

EIGER. I. A Large Sample of [O III]-emitting Galaxies at $5.3 < z < 6.9$ and Direct Evidence for Local Reionization by Galaxies

Journal Article

Author(s):

Kashino, Daichi; Lilly, Simon J.; Matthee, Jorryt; Eilers, Anna-Christina; Mackenzie, Ruari; Bordoloi, Rongmon; Simcoe, Robert A.

Publication date:

2023-06-10

Permanent link:

<https://doi.org/10.3929/ethz-b-000618852>

Rights / license:

[Creative Commons Attribution 4.0 International](#)

Originally published in:

The Astrophysical Journal 950(1), <https://doi.org/10.3847/1538-4357/acc588>



EIGER. I. A Large Sample of [O III]-emitting Galaxies at $5.3 < z < 6.9$ and Direct Evidence for Local Reionization by Galaxies

Daichi Kashino^{1,2} , Simon J. Lilly³ , Jorjyt Matthee³ , Anna-Christina Eilers⁴ , Ruari Mackenzie³ ,
Rongmon Bordoloi⁵ , and Robert A. Simcoe⁴

¹Institute for Advanced Research, Nagoya University, Nagoya 464-8601, Japan; kashino.daichi.v7@f.mail.nagoya-u.ac.jp

²Department of Physics, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

³Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, Zürich, 8093, Switzerland

⁴MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

⁵Department of Physics, North Carolina State University, Raleigh, NC 27695, USA

Received 2022 November 15; revised 2023 February 16; accepted 2023 March 12; published 2023 June 12

Abstract

We present a first sample of 117 [O III] $\lambda\lambda 4960, 5008$ -selected star-forming galaxies at $5.33 < z < 6.93$ detected in JWST/NIRCam $3.5 \mu\text{m}$ slitless spectroscopy of a $6' 5 \times 3' 4$ field centered on the hyperluminous quasar SDSS J0100+2802, obtained as part of the Emission-line galaxies and Intergalactic Gas in the Epoch of Reionization (EIGER) survey. Three prominent galaxy overdensities are observed, one of them at the redshift of the quasar. Galaxies are found within 200 pkpc and 105 km s^{-1} of four known metal absorption-line systems. We focus on the role of the galaxies in ionizing the intergalactic medium (IGM) during the later stages of cosmic reionization and construct the mean Ly α and Ly β transmission as a function of distance from the galaxies. At the lowest redshifts in our study, $5.3 < z < 5.7$, the IGM transmission rises monotonically with distance from the galaxies, as seen previously at lower redshifts. In contrast, at $5.7 < z < 6.14$, the transmission of both Ly α and Ly β first increases with distance but then peaks at a distance of 5 cMpc before declining. Finally, in the region $6.15 < z < 6.26$, where the additional ionizing radiation from the quasar dominates, the monotonic increase in transmission with distance is reestablished. This result is interpreted to represent evidence that the transmission of the IGM at $z \sim 5.9$ toward J0100+2802 results from the “local” ionizing radiation of galaxies that dominates over the much-reduced cosmic background.

Unified Astronomy Thesaurus concepts: [Reionization \(1383\)](#); [Intergalactic medium \(813\)](#); [Emission line galaxies \(459\)](#); [Circumgalactic medium \(1879\)](#); [High-redshift galaxies \(734\)](#); [Protoclusters \(1297\)](#); [Galaxy evolution \(594\)](#)

Supporting material: machine-readable table

1. Introduction

Understanding cosmic reionization is a central question in cosmology and galaxy evolution and was a primary scientific motivation for the development of JWST. Observational evidence from the analysis of the Gunn–Peterson troughs in quasar spectra (e.g., Fan et al. 2006; McGreer et al. 2015; Bosman et al. 2018; Eilers et al. 2018; Zhu et al. 2021; Bosman et al. 2022; Zhu et al. 2022), as well as analysis of the Thomson optical depth for the cosmic microwave background radiation (Planck Collaboration et al. 2020), have suggested that the epoch of reionization (EoR) extended down to $z \sim 6$, with a tail extending even later to $z \sim 5.5$ (Greig & Mesinger 2017). This is also supported by the reports of declines in the Ly α equivalent widths of high- z galaxies (e.g., Schenker et al. 2012; Tilvi et al. 2014; Mason et al. 2018) and the Ly α luminosity function at $z \gtrsim 6$ (Konno et al. 2014). In parallel, surveys of galaxies in the EoR, complemented by studies of their analogs at lower redshifts, have suggested that the bulk of the ionizing photon budget that is required to complete reionization could plausibly be coming from young star-forming galaxies with reasonable ionizing photon escape fractions (e.g., Robertson et al. 2015; Izotov et al. 2016, 2018; Nakajima et al. 2020; Cain

et al. 2021; Matthee et al. 2022). However, direct evidence for a connection between galaxies and IGM opacity fluctuations in the same volume has not been presented.

To date, there have been some limited observations of the physical connections between the IGM and galaxies during the EoR on different physical scales. Wide-field searches for Ly α emitters (LAEs) and Lyman-break galaxies (LBGs) along the sight lines toward $z \gtrsim 6$ quasars have suggested an inverse correlation at $z \sim 5.7$ between the effective IGM opacity and the galaxy density when both are averaged over very large scales $\gtrsim 50$ cMpc, suggesting that reionization proceeds more slowly in underdense regions (Becker et al. 2018; Kashino et al. 2020; Christenson et al. 2021; Ishimoto et al. 2022). On smaller scales, Meyer et al. (2020) measured the cross-correlation between the Ly α transmission spectra of high- z quasars and surrounding LAEs and LBGs with spectroscopic redshifts and reported a possible excess transmission ~ 10 – 60 cMpc away from the galaxies, likely reflecting the effects of ionizing radiation from these galaxies (see also Kakiichi et al. 2018). Similarly, excess Ly α transmission was also reported associated with C IV absorbers, implying that it too reflects local ionization by the galaxies hosting the absorbers (Meyer et al. 2019).

Meanwhile, state-of-the-art radiation hydrodynamic simulations, coupled with realistic galaxy formation models, have become more reliable in predicting the connections between galaxies and IGM opacity, including the excess transmission near galaxies (Garaldi et al. 2022; Kannan et al. 2022; Smith et al. 2022),



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

following theoretical investigations using hydrodynamic radiative transport computations (Gnedin 2014; Gnedin et al. 2017; Garaldi et al. 2019; Kulkarni et al. 2019; Keating et al. 2020). However, direct observational constraints of the local ionizing effects of galaxies have remained quite sparse, limiting our understanding of the relative contributions from different galaxy populations (e.g., Naidu et al. 2020) and the impact of individual galaxies on the surrounding IGM.

A different probe of diffuse gas during reionization is offered by the metal absorption systems that are seen in high-resolution spectra of high-redshift quasars. The abundance of metals and their ionization state must reflect a combination of the early chemical enrichment of the circumgalactic medium (CGM) and the local ionizing radiation from the host galaxies of the absorbers (Tumlinson et al. 2017). Particularly, as we approach the EoR, where the Ly α absorption saturates due to higher neutral fractions, metal absorption lines become the only discrete observational probe of the ambient baryonic matter. Observations of these metal absorption lines have suggested that the incidence of low-ionization metal absorption systems, represented by species like C II and O I, is roughly constant or even increases toward the EoR, whereas that of high-ionization species like C IV appears to drop quickly beyond $z \approx 5$ (e.g., Simcoe 2006; Simcoe et al. 2011; D’Odorico et al. 2013; Matejek et al. 2013; Chen et al. 2017; Codoreanu et al. 2018; Becker et al. 2019; Cooper et al. 2019; D’Odorico et al. 2022). This observed difference in behavior is of great interest for understanding changes in the CGM that may be linked to the global reionization process. Currently, however, the observational identification of the galaxies associated with the metal absorption systems at these redshifts is very limited (Díaz et al. 2021; Wu et al. 2021) because of the general difficulty of measuring spectroscopic redshifts of such high- z faint galaxies.

Another aspect of galaxy evolution in the early universe involves the growth of the supermassive black holes (SMBHs) that power the very luminous quasars seen at these redshifts. These earliest quasars provide strong constraints on SMBH formation scenarios, appearing to require super-/hyper-Eddington mass accretion, direct collapse of a high-mass seed, or otherwise high duty cycles (e.g., Mortlock et al. 2011; Bañados et al. 2018; Yang et al. 2020).

Unfortunately, estimates of the masses of SMBHs at these redshifts have relied on rest-frame UV lines like Mg II and C IV, which must be indirectly tied back to the H β line. With JWST, however, it has become possible to use the width of the rest-frame optical H β line that, coupled with the luminosity of the quasar, provides the most robust estimates of the SMBH that can be directly compared to the local calibrations of SMBH mass from reverberation mapping (Bentz et al. 2013). In parallel with the more reliable estimates of SMBH mass, observing the larger-scale environment and the nature of the host galaxies of these earliest quasars is of particular interest in the context of the structure formation in the Λ CDM paradigm and (co)evolution of galaxies and SMBHs.

Motivated by these multifaceted science topics, we initiated the survey project Emission-line Galaxies and Intergalactic Gas in the Epoch of Reionization (EIGER). This is built around a 116 hr JWST program of guaranteed time observation (GTO; PI: Lilly; PID 1243) using deep wide-field slitless spectroscopy (WFSS) and broadband direct imaging observations using the Near InfraRed Camera (NIRCam) on JWST. The key goal of this program is to examine the direct connections between

galaxies and the state of the gas in the circumgalactic and intergalactic environments toward the end of the EoR by carrying out deep and complete surveys for [O III] λ 5008–selected emission-line galaxies in the fields of high-redshift luminous quasars for which we have also obtained extremely deep high-resolution spectra.

This paper is based on the first EIGER quasar field to have been observed (of six planned). We present an overview of the observational design of EIGER and our first sample of [O III] λ 4960, 5008–selected galaxies at $5.3 < z < 6.9$ in this single field. We focus in this paper on early results from the cross-correlation analysis between the galaxy sample and the transmission of the IGM. The detailed characterization of the [O III]-emitting galaxies is presented in Matthee et al. (2023, hereafter Paper II). Other aspects of the data will be exploited in a series of companion papers.

The paper is organized as follows. In Section 2, we provide the survey overview and describe the JWST/NIRCam observations, data reduction, and identification of emission-line galaxies. The ground-based spectroscopy of the target quasar is described in Section 3. Section 4 presents the galaxy sample and a first look at the distribution of galaxies along the line of sight, including the identification of a number of strong overdensities, as well as galaxies associated with discrete metal absorption systems. We then present a more detailed analysis of the transmission of Ly α and Ly β in the IGM as a function of distance around the galaxies in Section 5. Section 6 summarizes our results. Throughout the paper, we adopt a flat Λ CDM cosmology with $\Omega_{\Lambda} = 0.69$, $\Omega_{\text{M}} = 0.31$, and $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Planck Collaboration et al. 2020). All magnitudes are quoted in the AB system.

2. The EIGER Survey

2.1. Survey Design

The JWST GTO program 1243, the EIGER survey, was primarily designed to explore the evolution of the IGM and circumgalactic environments at the tail end of cosmic reionization and thereby try to better understand the reionization process. The 116 hr JWST observing program consists of identical mosaics of simultaneous NIRCam WFSS and imaging observations (Greene et al. 2017) in six widely separated fields, each centered on a bright quasar at $5.98 < z < 7.08$. The program is complemented by 100+ hr of ground-based deep high-resolution spectroscopy of all of these quasars and, depending on the field, additional deep Hubble Space Telescope imaging data, optical wide-field integral-field spectroscopic campaigns with the Very Large Telescope (VLT) MUSE and Keck/KCWI, and Atacama Large Millimeter/submillimeter Array observations.

As a JWST GTO program, the key features of the EIGER program were defined about 5 yr ago. It became clear at that time that WFSS has a number of attractions for a study of this sort. First and foremost, spectra are obtained for all objects in the field of view, completely avoiding the need for the constrained preselection of targets that is inherent in the operation of multislit spectrographs such as NIRSPEC (as well as any practical difficulties associated with slits, including their alignment with targets and possibly poorly controlled slit losses). This 100% coverage of WFSS ensures that a complete census of galaxies in the field (above some straightforward flux criteria) is obtained. In addition, although deep WFSS

inevitably suffers from problems of contamination and overlapping spectra, because of the high surface number density of faint sources in the sky, the relatively high spectral resolution of the NIRCcam grisms suggested that it would be relatively straightforward to subsequently process the dispersed WFSS images so as to cleanly remove the continuum emission of all sources, leaving only the signals from emission lines.

The main difficulty of WFSS for a study of this kind concerns the identification of the source responsible for any given emission line, since there is an intrinsic degeneracy between spatial position and wavelength for any emission line detected in the dispersed image, plus the usual difficulties of line identification for single emission lines. This latter issue motivated our primary focus on the [O III] $\lambda\lambda 4960, 5008$ doublet while allowing for the possibility of ancillary science using other lines at lower redshift. As rest-frame visible emission, it is relatively unaffected by dust reddening. At the redshifts of interest, $5 < z < 7$, this emission is redshifted to $3.5 \mu\text{m}$, where the JWST background is lowest. Furthermore, previous photometric studies had suggested that these lines would be very strong, with high equivalent widths (e.g., Stark et al. 2013; Smit et al. 2014).

Most crucially, the fact that this is a well-defined doublet with a fixed line ratio and often appears as a triplet with the nearby $H\beta$ line suggested that [O III] $\lambda\lambda 5008, 4960$ emission could be easily isolated in the high-resolution ($R \sim 1500$) NIRCcam WFSS dispersed images. Furthermore, the observed $\Delta\lambda$ spacing between these two lines on the detector would already give an indication of the redshift and thus wavelength, significantly constraining the spatial positions of the sources responsible. Of course, for a given line flux, WFSS is always less sensitive than slit spectroscopy because of the elevated background in WFSS images, but the extraordinary sensitivity of JWST relative to any previous facility made this an acceptable trade, especially since, except at low spectral resolution, dispersed slit spectroscopy with NIRSpec is detector noise-limited.

NIRCcam WFSS observations with the F356W ($\approx 3.1\text{--}4.0 \mu\text{m}$) filter in the long-wavelength (LW) channel therefore allow isolation of [O III] $\lambda\lambda 5008, 4960$ at $5.3 < z < 7.0$, the target redshift for the EIGER survey of galaxies at the end of the EoR. A large number of other emission lines at generally lower redshift are also detected in the WFSS emission-line image. Scientific utilization of these other emission lines is, of course, possible but requires more careful treatment of the identification problem and will be addressed in later papers.

