The Astrophysical Journal, 999:177 (21pp), 2026 March 10
https://doi.org/10.3847/1538-4357/ae3a98

Lyα Intensity Mapping in HETDEX: Galaxy-Lyα Intensity Cross-power Spectrum

Maja Lujan Niemeyer[1,2], Eiichiro Komatsu[1,2,3], José Luis Bernal[4], Chris Byrohl[3,5], Robin Ciardullo[6,7], Olivia Curtis[6,7], Daniel J. Farrow[8,9], Steven L. Finkelstein[10,11], Karl Gebhardt[10], Caryl Gronwall[6,7], Gary J. Hill[10,12], Matt J. Jarvis[13,14], Donghui Jeong[6,7,15], Erin Mentuch Cooper[10], Deeshani Mitra[16], Shiro Mukae[10,17], Julian B. Muñoz[10], Masami Ouchi[3,18,19,20], Shun Saito[3,16], Donald P. Schneider[6,7], and Lutz Wisotzki[21]

[1] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany; maja@mpa-garching.mpg.de
[2] Ludwig-Maximilians-Universität München, Schellingstr. 4, 80799, München, Germany
[3] Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of Tokyo Institutes for Advanced Study (UTIAS), The University of Tokyo, Chiba 277-8583, Japan
[4] Instituto de Física de Cantabria (IFCA), CSIC-Univ. de Cantabria, Avda. de los Castros s/n, E-39005 Santander, Spain
[5] Universität Heidelberg, Institut für Theoretische Astrophysik, ZAH, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
[6] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[7] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[8] E. A. Milne Centre for AstrophysicsUniversity of Hull, Cottingham Road, Hull, HU6 7RX, UK
[9] Centre of Excellence for Data Science, Artificial Intelligence & Modelling (DAIM), University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[10] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA
[11] Cosmic Frontier Center, The University of Texas at Austin, Austin, TX, USA
[12] McDonald Observatory, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA
[13] Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK
[14] Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, 7535 Bellville, Cape Town, South Africa
[15] School of Physics, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea
[16] Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N. Pine Street, Rolla, MO 65409, USA
[17] MIRAI Technology Institute, Shiseido Co., Ltd., 1-2-11, Takashima, Nishi-ku, Yokohama, Kanagawa, 222-0011, Japan
[18] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[19] Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan
[20] Department of Astronomical Science, SOKENDAI (The Graduate University for Advanced Studies), Osawa 2-21-1, Mitaka, Tokyo, 181-8588, Japan
[21] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany

Received 2025 November 4; revised 2026 January 12; accepted 2026 January 15; published 2026 March 3

Abstract

We present a measurement of the Lyα intensity mapping power spectrum from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX). We measure the cross-power spectrum of the Lyα intensity and Lyα-emitting galaxies (LAEs) in a redshift range of $1.9 \leqslant z \leqslant 3.5$. We calculate the intensity from HETDEX spectra that do not contain any detected LAEs above a signal-to-noise ratio of 5.5. To produce a power spectrum model and its covariance matrix, we simulate the data using lognormal mocks for the LAE catalog and Lyα intensity in redshift space. The simulations include the HETDEX sensitivity, selection function, and mask. The measurements yield the product of the LAE bias, the intensity bias, the mean intensity of undetected sources, and the ratio of the actual and fiducial redshift-space distortion parameters, $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid} = (6.7 \pm 3.1)$, (11.7 ± 1.4), and $(8.3 \pm 1.5) \times 10^{-22}\,{\rm erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}\,\AA^{-1}}$ in three redshift bins centered at $\bar{z} = 2.1$, 2.6, and 3.2, respectively. The results are reasonably consistent with cosmological hydrodynamical simulations that include Lyα radiative transfer. They are, however, significantly smaller than previous results from cross-correlations of quasars with Lyα intensity. These results demonstrate the statistical power of HETDEX for Lyα intensity mapping and pave the way for a more comprehensive analysis. They will also be useful for constraining models of Lyα emission from galaxies used in modern cosmological simulations of galaxy formation and evolution.

