallicity (12 + log(O/H) = 7.77) and high ionization parameter (log $U = -1.74$). The direct method gives a similar value for the metallicity (12 + log(O/H) = 7.73). The UV shows low S/N detections of O III],1661 (EW = 8.0 ± 4.9 Å) and CIII] (EW = 10.9 ± 6.0 Å). Lyα is seen in the NIRSpec data with a large velocity offset to the systemic redshift ($\Delta v_{\text{sys}} = 458$ km s$^{-1}$), but with moderately lower EW ($9.8 \pm 2.4$ Å) than measured in Zitrin et al. (2015). As discussed in Section 2.3, we adopt the NIRSpec value in our analysis owing to the impact of a sky line on the MOSFIRE spectrum. The escape fraction of CEERS-1019 is low ($f_{\text{esc}, \text{Ly} \alpha} = 0.035 \pm 0.010$), suggesting the majority of Lyα is scattered out of the NIRSpec micro-shutter.

### 3.3.2 CEERS-1029

This bright galaxy ($H_{160} = 25.6$, $M_{\text{UV}} = -21.6$) was confirmed at $z_{\text{sys}} = 8.665$ via detection of Lyα emission in Larson et al. (2022). The Lyα EW (4.7 Å) is very low, likely signalling that this galaxy transmits a very small fraction of its line emission. The NIRSpec spectrum confirms the [O III] doublet, Hβ, and [O II], revealing a systemic redshift of $z_{\text{sys}} = 8.610$ with a distance of 3.9 pMpc away from CEERS-1019. This O32 ratio in CEERS-1029 (4.81) is not as large as most galaxies in our sample. The BEAGLE models suggest that this galaxy is more metal-rich (12 + log(O/H) = 8.01) than the other Lyα emitters, with a slightly lower ionization parameter (log $U = -2.33$). The NIRSpec spectrum reveals a tentative line emission at observed frame 11 758 Å, which recovers the Lyα emission detected in Larson et al. (2022) with nearly identical EW (4.2 ± 1.3 Å). The transmission of Lyα implied by the Lyα/Hβ ratio is low ($f_{\text{esc}, \text{Ly} \alpha} = 0.061 \pm 0.034$), as expected given the very low Lyα EW. The velocity offset between Lyα and the rest-frame optical lines is extremely large (1938 ± 162 km s$^{-1}$; Fig. 13). How Lyα is shifted to such large velocities is not entirely clear. Recent work has demonstrated that Lyα is very broad in similarly luminous reionization-era galaxies, with hints of multiple redshifted components (Endsley et al. 2022). Based on comparison to [C II] line profiles, these authors have speculated that such line profiles could arise if the different gas-rich clumps have significant proper motions with respect to one another. Along these lines, it is conceivable that the Lyα seen in CEERS-1029 could be backscattered off of a clump with significant proper motion. But regardless of the precise origin, the large velocity offset of the Lyα that escapes in CEERS-1029 will certainly facilitate its transmission through the galaxy and IGM, allowing ~70 per cent of the Lyα at this velocity offset to transmit through the neutral IGM even in very small ionized region ($R = 0.1$ pMpc) as discussed in Section 2.4.

### 3.3.3 CEERS-23

This galaxy is very faint ($H_{160} = 28.4$, $M_{\text{UV}} = -18.9$) and is detected in [O III] and Hβ. The systemic redshift measured from rest-frame optical emission lines is $z = 8.881$, placing this galaxy 5.5 pMpc away from CEERS-1019 and 8.6 pMpc away from CEERS-1029. The NIRCam SED indicates a moderately large [O III]+Hβ EW (1175 Å) and young stellar populations (12 Myr). We do not detect Lyα emission in the NIRSpec spectrum. By comparing its Lyα flux upper limit to Hβ flux we put a 3σ upper limit on the Lyα escape fraction of $<0.422$. Thus deeper spectroscopy is required to constrain whether this very faint source has Lyα emission.