Relatively short direct images are also mandatory in the same F356W filter to allow precise positional identification of sources. The two channels of NIRCcam allow deep direct images to be obtained in the F115W and F200W filters in the short-wavelength (SW) channel during the WFSS and direct imaging in F356W in the LW channel. The two SW filters were chosen to allow the characterization of the spectral energy distribution (SED) of galaxies at $z \sim 6$ by sampling the rest-frame UV spectrum in the two SW bands and, together with the F356W photometry, across the Balmer/4000 Å break band (although the F356W fluxes, by construction, contain multiple strong emission lines).

As described in detail below, a mosaic pattern of four offset field centers (each representing a separate JWST “visit”) was

Table 1
Target Quasars

Identifier	z_{QSO}	References ^a
SDSS J010013.02+280225.8	6.327	1, 2
ULAS J014837.63+060020.0	5.98	3
SDSS J103027.09+052455.0	6.280	4
PSO J159.2257–02.5438	6.381	5
ULAS J112001.48+064124.3	7.085	6, 7
SDSS J114816.64+525150.3	6.475	8

Note.

^a References: 1–Wu et al. (2015); 2–Venemans et al. (2020); 3–Jiang et al. (2015); 4–Fan et al. (2001); 5–Bañados et al. (2016); 6–Mortlock et al. (2011); 7–Venemans et al. (2020); 8–Alam et al. (2015).

chosen so as to cover a contiguous, reasonably wide survey area at more or less uniform depth, together with a central deep overlap region that was centered on each of the quasars. The full survey area of $6.5 \times 3.4 \text{ arcmin}^2$ around each quasar enables the characterization of the large-scale galaxy distribution on scales of up to $\sim 10 \text{ cMpc}$ around the sight line, enabling the cross-correlation of galaxies and the $\text{Ly}\alpha$ and $\text{Ly}\beta$ transmission of the IGM to be robustly examined on these scales with a complete galaxy sample (avoiding the danger of missing galaxies). The central deeper overlap area of $\approx 40 \times 40 \text{ arcsec}^2$ allows the full depth to be achieved close to the sight line, enabling the nature of the quasars, their host galaxies, and their surrounding environments to be examined and to carry out high-completeness searches for galaxies associated with metal absorption systems in the quasar spectra. The overall exposure time was set by the maximum depth that we required close to the quasar. The size of the positional offsets for the four centers in the mosaic pattern therefore at some level represents a trade-off between reducing the area of this central, maximally exposed region and increasing the total area observed at a 25% exposure level. In practice, the offset, at least in the x -direction, is largely determined by the gap between the two modules of the NIRCcam instrument.

2.2. Target Quasar Selection

The six target quasars (Table 1) were selected to address a range of complementary science goals. All six had very deep high-resolution spectra available from the optical to near-IR from ground-based observations with VLT/X-Shooter, the Magellan Folded-port InfraRed Echellette (FIRE), and the Keck High Resolution Echelle Spectrometer (HIRES). The quasars were chosen on the basis of what was known about the properties of the $\text{Ly}\alpha$ forest and/or Gunn–Peterson trough, the existence and ionization state of the known metal absorption systems along the line of sight, and the characteristics of the quasars themselves.

In the following, we describe the first EIGER observations that have been successfully completed in the field of SDSS J010013.02+280225.8 (hereafter J0100+2802) at $z = 6.327$ (Wu et al. 2015; Venemans et al. 2020) during 2022 August. It is anticipated that an almost identical set of data will be obtained on the other five quasars during the first year of JWST science observations.

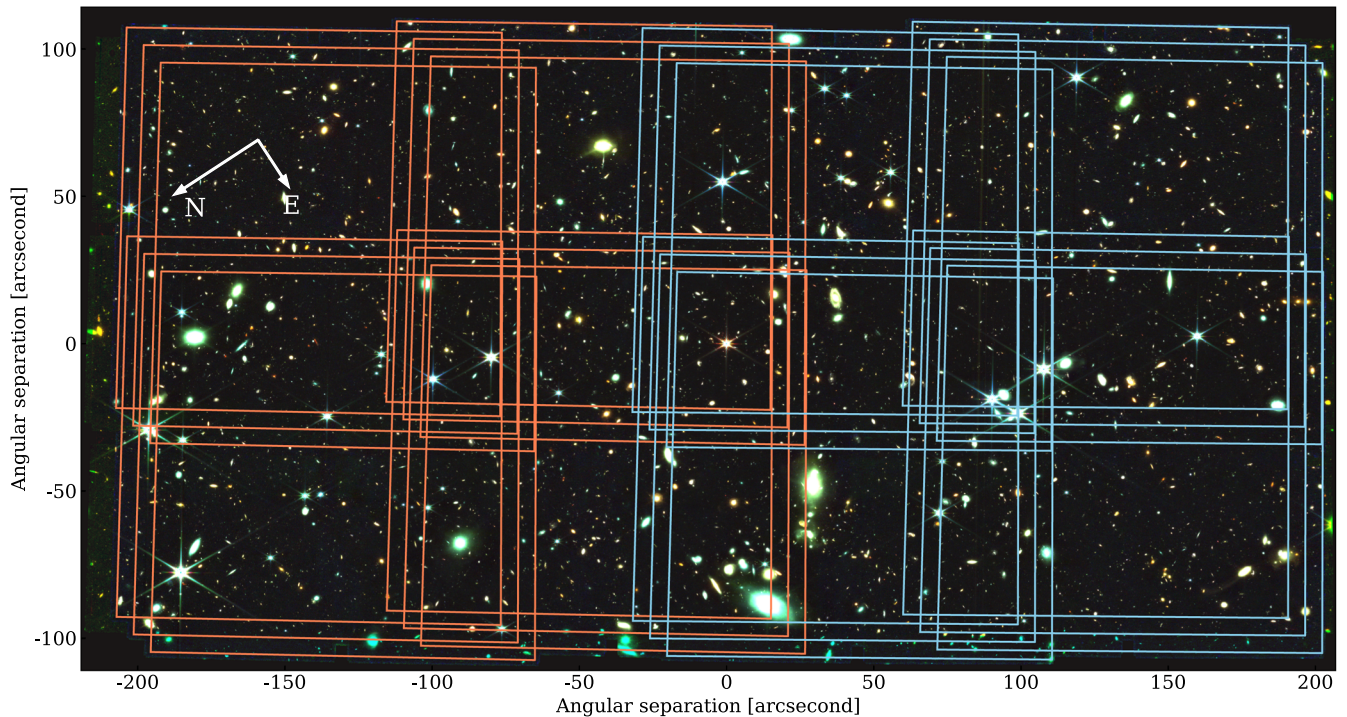


Figure 1. JWST NIRCcam observations of the field of the quasar J0100+2802. The background is the coadded pseudocolor image in the F115W (blue), F200W (green), and F356W (red) bands. The axis values are the angular separations from the position of the quasar. The detector fields of view of the individual WFSS exposures in our mosaic observations are shown by the large red and blue squares, respectively, for the NIRCcam module A and B detectors.

2.3. The NIRCcam Observations

The NIRCcam imaging and WFSS observations of the field of QSO J0100+2802 were carried out during 2022 August 22–24.⁶

NIRCcam has two modules (A and B), each with a $2'.2 \times 2'.2$ field of view separated by $44''$. The EIGER observations consist of a mosaic of four otherwise identical “visits” that form a 2×2 overlapping mosaic pattern. This gives a contiguous survey field of $\approx 6'.5 \times 3'.4$ centered on the target quasar and a central $\approx 40'' \times 40''$ region with a fourfold longer exposure time. The position angle of the mosaic (defined by the telescope V3 axis) for the observations of QSO J0100+2802 was 236° . This is usually arbitrary but was constrained for some of the quasar fields to ensure that the central region around the quasar in each grism image was not contaminated by spectra from other known brighter sources in the field. The mosaic pattern in the QSO J0100+2802 field is shown in Figure 1.

The WFSS observations use the LW channel and a combination of the grism *R* and F356W filters. This combination yields $3.1\text{--}4.0\ \mu\text{m}$ spectra dispersed along the detector rows. The spectral resolution for a point source is $R \sim 1500$ at around $3.5\ \mu\text{m}$, and the dispersion is $\approx 1\ \text{nm pixel}^{-1}$. The two modules A and B, with the grism *R*, disperse the spectra in opposite directions; i.e., the wavelength increases toward the $+x$ and $-x$ directions in the detector coordinates, respectively. In both modules, the zero deviation wavelength is $3.95\ \mu\text{m}$, close to the long wavelength end of the F356W filter response curve.

The adopted mosaic configuration thus yields two reversed dispersion spectra for all sources within a central vertical strip

of width $\approx 60''$ that is covered by both modules. Furthermore, two independent but not reversed spectra are obtained for all sources within two off-center vertical strips of width $\approx 60''$ and a central horizontal strip of width $\approx 70''$. The horizontal and center vertical strips intersect in the central square region centered on the quasar.

Together with the longer exposure time, the availability of the two reversed grism spectra helps to disambiguate the identification of source objects in the direct images, as well as the wavelengths of any emission lines, and to remove contaminating lines from other sources. When only one dispersion direction is available (which is the case over most of the extended survey field), the identification of the correct source in the direct image (to produce the correct wavelength solution) requires some care. However, extensive simulations before JWST launch indicated that this could be achieved with high confidence, and this has proved to be the case in practice.

Likewise, since the primary science goal was the detection of a very recognizable spectral feature, the [O III] $\lambda\lambda 4960, 5008$ doublet (or the triplet including $H\beta$), the difficulties produced by overlapping spectra (from interloping continuum source or line emission) were minimized. As will be described below, the relatively high spectral resolution of the NIRCcam WFSS means that contaminating continuum sources with slowly varying SEDs can be removed effectively by suitable filtering of the data.

The effective field of view for the dispersed images depends upon the observed wavelength and therefore the redshift of each source. The dispersed spectra with grism *R* through the F356W filter are ~ 940 pixels long, about half the width of the detector (2048 pixels), and extend across from the undeviated wavelength of $3.95\ \mu\text{m}$. Therefore, in module A, the dispersed spectrum extends across -850 to $+90$ pixels relative to the

⁶ All of the JWST data used in this paper can be found in MAST: [10.17909/2w6t-bn31](https://mast.stsci.edu/#/search/10.17909/2w6t-bn31).

undeviated position of the source, or +850 to −90 pixels in module B.

Given the LW detector size of 2048^2 pixels, the spectra of sources within ~ 850 pixels from one of the vertical edges (at $x = 1$ for module A and $x = 2048$ for module B) extend beyond the detector coverage, losing (in both modules) some length of the spectrum at the short wavelength end. When the redshift is based on a single spectral feature, this effectively reduces the survey area at the low-redshift end of the redshift range. The 2×2 mosaic pattern used in the EIGER survey means that the full spectra are in fact obtained for half the affected sources in another pointing. In the combined data set, full spectral coverage is obtained for a region of width $\approx 4'$. Provided the emission line is bright enough to be detected, the effective survey area therefore increases from ≈ 17 arcmin² at $\lambda = 3.1$ μm up to the maximum ≈ 25 arcmin² at $\lambda = 3.95$ μm , corresponding to redshifts of [O III] $\lambda 5008$ from $z = 5.3$ to 6.9.

During the WFSS observations, simultaneous direct (non-dispersed) imaging was obtained using the SW channel, first with the F115W filter and then with F200W. After the WFSS observations were completed, additional direct imaging was then obtained in the F356W filter in the LW channel, with further exposure time obtained in F200W in the SW channel. This last set of direct images includes two displaced field centers to ensure that direct images are obtained for all sources whose light can be dispersed by the grism onto the detector.

In summary, each visit has an identical sequence that consists of (1) F356W WFSS + F115W imaging, (2) (identical) F356W WFSS + F200W imaging, and (3) direct imaging in F356W + F200W. For both the WFSS and SW direct imaging, we employed a combination of the three-point INTRAMODULEX primary and four-point subpixel dithering, yielding a total of 12 exposures per visit and module in each of F115W and F200W and 24 WFSS images. Coadding these multiple exposures, the total nominal exposure time in a single visit is 4380 s filter^{−1} for the SW imaging and 8760 s for the WFSS.

Given the small-scale dithering and larger-scale mosaic pattern, the total exposure time varies across the field, ranging from 1.5 ks (for the extreme edge regions sampled by only a single primary dither) up to 13 ks for SW imaging and from 2.9 up to 35 ks for WFSS. The F356W direct and two “out-of-field” images have an exposure time of 526 s each, yielding a total exposure time of 1578 s over most of the field. Where the visits overlap, the exposure time increases to a maximum of 6.3 ks in the central region around the quasar.

2.4. Imaging Data Reduction and Photometry

The JWST/NIRCam imaging data were reduced using the `jwst` standard calibration pipeline (version 1.8.2⁷), adopting the Calibration Reference Data System context `jwst_0988.pmap` (PUB) including updated zero-points based on in-flight data (Boyer et al. 2022). All of the individual broadband raw images were first processed for detector-level corrections using the `Detector1` module of the pipeline. Then, the individual products (count-rate images, i.e., `_rate.fits` files) were calibrated through `Image2`, where flat-fielding and the flux calibrations were applied to convert the data from units of count rate to surface brightness (or flux density). Circular defects, known as “snowballs,” that arise from strong cosmic

ray hits were masked following Merlin et al. (2022). The individual calibrated images (`_cal.fits` files) were then resampled and coadded onto the common pixel grid with a pixel resolution of $0''.03$ pixel^{−1} through the `Image3` processing. The astrometry was calibrated to align stars from the Gaia Data Release 2 catalog (Gaia Collaboration et al. 2018).

The individual frames were further processed to subtract artificial scattered light, known as “wisps”; $1/f$ noise (correlated horizontal and vertical noise; e.g., Schlawin et al. 2020); and residual sky background using a source mask based on a preliminary coadded image before obtaining the final coadded images. The wisps were removed using the templates that were constructed by median-combining all exposures in the detector coordinates in a given filter and detector element. These templates were only applied to the SW filters, as the LW channel appears unaffected. This worked well, since our program conducted mosaic observation with large offsets in the sky. The $1/f$ noise is excluded by a sigma-clipped median subtraction in each 1/4 detector patch of a row that corresponds to one amplifier, then likewise a subtraction per column and finally per amplifier. The pseudocolor image constructed from the final coadded images in the F115W (blue), F200W (green), and F356W (red) bands is shown in Figure 1.