Unified Astronomy Thesaurus concepts: Lyman-alpha galaxies (978); Observational cosmology (1146); Large-scale structure of the universe (902)

1. Introduction

Line intensity mapping (LIM) is a novel tool for studying cosmology and the astrophysics of galaxies and intergalactic gas (see E. D. Kovetz et al. 2017; J. L. Bernal & E. D. Kovetz 2022, for reviews). Intensity maps of one or more emission lines in large volumes can be used as biased tracers of the underlying matter distribution. Instead of detecting galaxies as peaks in the intensity above the noise level, LIM can use noisy data to measure summary statistics of the intensity such as N-point correlation functions. It thus incorporates photons from all galaxies and diffuse gas within the survey volume that would otherwise remain undetected.

Several LIM surveys are currently in operation (O. Doré et al. 2014; M. Santos et al. 2016; D. R. DeBoer et al. 2017; CONCERTO Collaboration et al. 2020; K. A. Cleary et al. 2022; K. S. Karkare et al. 2022), or in preparation (J. Vieira et al. 2020; G. Sun et al. 2021; E. R. Switzer et al. 2021; CCAT-Prime Collaboration et al. 2023, P. Renard et al. 2024, 2025). The emission lines targeted by these surveys span from the ultraviolet (Lyman-α, or short, Lyα) to the radio (21 cm) and are emitted by atomic or molecular gas. LIM has produced detections of cross- and auto-power spectra for the 21 cm and CO lines (e.g., T.-C. Chang et al. 2010; G. K. Keating et al. 2016, 2020; S. Cunnington et al. 2023; S. Paul et al. 2023). Other surveys have provided upper limits on the auto-power spectrum (e.g., K. A. Cleary et al. 2022; R. P. Keenan et al. 2022) or employ stacking (e.g., D. A. Dunne et al. 2024, 2025; Z. Chen et al. 2025).

The Lyα emission line of atomic hydrogen enables observations of high-redshift galaxies, called Lyα emitters (LAEs), and of diffuse gas in the circumgalactic and intergalactic media (CGM and IGM; e.g., M. Ouchi et al. 2020). Stacking of LAEs (e.g., L. Wisotzki et al. 2018; R. Kakuma et al. 2021; S. Kikuchihara et al. 2022; M. Lujan Niemeyer et al. 2022a; S. Kikuta et al. 2023; R. F. Trainor et al. 2025) and other galaxies (C. C. Steidel et al. 2011; H. Kusakabe et al. 2022) has revealed the ubiquity of Lyα emission out to projected distances of $\simeq$1 Mpc from these galaxies. However, few Lyα LIM measurements have been performed on larger scales. All distance units refer to comoving distances unless otherwise specified.

R. A. C. Croft et al. (2016, 2018) and X. Lin et al. (2022) presented a measurement of the cross-correlation of quasars (QSOs) from the Baryon Oscillation Spectroscopic Survey and the Extended Baryon Oscillation Spectroscopic Survey with Lyα intensity from a projected radius of $r_\perp = 0.5\,h^{-1}$ Mpc to $150\,h^{-1}$ Mpc (see X. Lin et al. 2022, for a measurement from $r_\perp = 0.1\,\mathrm{h}^{-1}$ Mpc). The Lyα intensity is measured by subtracting the best-fit galaxy spectrum from spectra of luminous red galaxies, assuming that the residual consists of background Lyα intensity and noise. By measuring the QSO-Lyα intensity cross-correlation, they constrain the average cosmic Lyα luminosity density. However, these results are inconsistent with the upper limit on the cosmic Lyα luminosity density inferred from the Lyα forest, as reported in R. A. C. Croft et al. (2018). While R. A. C. Croft et al. (2018) argued that the intensity is dominated by Lyα emission from QSOs, X. Lin et al. (2022) concluded that the bulk of the Lyα luminosity originates from star-forming galaxies.

The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; K. Gebhardt et al. 2021) provides an ideal dataset for Lyα LIM. HETDEX is a galaxy survey aimed at detecting $\sim 10^6$ LAEs at $1.88 < z < 3.52$ in a comoving volume of 10.9 Gpc^3. HETDEX observes its survey area with the Visible Integral-Field Replicable Unit Spectrograph (VIRUS; G. J. Hill et al. 2021) on the Hobby-Eberly Telescope (HET; L. W. Ramsey et al. 1998; G. J. Hill et al. 2021) without target preselection. Therefore, most fiber spectra do not contain sources that are bright enough to be detected, with detected LAEs comprising only $\sim$0.01% of the signal (E. Mentuch Cooper et al. 2026, in preparation). This makes LIM a unique scientific target for HETDEX.