### 3.3.4 CEERS-1025

This moderately bright galaxy ($H_{160} = 26.2$, $M_{\text{UV}} = -21.1$) is covered by NIRCam observations, with an SED indicating very young stellar populations (1.6 Myr) and very efficient ionizing photon production (log [E_{ion}/erg$^{-1}$ Hz] = 25.8). NIRSpec detections of rest-frame optical lines confirm this source at $z_{\text{sys}} = 8.715$, with a distance of 1.4 pMpc from CEERS-1019 and 3.9 pMpc from CEERS-1029. Line ratios reveal moderately large O32 (8.12), with photoionization models suggesting a combination of metal poor gas (12 + log(O/H) = 7.75) and a reasonably high ionization parameter (log $U = -2.13$). The NIRSpec spectrum does not extend blue enough to provide constraints on Lyα in this galaxy, so we cannot place an upper limit on the escape fraction of Lyα.

### 4 DISCUSSION

In the previous section, we described NIRSpec observations of galaxies in three separate associations of Lyα emitters in the EGS field at $z \gtrsim 7$. In this section, we synthesize these new CEERS results and discuss implications for our understanding of the processes regulating Lyα visibility in the reionization era.

We first focus on the five strongest $z \gtrsim 7$ Lyα emitters (EW >9 Å) in our CEERS spectroscopic sample (see Table 5). This subset includes one of the highest redshift Lyα emitters (CEERS-1019; Zitrin et al. 2015), one of the largest EW Lyα emitters at $z \approx 7$ (CEERS-686; Jung et al. 2022), two newly confirmed Lyα emitters (CEERS-1027, CEERS-44), and a previously known Lyα emitting galaxy with strong rest-frame optical emission at $z = 7.5$ (CEERS-698; Roberts-Borsani et al. 2016; Stark et al. 2017). The rest-frame optical line ratios of the $z \gtrsim 7$ Lyα emitting galaxies are at the extreme end of the spectroscopic sample, with O32 values (14.72 to >30) that are only matched by nearby samples of LyC leakers (Izotov et al. 2018; Flury et al. 2022). The majority of Lyα emitting galaxies reported in Saxena et al. (2023b) also show similarly large O32 ratios (Fig. 11). For the subset of strong Lyα emitting galaxies with SED constraints, we find evidence for very large [O III]+Hβ EWs (1768 – 3276 Å) and very young stellar populations (3.4–10 Myr of CSFH) likely formed in a recent burst or upturn in star formation. As a result of the young and hot stellar populations (and/or combined with AGN contributions; e.g. Larson et al. 2023), the ionizing photon production efficiency in these galaxies (log [E_{ion}/erg$^{-1}$ Hz] = 25.7 – 26.0) is 1.4–3 × the average value at $z \approx 7$ – 8 (Prieto-Lyon et al. 2023; Endsley et al. 2023b). Because ionizing photons are reprocessed into recombination lines, the large values of $\xi_{\text{ion}}$ will translate into a higher-than-average production rate of Lyα photons (normalized to

---

1 This sample excludes the weaker Lyα emitting galaxy (CEERS-1029) identified in Larson et al. (2022). In addition to showing lower EW Lyα emission (4.2 Å), this galaxy also shows less extreme line ratios (O32 = 4.91). This galaxy stands out in its exceptionally large velocity offset (see Section 3.3), likely contributing to its visibility.
Figure 17. Lyα EW as a function of O32 for CEERS NIRSpec galaxies at $z > 7$ (red stars). Galaxies with Lyα emission detection are highlighted by solid stars. For those without Lyα emission detection, the upper limits are shown by open stars. We overplot data of $z > 6.5$ Lyα emitters (open red pentagons) from Saxena et al. (2023b), $z \sim 2$ EELGs (cyan squares) from Tang et al. (2021b), $z \sim 0$ Green Peas (green open circles) from Yang et al. (2017), and star-forming galaxies at $z \approx 0.2$–0.4 (grey diamonds) from LaLCS (Flury et al. 2022) as a comparison. The $z > 7$ galaxies tend to have on average lower Lyα EWs comparing to low redshift analogs at fixed O32.