Aperture-matched photometry was carried out using `SEXTRACTOR-2.25.0`⁸ with the F356W imaging data as a detection image. First, we convolved the higher-resolution F115W and F200W images to match the point-spread function (PSF) of the F356W image. We derived the modeled PSFs for each of the filters using `webbpsf`⁹ (Perrin et al. 2014) and then created a matching kernel using a `SplitCosineBell` window (in `photutils`). This methodology was tested extensively using `MIRAGE`¹⁰ simulations. Total magnitudes were measured with Kron apertures following Finkelstein et al. (2022). The Kron scaling factor was optimized to maximize the signal-to-noise ratio (S/N), and we applied aperture corrections. The uncertainty on the photometry was estimated using a model that fits the noise in random apertures in blank sky regions of the images as a function of aperture size and scaling for variations in the sensitivity. The average 5σ sensitivities are 27.8, 28.3, and 28.1 in the F115W, F200W, and F356W imaging data, respectively, reaching a maximum sensitivity of 28.7, 29.2, and 29.0 in the 10% deepest regions.

2.5. WFSS Data Reduction

The WFSS data were reduced with a combination of the `jwst` pipeline (version 1.7.0) and custom python-based processing steps. Figure 2 summarizes the procedure of the WFSS data processing. The development of this method was motivated to allow for the robust detection of emission lines in the dispersed grism images.

We first processed individual grism images using `Detector1` for detector-level corrections and used the `assign_wcs` step from `Spec2` to obtain the WCS solution for each grism image. We then processed the data using `Image2` for flat-fielding and subsequently removed the large-scale sky background variations and $1/f$ noise by straight

⁷ <https://github.com/spacetelescope/jwst>

⁸ <https://sextractor.readthedocs.io/en/latest/index.html>

⁹ <https://webbpsf.readthedocs.io>

¹⁰ <https://mirage-data-simulator.readthedocs.io/en/latest/>

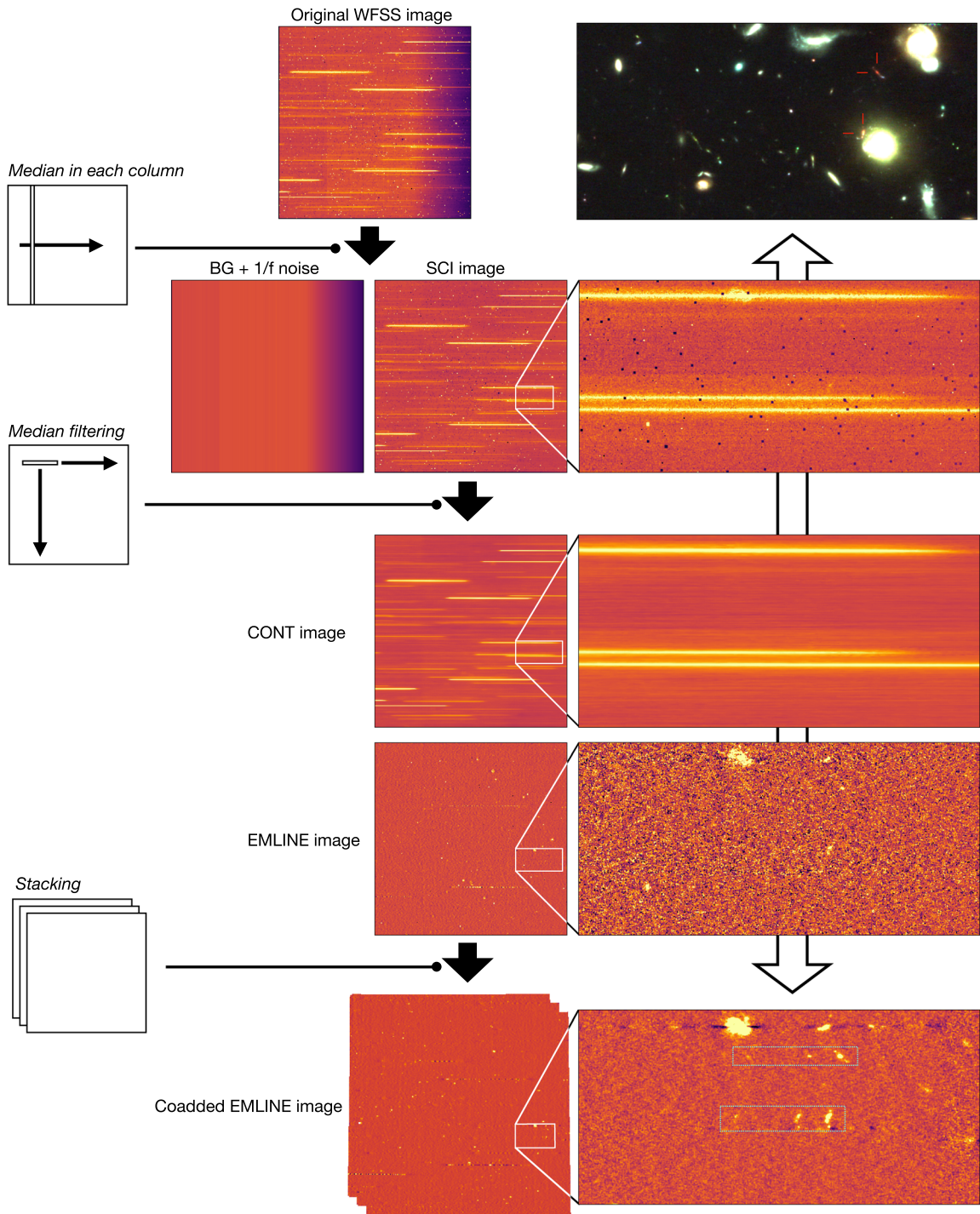


Figure 2. Processing of the WFSS data after the `Detector1`, `assign_wcs`, and flat-fielding processes. The WFSS images of individual exposures are first processed to subtract the background and $1/f$ noise. The resulting SCI images are then median-filtered to separate them into continuum (CONT) and emission-line (EMLINE) images; see text for details. These images are finally stacked using `Image3` to produce a single image for each visit and module. The rightmost column shows a zoom-in of a small region of the WFSS images, with a direct image of the corresponding region in the top panel. The green dotted rectangles mark the [O III] doublet and $H\beta$ lines that originate from the galaxies marked by the red reticles in the direct sky image.

median subtraction in each detector column (orthogonal to the dispersion direction). We refer to the output after this basic processing step as the “SCI” image.

The next step is to isolate emission lines in the SCI images.¹¹ We applied empirical median filtering using a boxcar-like

kernel along each detector row (i.e., parallel to the dispersion direction) to obtain filtered images that represent all of the different continuum components that appear in the grism data. This filtered continuum image may then be subtracted from the SCI image to yield, in principle, a pure emission-line image. The primary kernel for this median filtering had a length of 51 pixels with a central “hole” of 9 pixels ($\approx 770 \text{ km s}^{-1}$) so as to avoid self-subtraction around emission lines. Figure 3 presents

¹¹ The python programs for these steps will be made public upon publication.

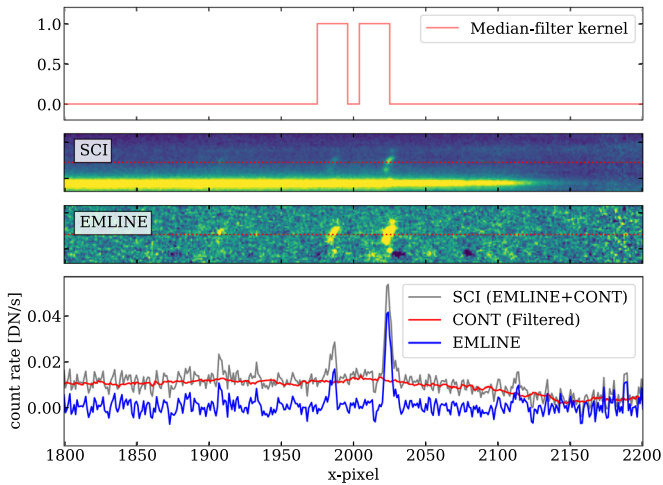


Figure 3. Demonstration of the median-filtering step to isolate emission lines and continuum. The top panel shows the kernel for median filtering that has a full width of 51 pixels and a central “hole” of 9 pixels. The second and third panels show the zoom-in of a small region of the SCI and EMLINE grism images around the [O III] doublet + H β lines. The bottom panel shows 1D spectra from the SCI (gray), CONT (red), and EMLINE (blue) frame along the single “row” marked by the horizontal red dotted lines. All of these panels have the same x -axis range across 400 pixels.

this kernel shape and demonstrates that this median filtering generally worked very well in removing the continua. This relatively long kernel, however, noticeably oversmooths the continua at each end of the spectrum, where the filter response curve produces large short-scale variations in the continuum level. This causes a spurious positive–negative pattern in the continuum-subtracted images at each end of the spectrum. To mitigate this end effect, we first identified the ends of continuum spectra by measuring the gradient of the brightness along each detector row and adopted a shorter kernel of 21 pixels in these regions. A composite continuum image was then produced from these two different-sized kernels with gradually changing fractions at the boundary to avoid discontinuity. This step was actually repeated with an emission-line mask obtained from the coadded grism image as described below. We note that the use of only the shorter kernel does not yield better results, particularly around the [O III] + H β lines, as it is much more affected by emission lines with close separations than the longer kernel. As shown in Paper II, this custom filtering could result in a $<5\%$ level oversubtraction of the continuum around emission lines.

The emission-line (hereafter “EMLINE”) images were then generated by subtracting these median-filtered continuum images (“CONT”) from the SCI data. These procedures were developed and tested extensively before launch using MIRaGe simulations based on the JAGUAR catalog (Williams et al. 2018) and worked well on the real data. The most noticeable residuals from the continuum-subtraction procedure involved bright galaxies around $z \sim 0.5$ where the strong CO bandheads are not completely removed. These residual features are easily recognized.

With the EMLINE images of individual exposures in hand, we then adopted two different approaches for image stacking and subsequent line identification. In the first method, we used the Image3 module of the `jwst` pipeline to coadd the set of 24 individual full detector-sized grism images (SCI, EMLINE, and CONT) per visit and module. Stacking of the full-sized images works only for images with small spatial offsets, up to

the primary dither size, and the same dispersion direction. The geometric relation between the position of a source and the location of the spectral trace on the detector has both spatial- and wavelength-dependent distortions. But the differential distortion over the spatial scale of the dither pattern is small enough to be ignored, allowing the coadding of these images.

The coadded EMLINE images produced in this way enable us to directly detect all emission lines in the coadded grism data using SExtractor without any knowledge of the sky locations of the source galaxies (see the bottom panels of Figure 2).

The second method extracts 2D spectra from the individual grism images for all sources detected in direct imaging (see Section 2.4). Despite the somewhat shorter exposure time of the F356W direct images, even pure emission-line sources from the WFSS frame are essentially always detected in the direct image, since the backgrounds are essentially the same in the two. In spectral extraction, we used `grismconf`¹² to calculate a position (x, y) in the EMLINE image for a given sky position and wavelength. We adopted the NIRCcam grism model (V4)¹³ based on commissioning observations. The spectral trace is not perfectly parallel to the detector row direction. Therefore, the extracted 2D spectra were first rebinned column by column to produce flat spectra and then resampled into a common linearly spaced wavelength grid in steps of 9.75 Å. These procedures were executed so that the fluxes are conserved. The data in units of count rate were then converted into flux density units ($\text{erg s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$) by using the total sensitivity curve as a function of wavelength. The sensitivity curve is provided for each of modules A and B in the grism model, and we slightly modified them by recalibrating with commissioning data. The average throughput is $\approx 28\%$ lower in module B than A. Stacking was then performed for each source selected in the direct image using all the available WFSS exposures separately for each module; recall that the EMLINE images tracing the source shape in the sky are flipped between the modules. Emission lines could then be identified in these stacked 2D spectra (although the same emission line may appear in the stacks of multiple objects).

These two methods of grism image stacking are linked to the two approaches for identification of emission lines and source galaxies, as described in the following subsection.

2.6. Identification of [O III]-emitting Sources

We developed two complementary methods of identifying [O III]-emitting galaxies from the EMLINE images, referred to hereafter as the “backward” and “forward” approaches. These identifications were independently carried out by two members of our team (D.K. and J.M., respectively), and then the candidate samples were reconciled to produce a final sample.

The backward approach starts with the coadded EMLINE images for each visit and module. We first detected all emission-line signals using SExtractor and searched for pairs or triplets of detections that could conceivably be a combination of the [O III] $\lambda\lambda 4960, 5008$ doublet and/or H β . In practice, we first selected all emission lines detected at 5σ or more as a primary detection, where the S/N is determined using the flux and associated errors in a fixed 5 pixel ($0''.315$) aperture. We then provisionally assigned this primary line to be

¹² <https://github.com/npirzkal/GRISMCONF>

¹³ https://github.com/npirzkal/GRISM_NIRCAM

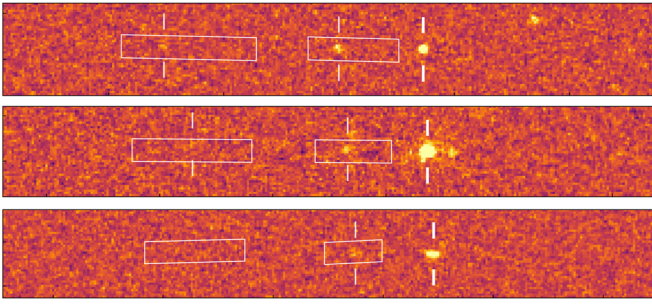


Figure 4. Three representative examples of the search for pairs or triplets of emission lines in the EMLINE images in the so-called “backward” approach. These are candidates for the [O III] doublet and $H\beta$. The primary line detections, putatively [O III] $\lambda 5008$, are marked by thick reticles. The white boxes then indicate the search boxes for the secondary lines, i.e., [O III] $\lambda 4960$ and $H\beta$, which include the vertical uncertainty in position and the horizontal uncertainty in the redshift and resulting $\Delta\lambda$. The left and right boundaries of these boxes therefore correspond, respectively, to the extreme cases where the primary [O III] $\lambda 5008$ is actually at 3.1 and 4.0 μm . The secondary line detections found within these search boxes are marked by thin reticles. All three sources were indeed securely identified as [O III] emitters.