M. Lujan Niemeyer et al. (2022a, 2022b) stacked the Lyα emission in spectra around HETDEX LAEs and independently detected Lyα-faint galaxies selected by their [O III] emission, and detected Lyα emission out to $\simeq$100 kpc (proper) in both samples. This stacking signal is closely related to the small-scale angular cross-correlation of galaxies with Lyα intensity. M. M. Khanlari et al. (2025) increased the sample size and stacked HETDEX spectra as a function of line-of-sight (LOS) distance and angular separation, finding Lyα absorption halos (see Section 9.4 for a discussion). These studies confirm the suitability of HETDEX data for LIM. The simulation study presented in M. Lujan Niemeyer et al. (2023) (see also J. Fonseca et al. 2017) predicts a high-significance detection of the LAE-Lyα intensity cross-power spectrum at wavenumbers of $k \simeq 0.08$–$1\,h\,\mathrm{Mpc}^{-1}$ for an ideal HETDEX survey without systematics, while masking the spectra of detected LAEs.

In this paper, we present Lyα LIM results from HETDEX observations. Using the HETDEX LAE catalog and HETDEX spectra, we detect the LAE-Lyα intensity cross-power spectrum with high statistical significance. By masking LAEs that are detected above a certain signal-to-noise ratio (SNR) threshold, we ensure that the intensity contains only undetected sources. We generate mocks with the SIMPLE code (M. Lujan Niemeyer et al. 2023) to calculate the transfer function and estimate the power spectrum covariance matrix. By shuffling the on-sky positions of the HETDEX intensity, we calculate the uncertainty due to HETDEX noise. We constrain the product of the LAE bias, the intensity bias, the mean intensity of undetected sources, and the ratio of the actual and fiducial redshift-space distortion factors, $b_\mathrm{g} b_I \langle I \rangle \bar{F}_\mathrm{RSD}/\bar{F}^\mathrm{fid}_\mathrm{RSD}$. Finally, we compare our findings to the expected mean intensity from integrating extrapolated LAE luminosity functions, and to a cosmological Lyα radiative transfer (RT) simulation.

Our measurement is complementary to those of R. A. C. Croft et al. (2016, 2018) and X. Lin et al. (2022) because we cross-correlate the Lyα intensity with different sources, with an emphasis on larger scales, and with higher sensitivity. Because this work does not exclusively use QSOs, but LAEs, our results should not be strongly affected by Lyα emission around QSOs.

This work represents an LIM detection in HETDEX data. We expect to improve data processing and increase the sample size of LAEs and spectra in the future.

This paper is structured as follows. Section 2 describes the LAE catalog and spectra obtained by HETDEX and the data processing. It also describes the separation of the observations into smaller regions and the creation of galaxy and intensity maps. Section 3 presents lognormal mocks from the SIMPLE code. Section 4 explains how we estimate the LAE-Lyα intensity cross-power spectrum. Section 5 describes how we fit the power spectrum model to the data. We show the power spectra and constraints on $b_\mathrm{g} b_I \langle I \rangle \bar{F}_\mathrm{RSD}/\bar{F}^\mathrm{fid}_\mathrm{RSD}$ in Section 6. We compare our constraints with the QSO-Lyα intensity and Lyα forest–Lyα intensity cross-correlations in Section 7. We explore origins of the Lyα intensity in Section 8. Section 9 discusses potential improvements in data processing and modeling, and the possible effect of Lyα absorption on our measurement. We conclude in Section 10.

We use the following Fourier convention:

$$\tilde{f}(\boldsymbol{k}) = \int \mathrm{d}^3\boldsymbol{x} f(\boldsymbol{x}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}},$$
$$f(\boldsymbol{x}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \tilde{f}(\boldsymbol{k}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{1}$$

where the tilde denotes quantities in Fourier space. We adopt a flat Λ cold dark matter cosmology with $H_0 = 67.66\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\mathrm{b},0}h^2 = 0.022$, $\Omega_{\mathrm{m},0}h^2 = 0.142$, $\sum m_\nu = 0.06$ eV, $\ln(10^{10}A_s) = 3.047$, and $n_s = 0.9665$ (Planck Collaboration et al. 2020). We refer to the specific intensity I_λ, the intensity per unit observed wavelength, as the intensity I for simplicity.