Figure 18. Lyα escape fraction ($f_{\text{esc}, \text{Ly}}$) versus O32 for CEERS NIRSpec galaxies at $z > 7$ (red stars). Galaxies with Lyα emission detection are plotted as solid stars. For Lyα non-detected sources, upper limits are plotted as open stars. For CEERS-686 (i.e. z8_13573 in Jung et al. 2022), which does not have [O II] and hence an O32 measurement, its $f_{\text{esc}, \text{Ly}}$ is shown by the red dashed line. We overplot data of $z > 6.5$ Lyα emitters (open red pentagons) from Saxena et al. (2023b), $z \sim 2$ EELGs (cyan squares) from Tang et al. (2021b), $z \sim 0$ Green Peas (green open circles) from Yang et al. (2017), and star-forming galaxies at $z \approx 0.2$–0.4 (grey diamonds) from LaLCS (Flury et al. 2022) as a comparison. At fixed O32, $z > 7$ galaxies have on average lower $f_{\text{esc}, \text{Ly}}$ compared to $z \sim 0$–3 analogs.

A large $\xi_{\text{ion}}$ can still contribute significantly to the Lyα visibility even if the galaxy is leaking a large fraction of LyC emission (e.g. $f_{\text{esc}, \text{LyC}} \gtrsim 0.2$). This is because the conditions producing a high $f_{\text{esc}}$ (e.g. very young stellar populations) often produce a high $\xi_{\text{ion}}$ and a very large intrinsic Lyα EW ($\gtrsim$ 400 Å; e.g. Saxena et al. 2023a). Therefore, even if the observed Lyα EW is weakened due to the high $f_{\text{esc}}$, its value can still be high (e.g. Naidu et al. 2022).
0–2 from LzLCS (Flury et al. 2022), GPs (Yang et al. 2017), and z ≲ 2 EELGs (Tang et al. 2021b). The median Lyα EW of z = 0–2 galaxies with O32 > 10 is 120 Å. In contrast, the z > 7 galaxies with O32 > 10 and Lyα detections in the three associations have Lyα EWs in the range 10–20 Å. We note that this does not include CEERS-686, a stronger Lyα emitter (EW ≈ 41.9 Å) in the z ≈ 7.7 group (see Section 3.2) as it lacks a constraint on [O II]. We will return to discuss this source later in the section. But overall, these results suggest that the Lyα emitters at z ≈ 7–9 in the EGS field tend to be significantly weaker (in Lyα emission) than those with similar O32 indices in low redshift samples, with a typical downward shift of 6–12 × in the Lyα EW.

There are likely a variety of effects contributing to the absence of very strong Lyα emission in the z ≳ 7 galaxies with large O32. First, while we have only observed Lyα redshifted with respect to systemic, it is likely that some fraction of the photons may have been emitted in a blue peak, which would be resonantly absorbed by optically thick residual neutral gas in the ionized IGM (e.g. Gunn & Peterson 1965; Mason & Gronke 2020). In our comparison sample at z ≲ 2 about 30 percent of the total Lyα flux is emitted in a blue peak, similar to Lyα emitter samples at z ≲ 0–3, which find 23 percent of the flux is in a blue peak (Hayes et al. 2021). Thus, unless an unexpectedly large fraction of Lyα photons were emitted in a blue peak in our z ≳ 7 sample, the resonant absorption should not be the dominant cause of the reduction in Lyα EW at z ≳ 7. If the IGM at z ≳ 7 is mostly neutral, depending on the size of the ionized region around the three Lyα associations in the EGS (which span 5–10 pMpc) we may expect significant attenuation from the damping wing of the surrounding IGM.