[O III] $\lambda 5008$ and searched for the required secondary lines, i.e., [O III] $\lambda 4960$ and/or $H\beta$. Because at this point, the true source position is unknown, the putative [O III] $\lambda 5008$ line could lie anywhere in the possible 3.1–4.0 μm window, i.e., anywhere in the $5.3 < z < 7.0$ redshift range, meaning that the other lines of the doublet/triplet could be displaced by a range of pixels. The search boxes for the other putative components were therefore defined accordingly using the grism model. Figure 4 shows examples of the detection of the primary lines and the search for the secondary lines. Note that these procedures are based purely upon the emission lines and do not rely on any knowledge from direct imaging. All pairs/triplets of lines that could conceivably be the desired [O III] $\lambda 4960$, 5008 combination, with or without $H\beta$, were thereby recorded. For each combination, an approximate redshift is derivable from the line separation in the EMLINE image.

The next step is to examine the robustness of these line pairs/triplets and identify the source galaxies of these emission lines. The grism model enables us to inversely calculate the sky location (R.A., decl.) for a given position in the grism image (x , y) for any and all wavelengths. We therefore first derived a “trace” in the sky (i.e., a locus in the F356W direct image) as a function of assumed wavelength between 3.1 and 4.0 μm for each of the primary lines that had at least one possible secondary line detection of [O III] $\lambda 4960$ and/or $H\beta$, and then we identified all sources that are located within $\approx 0''.15$ of this trace as candidates to be the source galaxy responsible for the emission lines in question. Finally, we visually inspected the possible combinations and determined which galaxy is the source of the emission lines, otherwise the emission line is discarded as a non-[O III] source. A key requirement for all sources is that the implied redshift (fixed once the source position is determined) must be consistent with the approximate redshift determined from the separation of the lines in the grism image. Further judgments were made on the basis of the available information, including the spatial offset from the expected position, the flux ratio of the doublet lines, the morphological properties of the source in the direct and dispersed images, and the observed broadband color of the candidate sources.

These procedures for line detection and line-source identification were performed separately for each visit and module. The results of these visual inspections were then merged into a single combined catalog of [O III]-emitter candidates by excluding duplicates and resolving conflicts in the overlap regions.

This backward approach provides a sensible lower limit to the number of [O III] emitters and has the potential advantage that it produces a purely line flux–limited set of emission-line sources, including sources that may be ambiguous in the direct images, due to their intrinsic faintness, blending with other bright sources, or any possible contamination. On the other hand, the quality of the spectral images cannot be optimal because of the stacking that can only be done per visit/module and the small but nonzero smearing of line signals due to the differential distortions, even between primary dithers ($\approx 6''$ in both the x - and y -directions) within the same visit.

We therefore also attempted the inverse approach using the best-quality coadded images of the spectra as follows. In this so-called “forward” approach, we start with a catalog of broadband sources detected in the F356W image (see Section 2.4). We first picked up all $\sim 20,000$ sources brighter than $m_{\text{F356W}} = 29$ mag. For each of these, we extracted 2D spectra from the grism (SCI, EMLINE, and CONT) images of the individual exposures and then coadded them as described in Section 2.5. We then ran `SExtractor` on each of these coadded EMLINE 2D spectral images to detect emission-line signals and selected those in which candidate line pairs/triplets were detected at $>3\sigma$, close to the expected locus of the spectral trace of the object and with a sensible separation in wavelength.

We conservatively selected all candidate pairs for the [O III] doublet with a relative flux between a factor of 1.5 and 6 (i.e., a factor of 2 around the expected 2.98 ratio), a maximum 3 pixel ($0''.19$) difference in the spatial centroid (i.e., in the column (y) direction) of the lines in the 2D spectrum, and a doublet separation within 15 \AA of the expected separation for the possible redshift solution. Through visual inspection, we removed candidates where one of the lines was likely to be an artifact, clearly came from a different galaxy, or was evidently a residual from the continuum filtering method, using the reversed dispersion images from both modules if available. Then we verified the association between the lines and the broadband source primarily based on the spatial position and doublet separation and, in some cases, based on color and/or morphology. In doing so, we carefully resolved the cases where a single emission-line pair is associated with multiple galaxies. The main benefit of the forward approach is that it can optimally combine data over multiple visits and modules and therefore yield fainter sensitivity than the backward approach described above.

The two samples of [O III]-emitter candidates obtained from these two complementary approaches were then carefully reconciled with an iterative fine-tuning of the search parameters so as to obtain the final catalog of [O III] emitters. A first comparison showed that the two samples agreed on $\sim 70\%$ of the final sample, with $\sim 20\%$ of the fainter sources not initially identified in the backward approach and $\sim 10\%$ missed by the forward approach. The sources initially missed in the forward approach included one severely blended under a bright nearby object, and some that were missed due to the inaccuracy of the grism model, or for which the forward approach missed [O III]

$\lambda 4960$ near the threshold S/N, or for which the line-source association was revised based on the backward sample during the reconciliation procedure. A handful of sources were discarded as unreliable, and other individual discrepancies were resolved by inspection by D.K. and J.M. The use of two independent and in a sense opposite approaches gives us more opportunities to double-check the associations between lines and sources, significantly reducing the risk of misidentification.

The main challenge in both approaches was to associate a unique set of emission lines with a unique galaxy, especially in crowded regions, where a single set of [O III] (+H β) lines may plausibly be associated with multiple galaxies. It should be noted that limiting ourselves to objects with at least two significant line detections significantly facilitated this procedure because the line separation limits the possible range of redshift and thus of position in the sky. This is a significant benefit of designing the survey to search for the [O III] lines instead of, for example, H α with a redder filter.

2.7. The Sample of [O III] Emitters in the J0100+2802 Field

After visual inspection and reconciliation, we identified 133 [O III]-emitting components over $z = 5.33\text{--}6.93$ with at least two detected emission lines above 3σ . As detailed in Paper II, we decided to classify clumps located within $2''$ as a single “system,” yielding a final sample of 117 unique [O III]-emitting systems, or galaxies. Measurements of the emission-line fluxes were carried out using the best-quality 1D spectra that were extracted from the coadded 2D spectral images (see Paper II). We note that these are subject to the $\sim 5\%$ -level systematic uncertainties in the absolute flux calibrations. We also estimated the systematic redshift uncertainties to be about $\sim 100 \text{ km s}^{-1}$ from inaccuracy of the grism model. This corresponds to ~ 1 pixel uncertainty in the wavelength calibration. For most cases, the systematic redshift uncertainty dominates the statistical errors. Figure 5 presents a representative set of stacked 2D spectra of [O III] emitters in order of redshift.

The full catalog of the [O III]-emitter sample is provided in Table 2. We classified 94 sources as high-confidence sources (CONFID = 2) and the other 23 as lower confidence (CONFID = 1) based on the measured S/N of the lines and any remaining ambiguity in the line–galaxy identification. This classification of confidence was independently made by D.K. and J.M., and any discrepancies were carefully reconciled. We estimate that these two classes are $>99\%$ and $>90\%$ reliable, respectively, for CONFID = 2 and 1 by repeating equivalent visual inspections for artificial signals based on the JAGUAR catalog (Williams et al. 2018) embedded in the real grism data. Further objective assessment of robustness, completeness, and sample contamination will be conducted using a larger sample including future observations of the remaining EIGER fields. Our present scope only concerns the early development of methods to process NIRCcam data, initial data quality, and findings from the J0100+2802 sight line. Throughout this paper, we use all of the [O III] galaxies for analysis. Exclusion of low-confidence galaxies does not change any conclusions, except for one galaxy that is identified as associated with a metal absorption system (see Section 4; this association makes it unlikely to be a chance spurious association with a wrong redshift).

3. Ground-based Spectroscopy of QSO J0100+2802

We used ground-based spectroscopic observations of QSO J0100+2802 to measure the transmission as a function of redshift in the Ly α and Ly β forest region and identify intervening absorption systems. The spectroscopic data used in this paper were obtained with the X-Shooter spectrograph (Vernet et al. 2011) on the VLT, FIRE (Simcoe et al. 2013) on the Magellan Telescope, and HIRES (Vogt et al. 1994) on the Keck Telescope.

The archival X-Shooter data (ESO program 096.A-0095) and the FIRE data (PI: Simcoe) were reduced, carefully corrected for telluric absorption using numerical atmospheric models, and optimally combined to maximize the S/Ns using the PyPeIt¹⁴ spectroscopic reduction pipeline. A detailed description of the data reduction will be provided in a companion paper (Eilers et al.). The HIRES data were reduced using the makee pipeline, and the coadded 1D spectrum was then corrected for telluric absorption using observations of a hot white dwarf spectrophotometric standard star (Cooper et al. 2019).

We used the X-Shooter spectrum obtained in the visible spectrograph to measure the Ly α and Ly β transmission fluxes across the redshift range of interest. The intrinsic continuum spectrum blueward of rest-frame 1280 \AA was estimated from the shape of its redward spectrum using a deep neural network trained with a sample of low-redshift ($0.1 \leq z \leq 3$) quasar spectra (Liu & Bordoloi 2021). The intrinsic spectrum that was estimated by this method extends down to the rest-frame wavelength of 1080 \AA and was complemented to shorter wavelengths by a power-law fit to the continuum redward of Ly α . In doing so, the power-law fit was forced to continuously connect the one predicted from the deep-learning network at rest-frame 1080 \AA . Figure 6 shows the observed X-Shooter spectrum and estimated intrinsic spectrum. We estimated the 1σ uncertainties of the predicted continuum spectrum to be $\approx 5\%$ for the Ly α and Ly β forest regions that we will use for analysis.

In the redshift window covered by [O III] ($z = 5.3\text{--}6.327$), we identified eight discrete intervening absorption systems using all of the available spectroscopic data. The high-resolution spectra suggest that these systems often contain multiple blended absorption components. We used absorption-line fitting routines that generate a single hierarchical model of absorbing components and ion transitions and project the same model into the respective spectral data space of each instrument, convolving to the appropriate spectral resolution. The model was optimized to fit the data from all instruments simultaneously using a Markov Chain Monte Carlo technique implemented in the emcee package (Foreman-Mackey et al. 2013). Full details of the identification and measurements of absorption lines will be provided in dedicated companion papers (R. Bordoloi et al. 2023, in preparation; R. Simcoe et al. 2023, in preparation).

In this paper, the absorption systems are simply categorized into low- and high-ionization systems according to the transitions and their strengths. Three absorption systems at absorption redshifts $z_{\text{abs}} \approx 5.34, 6.01, \text{ and } 6.19$ exhibit significant absorption by the CIV 1548, 1550 \AA transitions, and these are classified as high-ionization systems. The other five systems at $z_{\text{abs}} \approx 5.35, 5.80, 5.95, 6.11, \text{ and } 6.14$ are

¹⁴ <https://github.com/pypeit/PyPeIt>

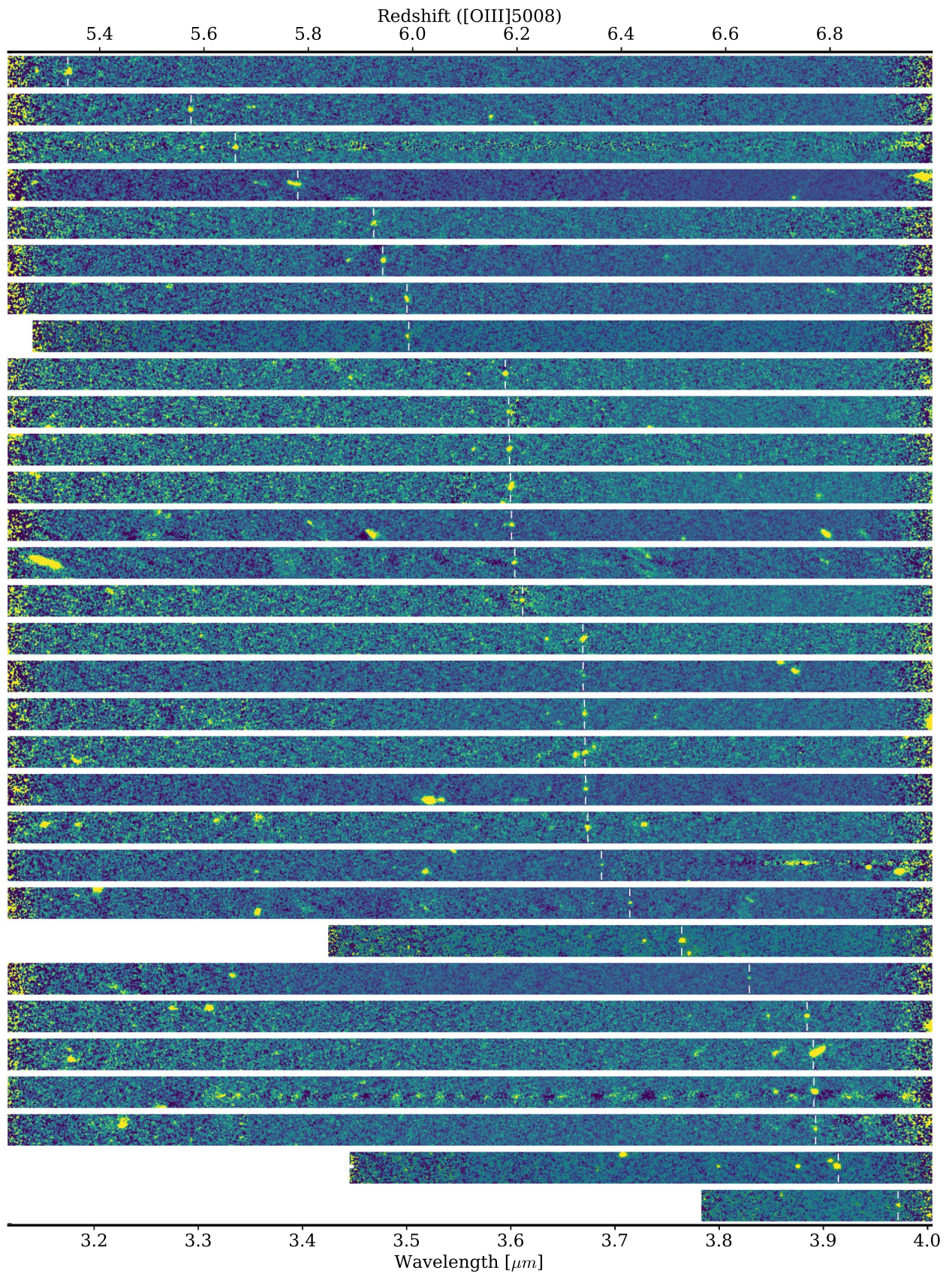


Figure 5. Examples of stacked 2D spectra of [O III] emitters, shown in order of redshift from top to bottom. The spectra are rebinned to be flat and resampled into a common wavelength grid in steps of $9.75 \text{ \AA pixel}^{-1}$. The [O III] $\lambda 5008$ line is marked by ticks. Some spectra were truncated because the bluer part of the spectrum was dispersed beyond the detector coverage (see Section 2.3).