2. HETDEX Data

Our analysis uses the data of the internal HETDEX data release HDR5 (E. Mentuch Cooper et al. 2026, in preparation). These data are collected from 2017 January 1 to 2024 July 31 and comprise the full main HETDEX survey. The spectra cover 90 deg^2 of sky, 87.7deg^2 of which meet the science quality criteria. These data include observations within an equatorial "Fall" field, whose footprint covers 150 deg^2, a higher decl. "Spring" field, which covers 390 deg^2, and several smaller fields, which were used for HETDEX science verification (K. Gebhardt et al. 2021). We also use observations of the Texas Euclid Survey for Lyα (TESLA), which covers a 10 deg^2 region of the North Ecliptic Pole (NEP; Ó. A. Chávez Ortiz et al. 2023). The TESLA data were observed in the same manner, processed in the same way, and have the same exposure times as the HETDEX spectra.

The HETDEX spectra were obtained with VIRUS on the 10 m HET (L. W. Ramsey et al. 1998; G. J. Hill et al. 2021). VIRUS is composed of up to 78 integral field unit (IFU) fiber arrays, each of which spans $51'' \times 51''$ on the sky and contains 448 fibers that are $1\farcs5$ in diameter. The fibers from each IFU feed into a spectrograph with two spectral channels. The spectrographs have a resolution of $R \simeq 800$, which corresponds to an FWHM of 5.6 Å at $\lambda = 4500$ Å, and cover the wavelengths between 3470 and 5540 Å. Each spectral channel is read out by a CCD with two amplifiers; thus, the fibers from each IFU are split into four amplifiers. Each HETDEX observation consists of three 6 minute exposures in a triangular dithering pattern that fills the gaps between the fibers on each IFU.

The IFUs are spread out over the HET's $18'$ field of view on a $100''$-grid, so that the gap between each IFU is roughly equal to the size of the IFU. As a result, the gaps between the IFUs lead to a fill factor of $\sim 1/4.6$ for each observation. The observations in the NEP field are dithered to fill the gaps between the IFUs, while these gaps remain in the Spring and Fall fields.

For this work, we only include observations that are not masked by the HETDEX flag for low-quality observations. As HETDEX requires a minimum effective throughput of 0.08 at 4540 Å for an LAE to be included in the catalog, we only select spectra from observations that meet this requirement. Additionally, we exclude HETDEX observations before 2018 because these early data contain more artifacts and larger sky emission residuals. This approach leaves 1949 observations in the Fall field, 3797 in the Spring field, and 484 in the NEP field. Figure 1 shows the IFU coordinates used in the Fall, Spring, and NEP fields.

2.1. LAE Catalog

We select LAEs from catalog 5.0.1 of the internal HETDEX data release HDR5 (E. Mentuch Cooper et al. 2026, in preparation). A public release of an early subset of this sample is found in E. Mentuch Cooper et al. (2023), including a description of source detection and classification. For each line detection, the catalog provides the central wavelength; the sky coordinates of the detection; the line width of the best-fit Gaussian to the emission line (σ); the line flux; the SNR of the detection; an aperture correction factor, which quantifies the fraction of a source's light that falls onto an IFU (as opposed to off an edge); and the source type, e.g., LAE or [O II]-emitting galaxy, and the confidence of this classification.

Because the LAE detections will be masked in the Lyα intensity map, there will be no shot noise in the cross-power spectrum. We restrict our LAE sample to high-confidence objects ("flag_best" detections with SNR > 5.5; see E. Mentuch Cooper et al. 2026, in preparation), preferring a smaller, higher-quality set of LAEs to a larger set with more contamination. The "flag_best" criterion of the HETDEX detection catalog summarizes quality criteria based on known bad data; masks of large galaxies, meteors, and satellites; goodness-of-fit of the line detection; detection confidence based on Machine Learning and citizen science (L. R. House et al. 2023, 2024); and other quality criteria (details can be found in E. Mentuch Cooper et al. 2023; E. Mentuch Cooper et al. 2026, in preparation). We also select only those sources with an aperture correction factor >0.4, and require the updated SNR, which incorporates systematics, to be $\mathrm{SNR}_{\mathrm{rres}} > 4.5$. The noise model used for $\mathrm{SNR}_{\mathrm{rres}}$ is the third model described in K. Gebhardt et al. (2021).