We estimate the transmission of Lyα through the neutral IGM with different ionized region sizes at z = 7–9. We first build an ‘intrinsic’ Lyα profile (defined as the profile emerging from the ISM and CGM before being impacted by the neutral IGM) of z ≳ 7 galaxies. We use the composite Lyα profile of EELGs at z ≲ 2–3 in Tang et al. (2021b) as a baseline (Tang et al. in preparation), as this sample matches the large [O III]+Hβ EWs (> 1000 Å) and large O32 ratios (>6) seen in CEERS z ≳ 7 galaxies but is situated at redshifts where the IGM is highly ionized and thus the impact of neutral IGM is minimal. We assume that the composite Lyα profile of z ≳ 2–3 EELGs is the same as the Lyα profiles of z ≳ 7 galaxies before being attenuated by the neutral IGM. Then we compute the damping wing optical depth of Lyα at z = 7–9 as a function of velocity offset from systemic (Miralda-Escudé 1998) to the composite Lyα profile of z ≳ 2–3 EELGs to estimate the IGM transmission. We consider a galaxy at z = 8 that is situated in ionized regions with different radii, and we assume the IGM within the ionized region is completely ionized and outside is completely neutral. The results show that if the galaxy is situated in a relatively large ionized bubble (R = 1 pMpc), about 30 percent of the Lyα flux (emerging from the ISM and CGM) will transmit through the neutral IGM. If the galaxy is situated in smaller bubbles (R = 0.5 and 0.1 pMpc), the transmission of emerging Lyα flux will decrease to 20 percent and 6 percent, respectively. In order for neutral IGM to contribute significantly to the observed weak Lyα emission at z ≳ 7.5 (a downward shift of 6–12 × in Lyα EW), the ionized regions surrounding the Lyα emitters must be smaller (≲0.5–1 pMpc) than was previously thought. Such small bubbles may not be unexpected at such high redshifts where the IGM is likely significantly neutral (Lu et al. 2023). In this case, we would expect galaxies situated in between the two bright Lyα emitters to show much lower Lyα transmission. If the Lyα transmission is low in very small ionized bubbles, as suggested by the weak EWs, we can only see the reddest tail of the Lyα emission line which could help to explain why the Lyα velocity offsets of CEERS z ≳ 7 galaxies are large.

But other factors may be more important in driving down the Lyα EWs in the EGS field at z ≳ 7.5. While the low redshift comparison samples are similar to the z ≳ 7 galaxies in their rest-frame optical spectroscopic properties (i.e. O32, [O III] EW), their UV luminosities are not comparable. The LzLCS galaxies have absolute magnitudes in the range of MUV = −21.5 to −18.3 with a median of MUV = −19.9. In contrast, the z ≳ 7 galaxies in the Lyα associations are more luminous, spanning MUV = −22.4 to −20.2 with a median of −21.1. If galaxies with larger UV luminosities have more neutral hydrogen in their circumgalactic medium (CGM), we may expect more scattering of line photons within the galaxy, thereby decreasing the Lyα flux through the NIRSpec micro-shutter. This would also broaden and shift the lines redward (e.g. Verhamme et al. 2006; Yang et al. 2017). To test the role of luminosity-dependent effects on the Lyα properties, we plot the Lyα EW as a function of absolute magnitude in Fig. 19 for the z ≳ 7 galaxies and the low redshift comparison samples. Several things stand out. Galaxies with the largest luminosities do tend to have weaker Lyα in both comparison and z ≳ 7 samples. This trend is also seen in other z ≳ 7 Lyα emitters (e.g. Endsley et al. 2022; Bunker et al. 2023; Saxena et al. 2023b; see also Fig. 13) and in large Lyα emitter and Lyman break galaxy samples at lower redshifts (e.g. Ouchi et al. 2008; Stark et al. 2010; Schaerer, de Barros & Stark 2011; Hashimoto et al. 2017; Oyarzun et al. 2017). Yet at moderate (and lower) luminosities (MUV ≲ −21 to −19), there does appear to be a population of intense Lyα emitters (EW > 100 Å) that have yet to be observed at z ≳ 7, even in regions where the IGM is thought to be ionized. If the bubbles in the EGS are very large, we may expect to see very intense Lyα once larger spectroscopic samples of lower luminosity galaxies are assembled. These sources should also present much lower velocity offsets than those of our current z ≳ 7 sample (see Fig. 13).