Table 2
[O III]-emitter Sample^a

ID	R.A.	Decl.	Redshift	CONFID ^b
J2000				
31	15.0815	28.0534	5.94	2
82	15.0571	28.0859	5.74	2
435	15.0614	28.0779	5.99	1
...
19964	15.0730	28.0115	6.18	2
20317	15.0055	28.0547	6.24	1
20379	15.0375	28.0112	6.77	1

Notes.

^a The coordinates and redshifts shown are tentative. The full catalog will be made publicly available upon publication.

^b Reliability flag: CONFID = 2 (1) for higher (lower) confidence.

(This table is available in its entirety in machine-readable form.)

classified as low-ionization systems and exhibit prominent neutral and/or singly ionized metal absorption lines and absent or very weak absorption in C IV and/or Si IV. All of these low-ionization systems, except the one at $z_{\text{abs}} = 5.35$, are likely to trace neutral gas, as indicated by the detection of O I 1302 Å absorption. A particular goal of the EIGER survey is to search for the host galaxies of these absorbing systems.

Here we highlight the presence of neutral gas at $z \approx 6.14$, approximately ~ 75 cMpc in front of the quasar, that is identified through the detection of a strong O I absorption system, previously reported in Cooper et al. (2019) and Becker et al. (2019). Like all luminous quasars, J0100+2802 enhances the ionization fraction of hydrogen in its local environment, as evidenced by a gradual increase in the Ly α transmission as the quasar’s redshift is approached. Gunn–Peterson absorption begins to saturate at ≈ 72 cMpc in the quasar’s foreground, but the presence of O I and detailed shape of the Ly α transmission both suggest the existence of a damped Ly α absorption system (DLA) at $z = 6.14$, which would fully absorb the Lyman continuum radiation from the quasar along the line of sight. This implies that the foreground IGM will be fully shielded from the ionizing radiation of the quasar below this redshift. This motivates us to use this redshift as a natural boundary to divide the regions with and without the effect of the quasar’s ionizing radiation. In this work, we define the region just above this division redshift (up to the quasar’s redshift) as the so-called “near zone” of the quasar (e.g., Carilli et al. 2010; Venemans et al. 2015; Eilers et al. 2017).

4. Galaxies along the QSO J0100+2802 Sight Line

Here we provide an overview of this first sample of 117 [O III]-emitting galaxies and some of the most immediate findings. These galaxies have [O III] $\lambda 5008$ luminosities $\log L_{[\text{O III}]\lambda 5008}/(\text{erg s}^{-1}) \approx 41.7\text{--}43.4$, UV magnitudes $M_{\text{UV}} \approx -17.7$ to -22.3 , and stellar masses $\log_{10} M_{\text{star}}/M_{\odot} \approx 6.8\text{--}10.1$ (median $10^8 M_{\odot}$). The average star formation rate of the sample is estimated to be $\approx 4 M_{\odot} \text{ yr}^{-1}$ from the H β flux in the stacked spectrum. The measurements of the physical quantities and the detailed characterization of the [O III] emitters are presented in Paper II.

As shown in Figure 7, we detected more galaxies (59%) in the right-hand side region mostly covered by module B than the other side covered by module A. Given the lower sensitivity in module B, this should reflect the real spatial distribution of the galaxies, rather than being an artificial bias.

4.1. Overdensities along the Line of Sight

Figures 7 and 8 show, respectively, the spatial distribution and the transverse separations of the [O III] emitters from the quasar sight line as a function of redshift along the line of sight. The galaxies show a strongly clustered redshift distribution with sharp overdensities at particular redshifts. In particular, three prominent overdensities are clearly visible at $z \approx 6.19$, 6.33, and 6.78, respectively, in front of, precisely at, and behind the quasar redshift. The galaxies within these three specific structures are highlighted in the figures. These member galaxies were defined by grouping galaxies in redshift space with a linking length of 200 km s^{-1} , which is comparable to the velocity dispersion of the two lower-redshift overdensities.

The $z \approx 6.19$ overdensity contains 20 galaxies within $6.176 \leq z \leq 6.195$ with a velocity dispersion $\sigma_v = 212 \text{ km s}^{-1}$. The galaxy spatial distribution is biased toward the right-hand side of the field of view (Figure 7). Figure 9 compares the distribution of [O III] sources and the Ly α transmission fluxes. This structure lies ≈ 57 cMpc in front of QSO J0100+2802, i.e., within the near zone, with strong transmission that is defined to be down to the “natural boundary” of neutral gas traced by the O I absorption system at $z = 6.14$ (see Section 3). Moreover, this overdensity, or at least the member galaxies of it, appears to be physically connected to the high-ionization metal absorption system at $z = 6.19$. The local effect of these galaxies on the surrounding IGM is examined in Section 5.

The second overdensity, containing 24 galaxies with $\sigma_v = 197 \text{ km s}^{-1}$, is exactly at the redshift of QSO J0100+2802, as shown in Figure 9. This is indeed the largest overdensity, in terms of the number of member galaxies, along the line of sight and implies that the SMBH that powers this most luminous quasar formed in a high overdensity of galaxies. The centroid of the galaxy positions is about $\approx 40''$, or ≈ 1.6 cMpc, away from the quasar position. On the other hand, Figure 9 shows some evidence that the velocity dispersion becomes larger at very small separations from the quasar. The velocity dispersion within $r < 1$ cMpc of the quasar is $\sigma_v = 300 \text{ km s}^{-1}$, twice as high as the $\sigma_v = 147 \text{ km s}^{-1}$ for $r > 1$ cMpc ($\sigma_v = 94 \text{ km s}^{-1}$ if limited to $1 < r/\text{cMpc} < 5$). This may reflect a Finger-of-God effect from virialization of the system. The nature of this structure will be examined in the companion papers (A.–C. Eilers et al 2023, in preparation; R. Mackenzie et al. 2023, in preparation).

The third overdensity, with 12 member galaxies and $\sigma_v = 100 \text{ km s}^{-1}$, is behind the quasar and comparable in redshift to the most distant protoclusters previously reported (Hu et al. 2021; Endsley & Stark 2022). Intriguingly for their very high redshifts, the median [O III] $\lambda 5008$ luminosity ($\sim 2.8 \times 10^{43} \text{ erg s}^{-1}$) of the member galaxies in this structure is $>$ six times higher than the average of the other galaxies in the sample ($\sim 4.2 \times 10^{43} \text{ erg s}^{-1}$).

4.2. Associations with Metal Absorption Systems

The deep NIRCcam WFSS observations enable a high-completeness search for host galaxies of the known metal absorption systems along this line of sight. Using the [O III]-emitter sample, we identified host galaxy candidates within $\pm 500 \text{ km s}^{-1}$ and 300 pkpc for four absorption systems ($z_{\text{abs}} = 5.34, 5.95, 6.01, \text{ and } 6.19$) out of the eight known systems. In Figure 8, the absorption systems with host candidates are marked by vertical dotted lines. These candidate

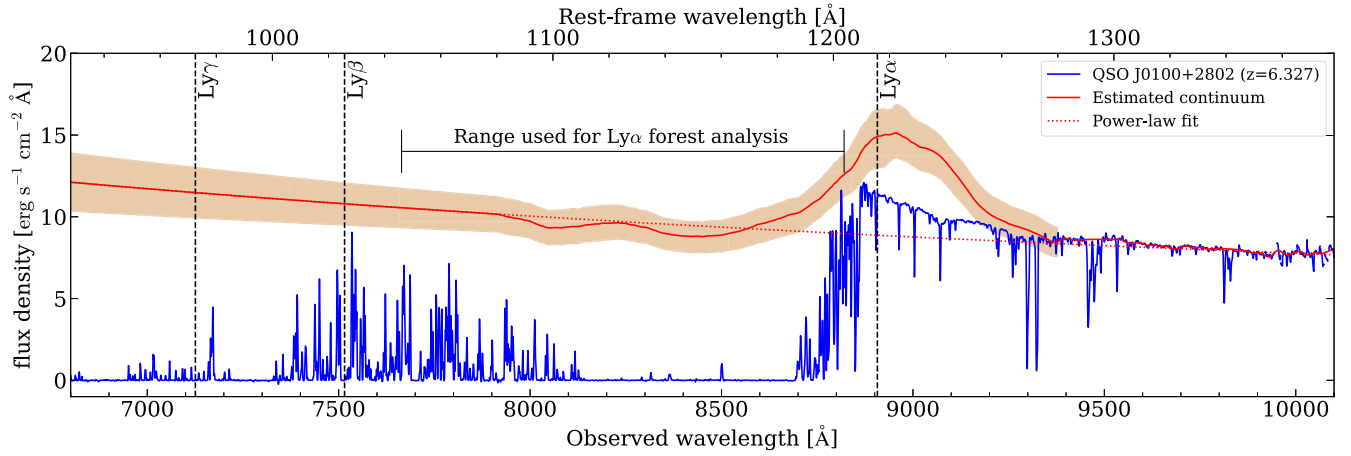


Figure 6. Spectrum of QSO J0100+2802 around the Ly α forest region. The blue line indicates the high-resolution X-Shooter spectrum, and the red solid line is the estimated intrinsic continuum from a deep-learning neural network with a 95% confidence interval (shaded region). Blueward of rest-frame 1080 Å is extended by a power-law fit to the continuum range (red dotted line). The dashed vertical lines indicate the wavelengths of the Ly α , Ly β , and Ly γ lines at the quasar redshift.

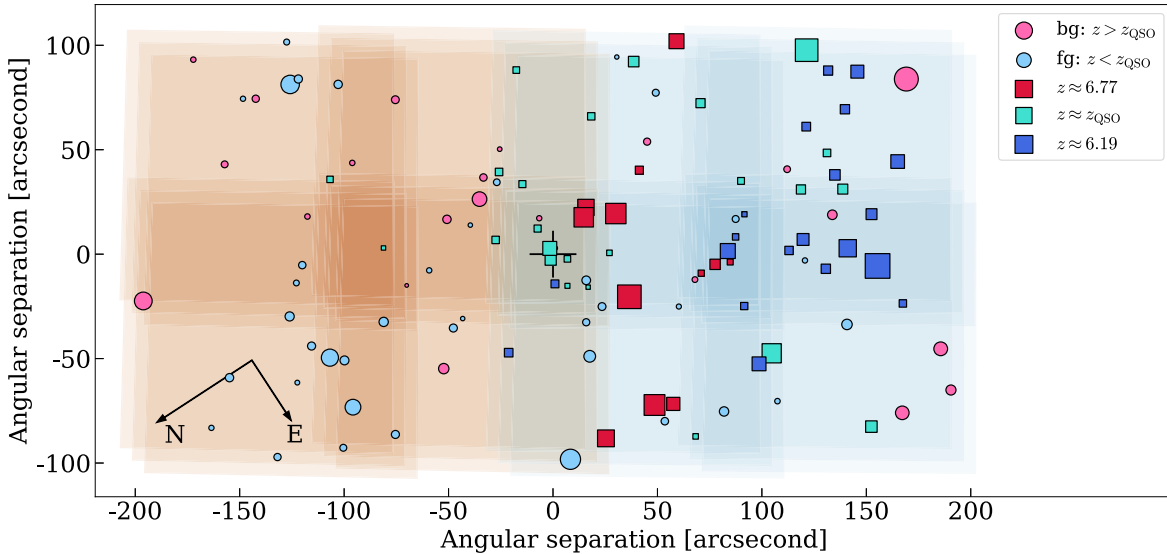


Figure 7. Sky distribution of the detected [O III]-emitting galaxies, with sizes coded according to the relative [O III] luminosity. Galaxies identified as being within the three strongest overdensities are marked by squares. Other galaxies are shown by circles and color-coded according to whether they are in the foreground (light blue) or background (pink) of the quasar. The detector fields of view of the individual exposures are shown so that the color intensity corresponds to the effective exposure time, with red and blue for the module A and B detectors, respectively. The axis values are the angular separations from the position of QSO J0100+2802, which is represented by the large plus sign.

associations in fact represent all of the sample galaxies that lie within 200 pkpc of the sight line within the redshift range studied; in other words, if an EIGER galaxy is within 200 pkpc of the quasar sight line, it has a metal absorption line associated with it.

As noted above, the galaxy associated with the metal absorber at $z=5.33$ was a priori classified as being in the second confidence class (CONFID=1) mostly because this object is at the bluer edge of the sensitivity curve, which makes it hard to assess the [O III] 4960–5008 ratio. However, this coincidence, with a remarkably small impact parameter of (only) 19 pkpc and a velocity offset of $\approx 120 \text{ km s}^{-1}$, further supports that this is truly a high- z [O III] source and associated with the absorber.

Interestingly, absorber-host candidates were identified for all three of the high-ionization systems, including the $z_{\text{abs}} = 6.19$ system coincident with the aforementioned prominent overdensity, but only for one out of the five low-ionization systems.

This could reflect the difference in the galaxy populations and/or physical conditions in the absorption systems at different ionization conditions. A more detailed study of these absorber–galaxy associations, including those identified at lower redshifts via other emission lines, such as H α , will be presented in the companion papers (R. Bordoloi et al. 2023, in preparation; R. Simcoe et al. 2023, in preparation).

In the remainder of this paper, we focus on the subsample of 59 [O III] emitters that are located within the redshift range where the transmission of Ly α and Ly β along the line of sight may be probed by the quasar spectrum.