We further impose thresholds on the LAE detection confidence parameters derived from machine learning models. We require $p_{\mathrm{conf}} > 0.5$ from a Random Forest classifier based on emission-line measurements (E. Mentuch Cooper et al. 2026, in preparation) and a score >0.4 from a convolutional neural network (CNN) trained on HETDEX spectra (S. Mukae et al. 2026, in preparation). These classifiers are developed to distinguish LAEs from false-positive detections due to systematic errors in the spectra and sky residuals. For the SNR > 5.5 sample, the adopted CNN threshold yields a precision of $\sim$96% when compared with visual classifications as described in L. R. House et al. (2023), effectively classifying LAEs while minimizing false detections. Using these criteria, 90.4% of detections in the COSMOS science verification field are confirmed in a different HETDEX observation of the same field or a separate galaxy catalog. Thus, the fraction of false-positive detections using these criteria is <9.6%.

We then remove detections within ±10 Å of the [O III] emission line at 5007 Å to mitigate contamination from high-Galactic latitude planetary nebulae. We also exclude LAE candidates within ±10 Å of the sky emission lines at 5200 and 5457 Å.

Figure 2 shows the redshift, luminosity, and line-width distributions of the LAE samples in the Fall, Spring, and NEP fields. The dip at $z \simeq 2.7$ in the redshift distributions is due to a mask that is applied at the center of 50% of the detectors as well as an increase in night sky emission (E. Mentuch Cooper et al. 2023). The dip at $z \simeq 2.3$ is also due to bright sky emission. There remain 37,773 LAEs in the Fall field, 95,774 in the Spring field, and 9372 in the NEP field. After removing detections in IFUs that we mask (see Section 2.2), 36,303 LAEs remain in the Fall field, 92,192 in the Spring field, and 8932 in the NEP field. This number is much lower than the total number of detected LAEs in HETDEX because of our strict selection criteria.

We mask these LAEs in the intensity map for the power spectrum measurement, as described in Section 2.2.3.

2.2. Intensity Spectra

In this section, we describe the additional data reduction steps we apply to the HETDEX spectra. The intensity mapping