Spectroscopic aperture effects must also be considered as a potential factor in the absence of strong Lyα at z ≳ 7. The small
area of the micro-shutters (0.20 × 0.46 arcsec) could potentially capture a smaller fraction of the resonantly scattered line emission. However, we have demonstrated consistency between the NIRSpec and Keck results (with the only exception being CEERS-1019 where there is significant skyline contamination), so based on existing data, it does not appear that the small size of the micro-shutters is leading to significant losses relative to slit-based spectroscopic surveys. However, it is conceivable that the HST/COS aperture may pick up more of the resonantly scattered Lyα emission from the z ≥ 0.3 LzLCS sources. Future investigation with the integral field mode of NIRSpec or NIRISS is needed to help clarify the impact of the small size of the micro-shutters on Lyα properties.

Finally, we consider two exceptions to the discussion above. CEERS-686 presents relatively strong Lyα (EW =41.9 Å) and a Lyα escape fraction (f_{esc,Lyα} = 0.325) that is very similar to what is seen in lower redshift Lyα emitters. This relatively faint galaxy (M_{UV} = −20.7) is located 0.6 pMpc behind the luminous galaxy EGS-zs8-1 (see Fig. 16 and Section 3.2). If EGS-zs8-1 is at the centre of a moderately small ionized region (≥1 pMpc in size), the transmission of CEERS-686 will be enhanced owing to its position on the far side of the bubble (see discussion in Jung et al. 2022).

The other exception to this discussion is CEERS-44, the strongest Lyα emitter in the NIRSpec sample (EW =77.6 Å) at z = 7.10. This galaxy is not in one of the known Lyα associations in the EGS. It is UV-faint (M_{UV} = −19.4) with a very large O32 value (>18.93) and strong [O III]+Hβ emission is evident in its NIRCam SED (EW =1786 Å). How Lyα is transmitted so readily in this galaxy is not immediately clear. There is no useful velocity offset constraint for this galaxy, as it was only observed with the low resolution prism. Future observations of the field are required to reveal whether there may be additional sources with similarly large Lyα transmission at this redshift.

Progress towards more quantitative estimates of the bubble sizes will require spectroscopic observations of the more numerous population of faint galaxies that span the distance between the known Lyα emitters in the EGS field. If the bubbles are very large (i.e. >1 pMpc) and the impact of the IGM on Lyα is minimal, we should expect to see very intense Lyα emission in a significant subset of faint galaxies. Such strong Lyα emitters should also present lower velocity offsets (≤ 200 km s^{-1}; e.g. Saxena et al. 2023a) than have been observed in this field at z ≥ 7 so far (Fig. 13). Failure to detect faint galaxies with large Lyα EWs and small velocity offsets would give indications that the IGM is significantly impacting the Lyα transmission in these ionized regions. While current Lyα samples in this field lack faint galaxies (see Fig. 19), the availability of deep NIRCam photometry over much of the EGS field will allow fainter galaxies to be efficiently targeted with future spectroscopic campaigns with NIRSpec.

5 SUMMARY

We discuss new JWST/NIRSpec observations of z ≥ 7 galaxies in the EGS field taken as part of the CEERS ERS program. This field has previously been shown to host three associations of Lyα emitting galaxies at redshifts (z = 7.5, 7.7, and 8.7) where the IGM is likely substantially neutral. In this work, we have detected rest-frame UV to optical emission lines and confirmed redshifts of 21 galaxies at z ≥ 7. Lyα emission lines have been detected for six galaxies in our sample, including four previously confirmed Lyα emitters (Zitrin et al. 2015; Roberts-Borsani et al. 2016; Stark et al. 2017; Jung et al. 2022; Larson et al. 2022) and two newly detected Lyα emitters at z > 7. Three (CEERS-1027, CEERS-686, CEERS-44) of these six galaxies present relatively high S/N (>7) Lyα detections and the other three (CEERS-1019, CEERS-1029, CEERS-698) show lower S/N (≈ 3–4) Lyα emission. By fitting the SEDs of a subset (12 out of 21) of our z ≥ 7 galaxies with JWST/NIRCam or high S/N HST+Spitzer photometry, we demonstrate that these galaxies have large [O III]+Hβ EWs (≈879–3276 Å) that are more extreme than the average values of z ~ 7–9 galaxies (≈ 700–800 Å; e.g. Labbé et al. 2013; De Barros et al. 2019; Endsley et al. 2021a, 2023b).