5. Correlation of Galaxies with IGM Transmission

In this section, we investigate the relation between galaxies and the surrounding IGM during the later stages of reionization by measuring the cross-correlation between the location of

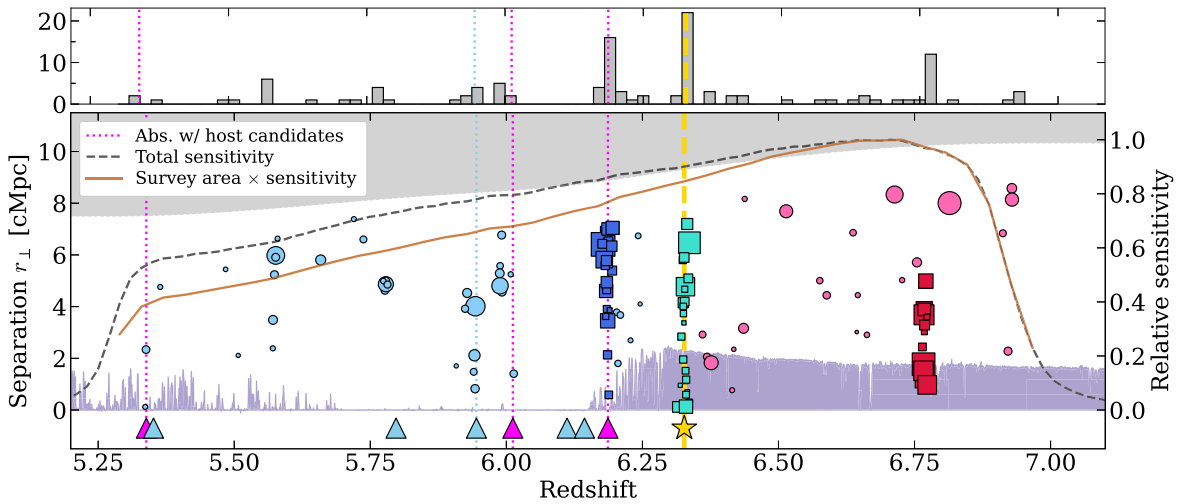


Figure 8. The transverse separation of the [O III] emitters from the quasar sight line is shown as a function of redshift (main panel) with the redshift histogram (subpanel). The symbols follow the same color and size coding as in Figure 7. For reference, the observed QSO J0100+2802 spectrum is shown in arbitrary flux units, where the wavelength is translated into the redshift for Ly α . The gold star and vertical dotted line mark the quasar redshift ($z_{\text{QSO}} = 6.327$). The triangles indicate metal absorption systems in the two classes of low- (cyan) and high-ionization (magenta) systems. The vertical dotted lines mark the absorption systems for which host galaxy candidates were identified. Essentially, all galaxies within 200 physical kpc (pkpc) of the sight line have an associated metal absorption system. The gray shaded region indicates a forbidden area, i.e., where an object is too far from the sight line to be within the field of view covered by the area of the detector. The black dashed curve gives an idea of the redshift sensitivity for the [O III] line because of the filter transmission curve and other effects. The red solid curve then multiplies this by the fraction of the field of view over which an [O III] line at that redshift will fall onto the detector. The scale of these curves is arbitrary and normalized to their peaks.

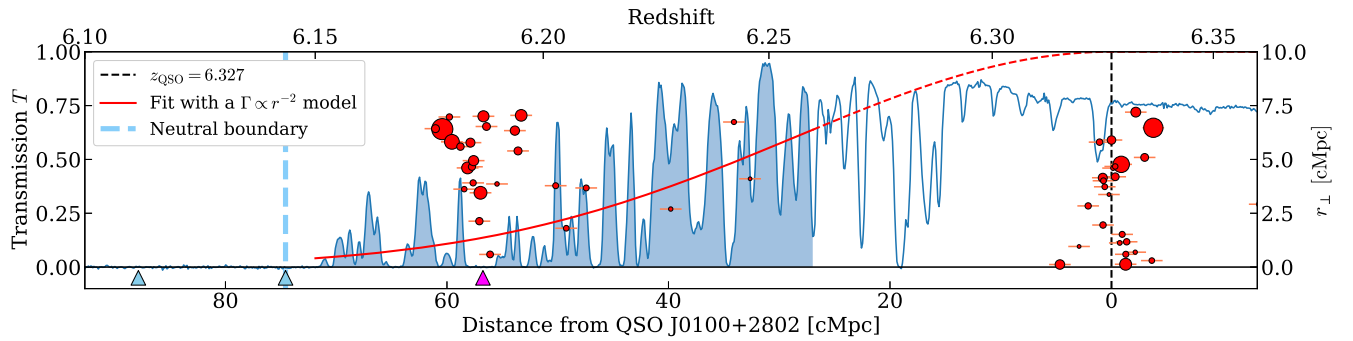


Figure 9. The Ly α transmission in the quasar near zone as a function of radial comoving distance from QSO J0100+2802. The vertical dashed line indicates the location of the quasar ($z_{\text{QSO}} = 6.327$). The red circles show the [O III] emitters, size-coded according to the relative [O III] luminosity. The horizontal error bars correspond to the nominal redshift uncertainties of $\pm 100 \text{ km s}^{-1}$. Strong overdensities are evidently seen at $z \approx 6.19$, within the near zone, and $z \approx 6.33$, coincident with the quasar redshift. The red curve indicates a characterization of the near-zone Ly α transmission profile, assuming that the dominant radiation field $\Gamma \propto r^{-2}$, fitted to the transmission data across the shaded region (see Section 5). The triangles mark the metal absorption systems in the same way as in Figure 8. The vertical blue dashed line represents a neutral system that, it can be safely assumed, blocks all ionizing radiation from the quasar from penetrating to lower redshifts.

[O III]-emitting galaxies and the transmission of the IGM in Ly α and Ly β .

As shown in Figures 6 and 8, the Ly α transmission varies strongly across the range of redshifts accessible. However, we may identify the three different regimes: (1) $5.3 < z < 5.7$, where the Ly α transmission is quite strong and much more ubiquitous than at higher redshifts; (2) $5.7 < z < 6.14$, a relatively dark Gunn–Peterson trough in which Ly α transmission is only seen in isolated discrete transmission spikes; and (3) $z \gtrsim 6.15$, where the transmission increases toward the quasar due to the ionizing radiation from the quasar itself. The division between the first two regimes is somewhat arbitrary. The second division is motivated by the “natural boundary” produced by the existence of the neutral medium traced by the $z = 6.14$ O I metal absorption system, as discussed in Section 4. This system should block ionizing radiation from the quasar from penetrating to all lower redshifts along the line of sight.

Figures 10 and 11 show a zoom-in along the line of sight in the first two redshift regimes identified above. For each galaxy,

the comoving transverse separation is shown vertically as a function of the comoving distance from the quasar with the same scaling as computed (assuming the Hubble flow) from the observed redshifts indicated along the upper horizontal axis. The y-axis on the right-hand side shows the fractional transmitted Ly α flux where the observed wavelength is converted into absorption redshift. Note that the transmitted fluxes are subject to 5%-level systematic uncertainties arising from the uncertainty in predicting the intrinsic continuum spectrum. Figure 9 shows the equivalent plot for the third near-zone redshift regime.

Comparing the Ly α transmission to the distribution of the identified galaxies, we can immediately see clear coincidences between galaxies and transmission spikes at $z \approx 5.785$, 5.99 , and possibly 5.94 (Figure 11). The transmission is, however, likely to be suppressed closer to the galaxies, as seen, e.g., at $z = 5.335$ (Figure 10) and around the $z = 6.19$ overdensity within the quasar near zone. It should also be noted that Ly α transmission is not seen near the known metal-absorber

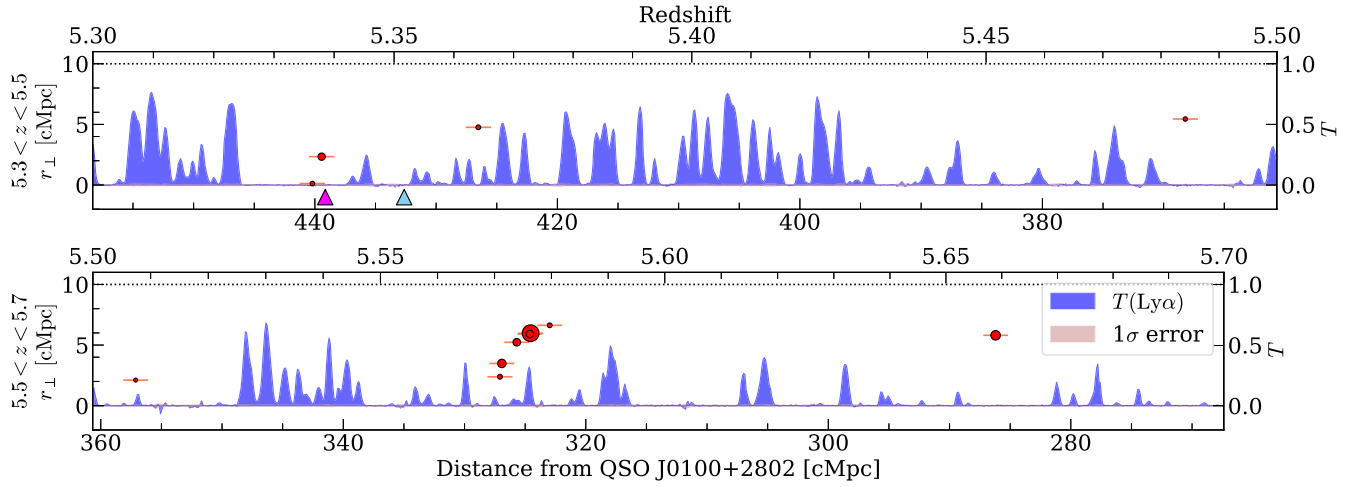


Figure 10. The $\text{Ly}\alpha$ transmission T along the line of sight to QSO J0100+2802 and the comoving transverse separation r_{\perp} of the galaxies from the quasar sight line in the redshift range $5.3 < z < 5.7$, the lowest in our study. The red circles show the radial distances to the detected [O III] emitters from the sight line with different sizes proportional to the [O III] luminosity. The horizontal error bars correspond to the nominal redshift uncertainties of $\pm 100 \text{ km s}^{-1}$. The triangles mark the positions of high- (magenta) and low-ionization (cyan) metal absorption systems. In both panels, the x - and y -axes are scaled equally in comoving distance.

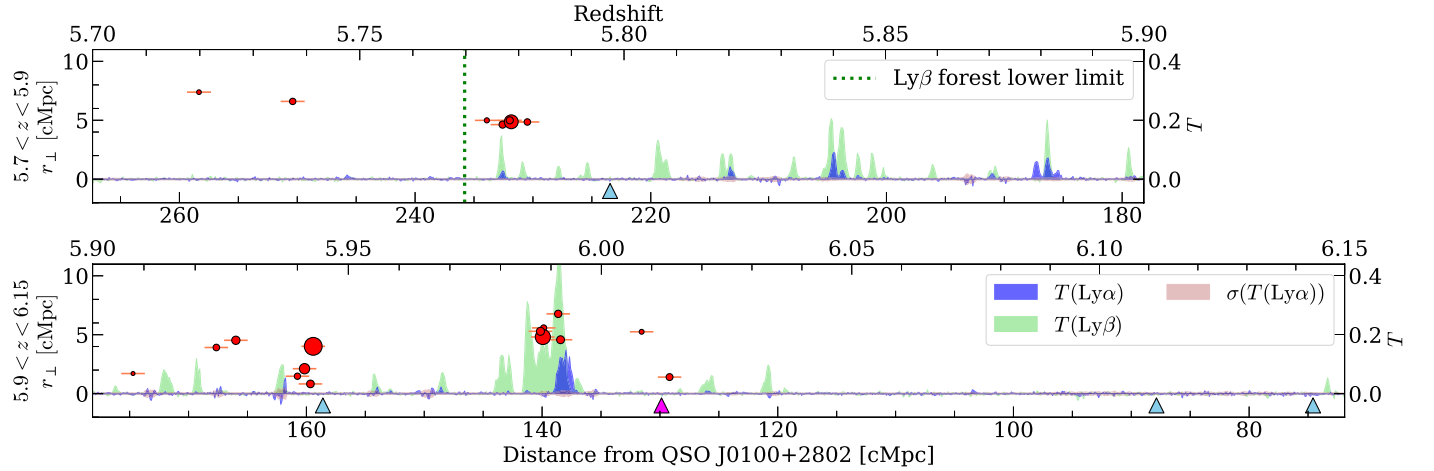


Figure 11. Same as Figure 10 but for the higher redshift range of $5.7 < z < 6.15$. The corresponding $\text{Ly}\beta$ transmission, which is also available in this redshift interval, is also shown in green. The vertical dotted line indicates the lower limit of the $\text{Ly}\beta$ forest range used for analysis ($z_{\text{abs}} = (1 + 6.14)\lambda_{\gamma}/\lambda_{\beta} - 1 = 5.77$; see text).

systems (and thus near their host candidate galaxies). All of these findings accumulate evidence that star-forming galaxies directly impact the surrounding IGM but in a complex manner.

5.1. Average Transmission around Galaxies

To quantify the effects of galaxies on the IGM, we measured the average $\text{Ly}\alpha$ transmission, denoted $\mathcal{T}(r)$, as a function of comoving distance r from the [O III] emitters in these three redshift regimes ($5.3 < z < 5.7$, $5.7 < z < 6.14$, and $6.15 < z < 6.26$).

Figure 12 shows the galaxy- $\text{Ly}\alpha$ transmission cross-correlation, or the transmission curves $\mathcal{T}(r)$. In each of the two lower redshift regimes, we took out the effect of the (relative) redshift evolution within each redshift range by dividing by the cosmic mean transmission as a function of redshift (Eilers et al. 2018) and then multiplying back by the average within the redshift range in question (to revert back to the absolute scale), i.e. we use a modified transmission curve

$$\tilde{\mathcal{T}}(z_{\text{Ly}\alpha}) = \frac{T(z_{\text{Ly}\alpha})}{T_{\text{mean}}(z_{\text{Ly}\alpha})} \times \frac{\int_{z_{\text{min}}}^{z_{\text{max}}} T_{\text{mean}}(z) dz}{z_{\text{max}} - z_{\text{min}}}, \quad (1)$$

where T_{mean} is the cosmic mean transmission. The results change negligibly if we adopt the more recent measurements of the cosmic mean transmission from Bosman et al. (2022).

In the near-zone regime where the ionizing radiation is dominated by the quasar, we adopted a modeled average transmission profile as a function of distance from the quasar, instead of the cosmic mean. This is shown in Figure 9. The model profile assumes that the surrounding IGM density is constant and that the radiation field Γ scales as r^{-2} , where r is the distance from the quasar. This produces $T \propto \exp(-ar^2)$ when the neutral fraction is substantially small. Here the (constant) parameter a , to be determined from the data, reflects the combination of the IGM density and the luminosity of the ionizing radiation.