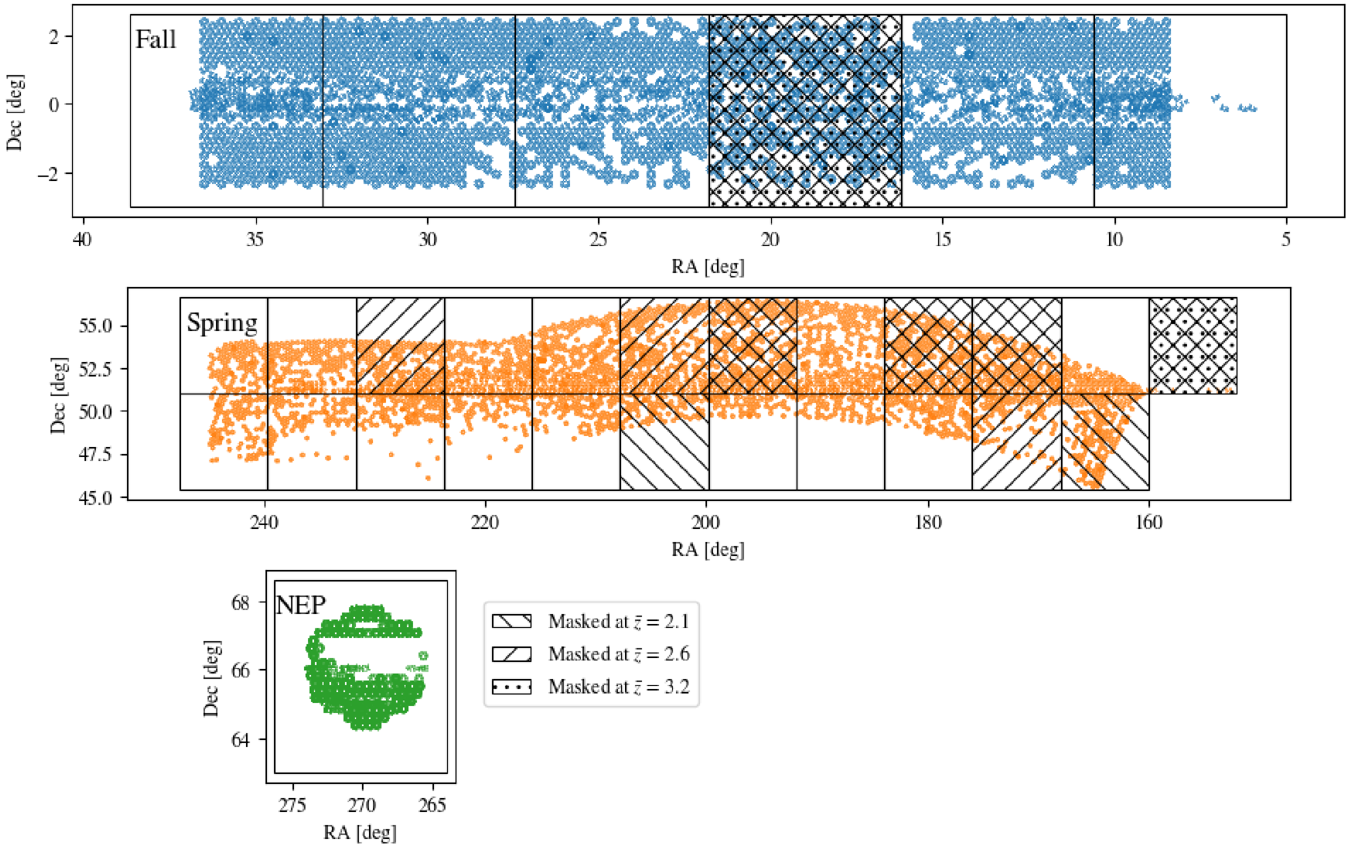

Figure 1. IFU coordinates in the Fall (top panel), Spring (middle panel), and NEP (bottom panel) fields. The fields are divided into 5.°6 wide regions for the power spectrum, shown as black squares (see Section 2.3). Masked maps are shown as hatched regions (see Section 4.2). The backward-facing (forward-facing) diagonal hatchings represent boxes masked in the low-*z* (medium-*z*) bin, and the dotted hatchings indicate those masked in the high-*z* bin.

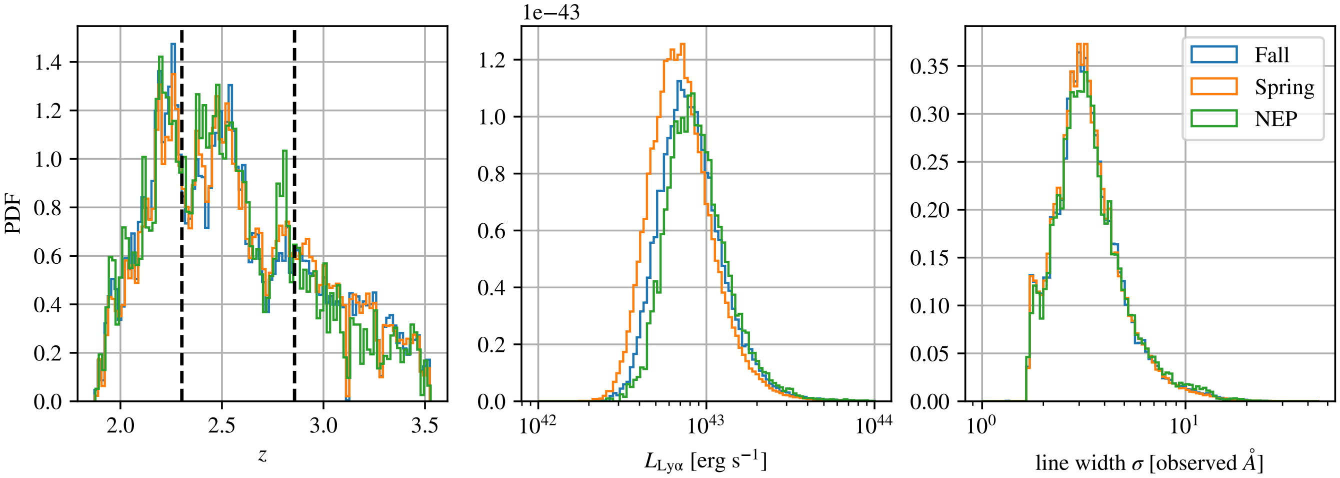

Figure 2. The normalized distributions of redshifts (left panel), luminosities (middle panel), and the best-fit Gaussian σ line widths in observed wavelength units (right panel) of HETDEX LAEs in the Fall (blue), Spring (orange), and NEP (green) fields. The black dashed lines in the left panel show the boundaries of the three redshift bins used to calculate the power spectra.

results can be sensitive to reduction artifacts, and our goal is to reduce those as much as possible.