Using available emission line ratios, we discuss the properties of the ionizing spectra and nebular gas of galaxies in our z ≥ 7 NIRSpec sample. We summarize our results below.

(i) We present the O32 and NeO2 ratios of z ≥ 7 CEERS NIRSpec galaxies, providing insight into the ionization state of the nebular gas. The z ≥ 7 galaxies show that average O32 and NeO2 (O32 = 17.84 and NeO2 = 0.89 measured from the composite spectrum) are consistent with other JWST measurements at z ≥ 7 in literature and ≈10 × larger than typical star-forming galaxies at z ~ 2. The O32 versus [O III]+Hβ EW of CEERS z ≥ 7 galaxies is consistent with the the relation derived at lower redshifts, with the largest O32 values (>10) associated with the largest [O III]+Hβ EWs (≥1500 Å). Our sample is biased toward strong [O III] emission and therefore the large O32 ratios may not be representative of the full population. The Lyα emitters (EW_{Lyα} > 9 Å) in our sample show even larger O32 (>14) values than the non-LAEs, suggesting even more extreme ionization conditions in this population.

(ii) We interpret the ionized gas properties of our CEERS z ≥ 7 spectroscopic sample by fitting rest-frame optical emission lines with the BEAGLE tool. The model fits suggest high ionization parameters (log U = −2.59 to −1.31 with a median of −2.11) and low gas-phase oxygen abundance (12 + log (O/H) = 7.45 − 8.22 with a median of 7.84). For the small subset galaxies with tentative auroral [O III]4363 line detections, we derive very low gas-phase metallicities (12 + log (O/H) = 7.6 − 7.9) that are consistent with the values inferred from photoionization models. Deeper spectra will allow investigation of whether these properties extend to galaxies with lower [O III] EW which are absent from the current spectroscopic sample.

(iii) We detect C III] emission in three of the galaxies in the z ≥ 7 CEERS NIRSpec sample. Each of these galaxies also shows very large O32 values (>18), consistent with the trend between C III] EW and O32 found in lower redshift galaxies. We detect high ionization emission lines (C IV, He II) in one of our galaxies (CEERS-1027), indicating a very hard ionizing spectrum. The UV line ratios in this system are consistent with expectations for metal poor massive stars.

(iv) The CEERS NIRSpec sample at z ≥ 7 contains 10 galaxies in the 3 Lyα emitter groups (a z = 7.5, 7.7, and 8.7) in the EGS field. We identify one new Lyα emitter (CEERS-1027) situated 5 pMpc from the brightest galaxy in the z ≈ 7.7 association and demonstrate that not all galaxies at these redshifts show Lyα. The strong Lyα emitters (EW > 9 Å) have among the largest O32 ratios in the entire sample (14 to 30) suggesting very high ionization parameters. Their SEDs suggest very large [O III]+Hβ EWs (>1700 Å) indicating very young stellar populations (<10 Myr assuming CSFH) and efficient ionizing photon production efficiency (log [f_{ion/erg-1Hz}] = 25.7 − 26.0), potentially boosting the Lyα visibility.

(v) The z ≥ 7 CEERS NIRSpec galaxies present relatively low Lyα escape fractions and large Lyα velocity offsets. By comparing the Lyα flux to Hβ flux, we derive Lyα escape fraction ranging from 0.035 to 0.085, suggesting that a very small fraction of the Lyα emerges through the NIRSpec micro-shutter. These values are lower than the typical f_{esc,Lyα} of galaxies at lower redshifts with similarly large O32. The Lyα velocity offsets of Lyα emitting galaxies in our
z ≥ 7 sample range from 323 to 1938 km s⁻¹. Such large velocity offsets may help boost the transmission of Lyα through the IGM, contributing to their visibility.