Because we want to see if the IGM transmission is affected by nearby galaxies, we compared the observed transmission curves, $\mathcal{T}(r)$, with Monte Carlo realizations using random transmission spectra. The random data were generated by partitioning the real quasar spectrum into 0.7 cMpc width bins and shuffling their order within the redshift range. This reshuffling bin size is comparable to the instrumental resolution and thus also to the narrowest possible width of discrete

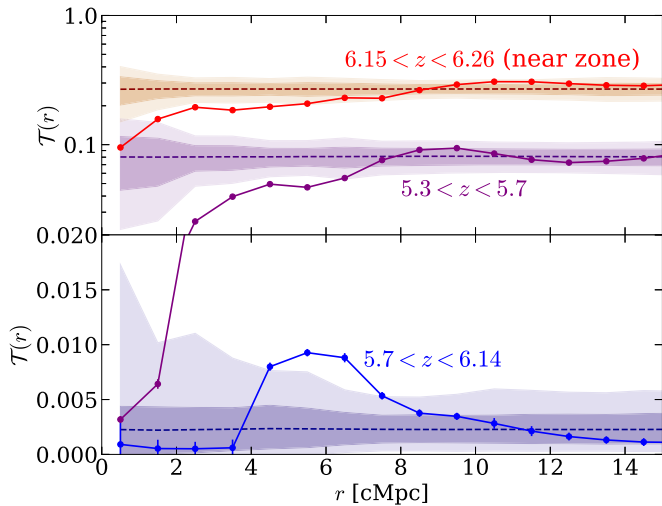


Figure 12. Average Ly α transmission as a function of distance from the [O III]-emitter galaxies in the three redshift regimes $5.3 < z < 5.7$, $5.7 < z < 6.14$, and in the quasar near zone $6.15 < z < 6.26$. The upper and lower panels show the values with log and linear scales to accommodate the wide range in the data. The dashed lines and shaded regions indicate the mean and the central 68th and 95th percentiles of the probability distribution obtained when the transmission along the line of sight is randomly scrambled (see text for details). At the lowest redshifts, the transmission increases monotonically to larger distances. At higher redshifts, as the EoR is approached, the transmission curve shows a pronounced peak around 5–6 cMpc. Finally, in the quasar near zone, where the ionizing radiation from the quasar dominates, the transmission reverts to the monotonic behavior seen at the lowest redshifts.

transmission spikes. We applied Gaussian smoothing with $\sigma = 2$ cMpc to both real and random transmission spectra before measuring the average transmission around the galaxies. As a sanity check, we also evaluated the statistical significance of the measurements by using randomized galaxy catalogs and a bootstrap approach. In the former, we estimated the scatter in the random case (i.e., the case where there is no correlation between galaxies and transmission spikes) by randomizing all of the [O III] emitters in the sky positions and redshift. The obtained scatters are comparable to what is seen in the case shuffling the spikes (as shown in the figure). In the latter, we estimated the error bars around the correlation measurements by using a set of artificial galaxy catalogs that are produced by bootstrapping the sample. These tests have ensured that our results are statistically significant. We limited our analysis to within $r < 15$ cMpc, i.e., $\sim 2\times$ the maximum transverse distance covered by our data (~ 7 cMpc at $z \sim 6$) in the survey area.

Figure 12 illustrates that the overall level of Ly α transmission, observed at large r in the random realization, varies strongly between the redshift ranges. This reflects the global evolution of the cosmic ionizing radiation field with redshift plus the strong near-zone ionization effect due to the quasar in the highest redshift regime.

At the lowest redshift range, the transmission $\mathcal{T}(r)$ is very low near the galaxies and monotonically increases with distance away from them, reaching an asymptotic value at $r \sim 8$ cMpc. This monotonic increase in transmission away from the galaxies is similar to what is observed at lower redshifts (e.g., Tummuangpak et al. 2014; Bielby et al. 2017; Chen et al. 2020; Mukae et al. 2020; Momose et al. 2021). This monotonic increase in $\mathcal{T}(r)$ can be interpreted to be reflecting the ionization of the IGM by a more or less uniform cosmic

ionizing background, coupled with a higher gas density surrounding the galaxies.

In contrast, in the intermediate redshift regime, the average transmission $\mathcal{T}(r)$ first appears to increase at small distances, likely due to the density effect, but then $\mathcal{T}(r)$ clearly peaks at ~ 5 – 6 cMpc away from galaxies before strongly declining toward the lower “average” value at still larger distances. This dramatic change in behavior is evidence that the local ionizing radiation from nearby galaxies is dominating over the cosmic ionizing background on these 5 cMpc scales because of the reduction in the latter, and that the local radiation from these galaxies, or fainter galaxies that are similarly distributed, is responsible for increasing the ionization of the surrounding IGM and thus producing the observed transmission spikes. Establishing whether the ionization is really due to the galaxies that we actually detect or to undetected fainter sources similarly distributed may require, among other things, an estimate of the ionizing photon escape fraction of the detected sources.

Finally, in the quasar near zone ($6.14 < z < 6.26$), the behavior of $\mathcal{T}(r)$ appears to qualitatively revert to the monotonic trend with distance. The average transmission increases with r out to ~ 8 cMpc. While it should be noted that the cross-correlation signal is dominated by the single strong overdensity at $z = 6.19$, this behavior is similar to what is seen in the lowest redshift range (without the quasar) and can be consistently explained as a natural consequence of denser gas around galaxies in a region dominated by a large-scale ionizing radiation field, in this case coming from the exceptionally luminous quasar. This situation may be more similar to that at lower redshifts, $z \sim 4.5$, where the cosmic ionizing background, and thus transmission, is higher.

As stressed earlier, the EIGER galaxy sample is complete in the sense that all galaxies that were detectable in the survey area above our flux limit (which has some spatial variations) should indeed appear in our galaxy catalog. It is clear that some transmission spikes appear along the line of sight without any associated galaxies. Unless there are galaxies similar to those detected in our survey lying just outside the field of view (whose minimum dimension is only ~ 3.8 cMpc from the sight line), the absence of detected galaxies associated with these spikes suggests that they are caused by fainter galaxies.

5.2. Ly β Transmission

To add confidence in this analysis, we also examined the transmission of Ly β in the intermediate redshift range (where the Ly β forest is still observationally accessible). Note that we will only see transmission in the Ly β forest if there is transmission in both Ly β at the redshift in question and Ly α at some lower redshift. Thus, absorption by Ly α (at the lower redshift) may suppress the appearance of Ly β transmission but cannot enhance it. Figure 11 shows that the inferred Ly β transmission is clearly coincident (i.e., occurs at the same redshifts) with the Ly α transmission but also emerges at redshifts where the Ly α absorption has saturated because of its lower oscillator strength.

As noted above, the observed transmission flux gives a lower limit on the Ly β transmission because it may be affected by lower-redshift Ly α absorption. There will be additional contamination from higher-redshift Ly γ absorption at $z_{\text{abs}} < 5.95 \approx (1 + z_{\text{QSO}})\lambda_{\gamma}/\lambda_{\beta} - 1$, where $\lambda_{\gamma} = 972.54 \text{ \AA}$ and $\lambda_{\beta} = 1025.72 \text{ \AA}$. We therefore adopted the redshift range down to $z_{\text{abs}} = 5.77 \approx (1 + 6.14)\lambda_{\gamma}/\lambda_{\beta} - 1$, where the Ly γ

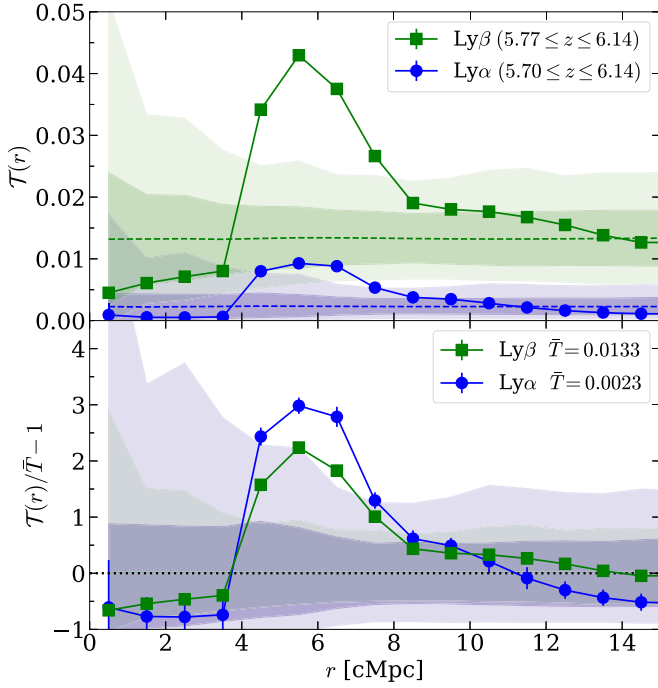


Figure 13. Upper panel: average transmission of Ly β (green squares) as a function of distance from the galaxies at $5.77 < z < 6.14$ compared to that of Ly α in the same redshift range (blue circles; taken from Figure 12). In practice, the Ly β transmission represents a lower limit, since the spectrum is affected by lower-redshift Ly α and higher-redshift Ly γ absorption, but the effect should be independent of distance. Strong enhancement of transmission at 5–6 cMpc is seen in both Ly α and Ly β . Lower panel: transmission curves normalized by the mean transmission \bar{T} at that redshift, which should statistically remove the effect of contaminating Ly α and Ly γ absorption on the Ly β transmission. This comparison emphasizes the similarity of the normalized transmission curves in Ly α and Ly β .

absorption should not be very strong, as this lies within the quasar near zone. Indeed, we do see Ly β transmission down to this redshift.

Figure 13 shows the average inferred (i.e., lower limit) Ly β transmission curve $\mathcal{T}(r)$ around the galaxies compared with that derived above for Ly α in the same redshift regime. Similarly to the analysis of Ly α transmission, we removed the effect of the cosmic evolution within the redshift interval in question using the cosmic mean Ly β transmissivity and the mean foreground Ly α absorption at $z'_{\text{abs}} = (1 + z_{\text{abs}})\lambda_{\beta}/\lambda_{\alpha} - 1$, where $\lambda_{\alpha} = 1215.67 \text{ \AA}$ (Fan et al. 2006; Eilers et al. 2018). Furthermore, the effects of Ly α and Ly γ absorption at different unrelated redshifts (and thus unrelated distances) are taken out by normalizing the transmission by the mean transmission within the redshift range in question, as shown in the lower panel. The normalized transmission curves show quite similar behaviors in Ly α and Ly β , and the coincident strong enhancement in transmission at $r \sim 5 \text{ cMpc}$ is seen in both Ly α and Ly β . This strengthens the evidence for local ionizing effects operating at distances of around 5 cMpc as being responsible for producing the observed transmission spikes.

5.3. Comparison with the THESAN Simulation

To gain further insights into the transmission–distance behavior, we compared these results to the state-of-the-art hydrodynamic simulation THESAN (Garaldi et al. 2022; Kannan et al. 2022; Smith et al. 2022), which is a large-volume ($L_{\text{box}} = 95.5 \text{ cMpc}$) radiation–magnetohydrodynamic

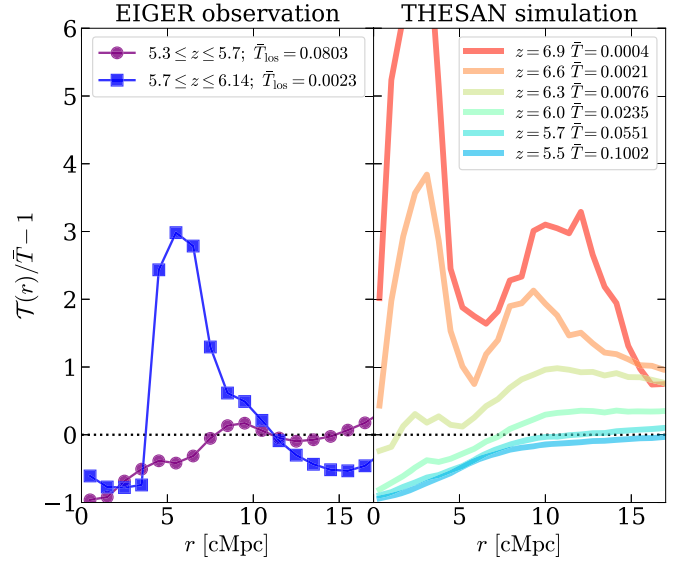


Figure 14. Comparison of the normalized transmitted fluxes that we have observed outside of the quasar near zone (left panel) with the predictions from the hydrodynamic THESAN simulation (right panel). The average transmission in each case is shown in the legend. The monotonic trend with distance from the galaxies that is observed at $z = 5.3\text{--}5.7$ is similar to that in the simulation at $z = 5.5\text{--}5.7$. The strong enhancement of transmission at distances of around 6 cMpc from galaxies at $z = 5.9\text{--}6.14$ is qualitatively similar to that seen in the simulation at $z \gtrsim 6.6$, when the mean transmissivity \bar{T} is actually comparable, but is at somewhat different distances from the galaxies ($r \lesssim 3 \text{ cMpc}$).

simulation that simultaneously models galaxy formation and radiative transport into the IGM. It is convenient for comparisons to convert the average transmission $\mathcal{T}(r)$ relation into the transmission relative to the mean transmission at that redshift. Figure 14 shows this normalized Ly α transmission, $\mathcal{T}(r)/\bar{T} - 1$, as a function of distance from nearby galaxies as obtained from our observations in the lower and intermediate redshift regimes (right panel) and those seen in the THESAN simulations (Garaldi et al. 2022, private communication). The observed transmissions are renormalized to the mean value (\bar{T}_{los}) along our specific sight line within each redshift range under consideration— $\bar{T}_{\text{los}} = 0.0803$ for $5.3 < z < 7.0$ and $\bar{T}_{\text{los}} = 0.0023$ for $5.7 < z < 6.14$. The analysis of the THESAN simulation computes the transmission around galaxies of $M_* > 10^9 M_{\odot}$ at different redshifts $z = 5.5\text{--}6.9$ in the (fiducial) THESAN-1 simulations. The mean transmission value, \bar{T} , at each redshift is denoted in the figure legend.