2.2.1. Removing Sky Emission

K. Gebhardt et al. (2021) described the HETDEX data processing pipeline. A crucial part of the processing for low-surface brightness measurements is sky subtraction. “Sky” emission, including airglow, zodiacal light, and foreground emission from the Milky Way, dominates the observed spectra and must be removed.

HETDEX employs two methods for sky subtraction. The first method uses a “local” sky estimator, measured using the 112 fibers present on each CCD amplifier; these fibers cover a

$\sim$13″ × 51″ field. This sky subtraction technique is optimized for LAE detection. The second method computes a "full-frame" sky, derived from the photons detected on all $\sim$35k fibers of the 18′ field of view of the VIRUS instrument within one exposure (see M. Lujan Niemeyer et al. 2022a, for details). As demonstrated in M. Lujan Niemeyer et al. (2023), the galaxy-intensity cross-power spectrum is suppressed on scales larger than the sky subtraction scale. To keep information on a scale as large as possible, we use the "full-field" sky subtraction in this work.

Our IFU data can be visualized in a two-dimensional figure, where the x-axis is the wavelength, and the y-axis is the fiber number. When averaging the full-frame sky-subtracted spectra of each amplifier over all observations taken within a month, one would expect a mean of zero. However, we find systematic two-dimensional patterns that are unique to each amplifier. To remove these residuals, we first divide each fiber spectrum by the full-frame sky spectrum of the same fiber. We then calculate the bi-weight location (BL) of these relative residuals across observations within a month. The BL and bi-weight scale ($\sigma_{\rm bw}$) are estimators of the central location and scatter of a distribution that are robust to outliers (T. C. Beers et al. 1990). In each observation, we subtract this BL relative residual multiplied by the respective sky spectrum for each fiber.

2.2.2. Masking Bad Detectors, Bright Voxels, Continuum, and [O II] Detections

Processed HETDEX spectra are output in 2 Å bins between 3470 and 5540 Å. When all 78 IFUs are operational, 3 × 34,944 spectra are obtained from each observation. However, some of the CCD amplifiers and a few of the individual fibers produce bad data in certain observations. The bad amplifiers, fibers, and pixels (including those affected by cosmic rays) are flagged by the HETDEX data processing software as described by criteria set in Table 2 of E. Mentuch Cooper et al. (2023), and are masked in our analysis. We also mask spectra containing large galaxies and meteors using the fiber mask provided by the HETDEX Application Programming Interface (API; see E. Mentuch Cooper et al. 2023).

For each observation, we also mask outlier pixels individually. First, we calculate the BL and $\sigma_{\rm bw}$ of all of the spectra, and then mask all pixels with fluxes that deviate by $>3\sigma_{\rm bw}$ from the BL of the fluxes.

Since HETDEX is an untargeted survey, the spectra of continuum objects, such as foreground stars and galaxies, are also in the HETDEX database, and these objects can contaminate the Lyα intensity map. We therefore remove all fiber spectra containing continuum emission. To do this, we estimate the continuum within each fiber by calculating the BL of its spectrum within the wavelength range of 4100–5100 Å, and mask out those fibers with a $3\sigma_{\rm bw}$ deviation from the BL of the continuum values within each observation.

Because light can scatter within the detector, we also mask out fibers adjacent to each masked fiber with continuum emission on the CCD. Spectra adjacent in the array obtained from the HETDEX API are adjacent on the CCD, unless they are the first or last fiber on the CCD. We mask all adjacent fibers in the array, which sometimes includes the first or last fiber on an adjacent amplifier.

We set the spectra of the fibers within 10″ of all [O II]-emitting galaxies to zero. The detection criteria match those used for LAEs, as described in Section 2.1, except that the source type indicates [O II]-emission. For these objects, we do not apply cuts based on the CNN score, $p_{\rm conf}$, or SNR$_{\rm rres}$.