(vi) We find that the Lyα emission in z ≥ 7 galaxies is much weaker than the Lyα of lower redshift galaxies with similar rest-frame optical spectral properties. The Lyα EWs (≥ 10–20 Å) of Lyα emitters in the 3 LAE associations at z ≃ 7.5–8.7 is 6–12 × lower than the median Lyα EW (≈ 120 Å) of galaxies with similarly large O32 (≥ 10) at z ∼ 0–2. Several effects likely contribute to this downturn of Lyα strength. If the ionized bubbles around the Lyα emitters at z ≥ 7 are relatively small (≤ 1 Mpc), we may expect a decrease of Lyα transmission even in ionized regions. Our z ≥ 7 Lyα emitters are also on average ∼1 mag brighter than z ∼ 0–2 analogs. If brighter galaxies have more H I in their CGM, the Lyα photons may also be more significantly scattered within galaxies, leading to weaker Lyα emission. Finally, the small area covered by NIRSpect MSA micro-shutter (0.20 × 0.46 arcsec²) may capture smaller fraction of the resonantly scattered Lyα line emission than lower redshift galaxies, such as LzLCS sample which uses HST/COS. Future observations of Lyα in fainter galaxies that span these fields should enable more quantitative constraints on the bubble sizes and associated impact of IGM on Lyα transmission.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for insightful comments. We thank the entire CEERS team for their effort designing and executing this program. MT acknowledges funding from the JWST Arizona/Steward Postdoc in Early galaxies and Reionization (JASPER) Scholar contract at the University of Arizona. DPS acknowledges support from the National Science Foundation through the grant AST-2109066. CAM and TYL acknowledge support by the VILUM FONDEN under grant 37459. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140. PS was generously supported by a Carnegie Fellowship through the Carnegie Observatories during the completion of this manuscript. LW acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2137419. The authors sincerely thank Ramesh Mainali for providing data of C III] versus O32 relation at z ∼ 0–3, John Chisholm for providing data of LzLCS sample, and Jacopo Chevallard for providing access to the BEAGLE tool used for SED fitting analysis.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for JWST. These observations are associated with Early Release Science program 1345. Part of this work is based on observations taken by the 3D-HST Treasury Program (GO 12177 and 12328) with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. This work is based in part upon High Performance Computing (HPC) resources supported by the University of Arizona Technology and Research Initiative Fund (TRIF), University Information Technology Services (UITS), and Research, Innovation, and Impact (RII) and maintained by the UArizona Research Technologies department.

This research made use of the following software: ASTROPy, a community-developed core PYTHON package for Astronomy (Astropy Collaboration 2013, 2018, 2022; NUMPY (Harris et al. 2020); SCIPY (Virtanen et al. 2020); MATPLOTLIB (Hunter 2007); SOURCE EXTRACTOR (Bertin & Arnouts 1996) via SEP (Barbary 2016); BEAGLE (Chevallard & Charlot 2016); and PROSPECTOR (Johnson et al. 2021).

DATA AVAILABILITY

The JWST/NIRSpec spectra and NIRCam imaging data as well as HST images used in this work are available on the Mikulski Archive for Space Telescopes (https://mast.stsci.edu/). Other data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Astropy Collaboration, 2018, AJ, 156, 123
Bouwens R. J. et al., 2021, AJ, 162, 47
Figure A1. MR ($R \sim 1000$) NIRSpec 1D spectra of $z \gtrsim 7$ galaxies with three or less detected emission lines at rest-frame optical ([O III], Hβ; five objects). The spectra have been shifted to the rest frame. The grey shaded regions show the uncertainty.

Figure A2. 2D NIRSpec spectra (MR, $R \sim 1000$) of Lyα emission lines of $z \gtrsim 7$ galaxies (four objects, see Table 5). Lines are marked by red rectangles, with black-white showing positive-negative features.
Figure A3. 2D NIRSpec spectra (MR, R ∼ 1000) of detected rest-frame UV emission lines of z ≳ 7 galaxies. Lines are marked by red rectangles, with black-white showing positive-negative features. Unresolved C III λ1908 emission lines have been detected in three galaxies: CEERS-1019, CEERS-1149, and CEERS-1027. Unresolved C IV λ1549 and He II λ1640 lines have also been detected in CEERS-1027, and O III λ1661 have been detected in CEERS-1019.

This paper has been typeset from a TeX/LaTeX file prepared by the author.