The monotonic increase of transmission with r from the very low value near to the galaxies up to the cosmic average transmission level at large distances that we observed at $z = 5.3\text{--}5.7$ is similar to the predicted behavior at the same redshift $z \sim 5.5\text{--}5.7$ in THESAN, where the overall mean transmission is also comparable to the observed value.

Interestingly, the THESAN simulations at higher redshifts do show a qualitatively similar peaked transmission curve to that seen in the observational data that we have presented in this paper. Quantitatively, there are differences; the peak first really becomes evident in the THESAN simulation at $z = 6.6$, as opposed to the $z \sim 5.9$ of our observations. It is quite noticeable, however, that the mean transmission along our line of sight at $z \sim 5.9$ is considerably lower than the THESAN \bar{T} at $z = 6.0$ and actually rather similar to that at $z = 6.6$, implying that the reionization history along this line of sight differs from the average but within the observed spatial variations (Bosman

et al. 2018; Eilers et al. 2018; Bosman et al. 2022). In other words, the THESAN curve and our own appear to be much more similar when compared at the same mean transmission value \bar{T} than at the same redshift. Having said that, there are other differences, not least that the peak shows up at somewhat smaller r in THESAN, and the behavior at much larger r also looks different.

Regardless of some differences, this general qualitative similarity of our observational transmission curve with the results of these state-of-the-art simulations should be taken to strengthen our conclusion that the ionizing radiation determining the Ly α transmission of the IGM is dominated by the cosmic ionizing background at low redshift ($z \sim 5.5$) but, at least along this line of sight, by the local ionizing effects of nearby galaxies at $z \sim 5.9$.

We stress that all of the results presented here are based on only a single line of sight. Indeed, the key features in Figure 12, i.e., the peaked shape of $\mathcal{T}(r)$ that is seen at $z \sim 5.9$, are coming from a small number of galaxy overdensities and transmission spikes, and the association between the two. It should be emphasized, however, that these data represent just the first 16% of the EIGER survey, with five additional sight lines to be observed. Their combined measurements will improve these constraints and inform new generations of models. These will also enable us to examine the nature of galaxies owing to the local ionizing radiation (e.g., the mass dependence of the escape fraction), which will be carried out in future papers.

6. Summary

In this paper, we introduced the EIGER survey, a large JWST/NIRCam WFSS campaign that has been designed to produce large samples of [O III]-emitting galaxies at $5.33 < z < 6.93$ in the fields of quasars for which we have very deep high-resolution spectroscopy. We have presented our first set of 117 [O III]-emitting galaxies along the line of sight toward the ultraluminous $z = 6.327$ quasar J0100+2802.

The quasar resides in a strong overdensity traced by 24 galaxies, and two other prominent overdensities also exist, one within the quasar near zone ($z = 6.18$) and the other in the background ($z = 6.78$).

One or more galaxies were identified within 200 pkpc and 105 km s^{-1} of four metal absorption systems, either as their physical host galaxies or as neighboring gas-galaxy associations embedded in a common larger-scale structure.

As the main focus of this paper, we carried out a cross-correlation analysis between the galaxies and the IGM transmission of Ly α and Ly β in three redshift regimes and determined the average transmission as a function of distance from the galaxies, $\mathcal{T}(r)$.

At the lowest redshifts, $5.3 < z < 5.7$, $\mathcal{T}(r)$ increases monotonically with radius away from the galaxies. This effect is observed at lower redshifts and reflects a higher density of gas in the immediate vicinity of the galaxies and the effect of a more or less uniform cosmic ionizing background, similar to that observed at $z \sim 2-3$.

At higher redshifts, $5.7 < z < 6.14$ (where we can be sure that the IGM is shielded from the quasar's ionizing radiation by a DLA), we see a different behavior. There are clear associations between galaxies and transmission spikes with separations of ~ 5 cMpc. The $\mathcal{T}(r)$ curve initially rises away from the galaxies, as before, but sharply peaks at a distance of ~ 5 cMpc, beyond which it rapidly drops again. We interpret

this peak as indicating that the IGM transmission at this redshift is being produced by “local” galactic sources of ionizing radiation that dominate, at 5 cMpc, over the much weakened cosmic ionizing background.

Finally, as we enter the near-zone regime around the quasar, at $6.15 < z < 6.26$, where the extra ionizing radiation from the quasar overwhelms other ionizing sources, the $\mathcal{T}(r)$ curve reverts to the monotonic behavior seen in the lowest redshift range.

Overall, these results represent convincing evidence that local ionizing radiation from galaxies is responsible for producing the last traces of transmission of the IGM at $z \sim 5.9$, at least along this line of sight.

These early results demonstrate the unprecedented power of JWST, particularly the WFSS observing mode of NIRCam, to produce impressively large and complete samples of [O III]-emitting galaxies at $5.3 < z < 7.0$. Companion papers will explore other science topics that are enabled by this data set.

Our EIGER survey will conduct almost identical observations in five other fields, all centered on luminous quasars in the EoR. The full survey, to be completed in the first year of JWST science operations, should enable us to elucidate the processes of cosmic reionization and early galaxy assembly.

We thank Enrico Garaldi, Rahul Kannan, Aaron Smith, and the THESAN team for providing us with their simulation data and Norbert Pirzkal for helpful advice on the NIRCam observing program.

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 were taken under GTO program No. 1243.

The participation of S.L. as an Interdisciplinary Scientist in the JWST Flight Science Working Group 2002–2022 has been supported by the European Space Agency. His earlier involvement in the development of the JWST 1996–2001 was supported by the Canadian Space Agency.





Some of the data used in this work were obtained at the European Southern Observatory under ESO program 096.A-0095 with the 6.5 m Magellan Telescopes located at Las Campanas Observatory, Chile, and at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration.




This work has been supported by JSPS KAKENHI grant No. JP21K13956 (D.K.) and by NASA grant No. HST-GO-15085.001 (R.S.).

Facility: JWST (NIRCam).

Software: Python, numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), Astropy (Astropy Collaboration et al. 2013), Matplotlib (Hunter 2007), MIRaGe (Hilbert et al. 2022), PyeIT (Prochaska et al. 2020a, 2020b).

ORCID iDs

Daichi Kashino  <https://orcid.org/0000-0001-9044-1747>
 Simon J. Lilly  <https://orcid.org/0000-0002-6423-3597>
 Jorryt Matthee  <https://orcid.org/0000-0003-2871-127X>
 Anna-Christina Eilers  <https://orcid.org/0000-0003-2895-6218>

Ruari Mackenzie  <https://orcid.org/0000-0003-0417-385X>
 Rongmon Bordoloi  <https://orcid.org/0000-0002-3120-7173>
 Robert A. Simcoe  <https://orcid.org/0000-0003-3769-9559>

References

- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
 Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, *ApJS*, 227, 11
 Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Natur*, 553, 473
 Becker, G. D., Davies, F. B., Furlanetto, S. R., et al. 2018, *ApJ*, 863, 92
 Becker, G. D., Pettini, M., Rafelski, M., et al. 2019, *ApJ*, 883, 163
 Bentz, M. C., Denney, K. D., Grier, C. J., et al. 2013, *ApJ*, 767, 149
 Bielby, R. M., Shanks, T., Crighton, N. H. M., et al. 2017, *MNRAS*, 471, 2174
 Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, *MNRAS*, 514, 55
 Bosman, S. E. I., Fan, X., Jiang, L., et al. 2018, *MNRAS*, 479, 1055
 Boyer, M. L., Anderson, J., Gennaro, M., et al. 2022, *RNAAS*, 6, 191
 Cain, C., D'Aloisio, A., Gangolli, N., & Becker, G. D. 2021, *ApJL*, 917, L37
 Carilli, C. L., Wang, R., Fan, X., et al. 2010, *ApJ*, 714, 834
 Chen, S.-F. S., Simcoe, R. A., Torrey, P., et al. 2017, *ApJ*, 850, 188
 Chen, Y., Steidel, C. C., Hummels, C. B., et al. 2020, *MNRAS*, 499, 1721
 Christenson, H. M., Becker, G. D., Furlanetto, S. R., et al. 2021, *ApJ*, 923, 87
 Codoreanu, A., Ryan-Weber, E. V., García, L. Á., et al. 2018, *MNRAS*, 481, 4940
 Cooper, T. J., Simcoe, R. A., Cooksey, K. L., et al. 2019, *ApJ*, 882, 77
 Díaz, C. G., Ryan-Weber, E. V., Karman, W., et al. 2021, *MNRAS*, 502, 2645
 D'Odorico, V., Cupani, G., Cristiani, S., et al. 2013, *MNRAS*, 435, 1198
 D'Odorico, V., Finlator, K., Cristiani, S., et al. 2022, *MNRAS*, 512, 2389
 Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017, *ApJ*, 840, 24
 Eilers, A.-C., Davies, F. B., & Hennawi, J. F. 2018, *ApJ*, 864, 53
 Endsley, R., & Stark, D. P. 2022, *MNRAS*, 511, 6042
 Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, *AJ*, 122, 2833
 Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, *AJ*, 132, 117
 Finkelstein, S. L., Bagley, M., Song, M., et al. 2022, *ApJ*, 928, 52
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1
 Garaldi, E., Gnedin, N. Y., & Madau, P. 2019, *ApJ*, 876, 31
 Garaldi, E., Kannan, R., Smith, A., et al. 2022, *MNRAS*, 512, 4909
 Gnedin, N. Y. 2014, *ApJ*, 793, 29
 Gnedin, N. Y., Becker, G. D., & Fan, X. 2017, *ApJ*, 841, 26
 Greene, T. P., Kelly, D. M., Stansberry, J., et al. 2017, *JATIS*, 3, 035001
 Greig, B., & Mesinger, A. 2017, *MNRAS*, 465, 4838
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, 585, 357
 Hilbert, B., Sahlmann, J., Volk, K., et al. 2022, MIRAGE: Multi Instrument Ramp Generator, Astrophysics Source Code Library, ascl:2203.008
 Hu, W., Wang, J., Infante, L., et al. 2021, *NatAs*, 5, 485
 Hunter, J. D. 2007, *CSE*, 9, 90
 Ishimoto, R., Kashikawa, N., Kashino, D., et al. 2022, *MNRAS*, 515, 5914
 Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016, *MNRAS*, 461, 3683
 Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018, *MNRAS*, 478, 4851
 Jiang, L., McGreer, I. D., Fan, X., et al. 2015, *AJ*, 149, 188
 Kakiichi, K., Ellis, R. S., Laporte, N., et al. 2018, *MNRAS*, 479, 43
 Kannan, R., Garaldi, E., Smith, A., et al. 2022, *MNRAS*, 511, 4005
 Kashino, D., Lilly, S. J., Shibuya, T., Ouchi, M., & Kashikawa, N. 2020, *ApJ*, 888, 6
 Keating, L. C., Weinberger, L. H., Kulkarni, G., et al. 2020, *MNRAS*, 491, 1736
 Konno, A., Ouchi, M., Ono, Y., et al. 2014, *ApJ*, 797, 16
 Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, *MNRAS Lett.*, 485, L24
 Liu, B., & Bordoloi, R. 2021, *MNRAS*, 502, 3510
 Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, *ApJ*, 856, 2
 Matejek, M. S., Simcoe, R. A., Cooksey, K. L., & Seyffert, E. N. 2013, *ApJ*, 764, 9
 Matthee, J., Mackenzie, R., Simcoe, R., et al. 2023, *ApJ*, 950, 67
 Matthee, J., Naidu, R. P., Pezzulli, G., et al. 2022, *MNRAS*, 512, 5960
 McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, *MNRAS*, 447, 499
 Merlin, E., Bonchi, A., Paris, D., et al. 2022, *ApJL*, 938, L14
 Meyer, R. A., Bosman, S. E. I., Kakiichi, K., & Ellis, R. S. 2019, *MNRAS*, 483, 19
 Meyer, R. A., Kakiichi, K., Bosman, S. E. I., et al. 2020, *MNRAS*, 494, 1560
 Momose, R., Shimasaku, K., Kashikawa, N., et al. 2021, *ApJ*, 909, 117
 Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Natur*, 474, 616
 Mukae, S., Ouchi, M., Hill, G. J., et al. 2020, *ApJ*, 903, 24
 Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2020, *ApJ*, 892, 109
 Nakajima, K., Ellis, R. S., Robertson, B. E., Tang, M., & Stark, D. P. 2020, *ApJ*, 889, 161
 Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, *Proc. SPIE*, 9143, 91433X
 Prochaska, J., Hennawi, J., Westfall, K., et al. 2020a, *JOSS*, 5, 2308
 Prochaska, J. X., Hennawi, J., Cooke, R., et al. 2020b, pypeit/Pypeit: Release v1.0.0, Zenodo, doi:10.5281/ZENODO.3743493
 Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A&A*, 641, A6
 Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJL*, 802, L19
 Schenker, M. A., Stark, D. P., Ellis, R. S., et al. 2012, *ApJ*, 744, 179
 Schlawin, E., Leisenring, J., Misselt, K., et al. 2020, *AJ*, 160, 231
 Simcoe, R. A. 2006, *ApJ*, 653, 977
 Simcoe, R. A., Burgasser, A. J., Schechter, P. L., et al. 2013, *PASP*, 125, 270
 Simcoe, R. A., Cooksey, K. L., Matejek, M., et al. 2011, *ApJ*, 743, 21
 Smit, R., Bouwens, R. J., Labbé, I., et al. 2014, *ApJ*, 784, 58
 Smith, A., Kannan, R., Garaldi, E., et al. 2022, *MNRAS*, 512, 3243
 Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, *ApJ*, 763, 129
 Tilvi, V., Papovich, C., Finkelstein, S. L., et al. 2014, *ApJ*, 794, 5
 Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, 55, 389
 Tummuangpak, P., Bielby, R. M., Shanks, T., et al. 2014, *MNRAS*, 442, L019
 Venemans, B. P., Bañados, E., Decarli, R., et al. 2015, *ApJL*, 801, L11
 Venemans, B. P., Walter, F., Neeleman, M., et al. 2020, *ApJ*, 904, 130
 Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, 536, A105
 Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, 17, 261
 Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, *Proc. SPIE*, 2198, 362
 Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, *ApJS*, 236, 33
 Wu, X.-B., Wang, F., Fan, X., et al. 2015, *Natur*, 518, 512
 Wu, Y., Cai, Z., Neeleman, M., et al. 2021, *NatAs*, 5, 1110
 Yang, J., Wang, F., Fan, X., et al. 2020, *ApJL*, 897, L14
 Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2021, *ApJ*, 923, 223
 Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2022, *ApJ*, 932, 76