2.2.3. Separating Intensity Contributions from LAEs and Undetected Sources

We separate the spectra into contributions from detected LAEs and undetected sources as described in this subsection. In this step, we consider all LAEs selected in Section 2.1 as "detected." We select voxels within 10″ of an LAE's positional centroid and within 2.5 times the line width of its wavelength centroid. For the "undetected" spectra, we set these voxels to zero. For the "detected" spectra, we instead set all other voxels to zero. Adding the "detected" and "undetected" spectra therefore yields the total spectra. This process translates directly to the power spectra: adding the cross-power spectrum of LAEs with the "detected" intensity to that of LAEs with the "undetected" intensity has to be equal to the cross-power spectrum of LAEs with the "total" intensity. We use this property as a check for the power spectrum pipeline. Henceforth, we apply the same data processing to the "detected," "total," and "undetected" spectra, unless specified otherwise.

2.2.4. Lowering Angular Resolution

The angular resolution of the HET is much better than that of typical LIM experiments: the FWHM of the VIRUS point-spread function (PSF) is between 1.″2 and $\simeq$4″. For computational simplicity, we condense 3 × 448 spectra taken by each IFU during a single HETDEX observation into one spectrum. We have 132,051 IFU spectra in the Fall field, 241,969 in the Spring field, and 34,239 in the NEP field. Each spectrum represents the average flux within 51″ × 51″ × 2 Å voxels, which corresponds to 1.5 Mpc × 1.5 Mpc × 1.9 Mpc at $z = 2.5$. We will henceforth refer to these as "IFU spectral elements."

We calculate the mean of the "total," "detected," and "undetected" fiber spectra within each IFU. We also calculate the standard deviation of the "total" fiber spectra within each IFU and save the mean coordinates of all of the IFU's fibers.

Finally, we calculate the inverse variance of the "total" fiber spectra of each pixel. Some voxels have much higher values than the typical distribution because few unmasked elements of the fiber spectra contribute to it. We therefore set all pixels with inverse variances of $>3600/u_I^2$ to zero, where $u_I = 10^{-17}\,{\rm erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}\,\text{Å}^{-1}}$. This removes everything with a deviation $\gtrsim 5.7\sigma_{\rm bw}$ from the BL of the inverse-variance distribution.

2.2.5. Calculating the Intensity and Removing Extinction

The spectra are output as specific fluxes, f_λ, per 2 Å bin in each fiber. The intensity is obtained from f_λ as $I(z) = f_\lambda(z)/A_{\rm fiber}$, where the fiber area is $A_{\rm fiber} = \pi(0.''75)^2$.

To obtain the intensity as would be seen outside the Milky Way, we remove the reddening caused by Galactic extinction. We collect the color-excess values, $E(B-V)$, as measured by D. J. Schlegel et al. (1998) at the coordinates of each IFU, using the *dustmaps* API (G. Green 2018). Following E. F. Schlafly & D. P. Finkbeiner (2011), we assume an $R_V = 3.1$ reddening model of E. L. Fitzpatrick (1999) and set

$A_V = 2.742E(B-V)$. We calculate and apply the extinction correction using the open-source Python module extinction.[22]

2.2.6. Removing More Systematics

We combine all IFU spectra within each of the three fields for the following steps.

First, we again mask those voxels that are much fainter or brighter than the rest of the distribution, i.e., those with values above $0.0125u_I$ or below $-0.0125u_I$. We determined this threshold empirically; this cut removes everything with a $\gtrsim 7\sigma_{\rm bw}$ deviation from the BL of the pixel distribution, which is $\simeq 0$.

We expect that, after the subtraction of foreground sources and sky emission, the spectra of each IFU will be dominated by sky noise. However, some spectra exhibit systematic patterns, indicating imperfect sky subtraction or other detector problems. Since these spectra typically have a higher standard deviation along the wavelength direction, we mask out those spectra with a standard deviation along the wavelength direction $>0.0035\,u_I$. This maximum value was determined empirically; the cut removes spectra whose standard deviation along the wavelength direction deviates $\gtrsim 3.3\sigma_{\rm bw}$ from the BL of the distribution.

The residuals from sky subtraction are most prominent around the emission lines in the sky foreground, including light from street lamps. We therefore mask all pixels of the IFU spectra at 4350–4370 Å, 5190–5210 Å, and 5447–5467 Å. We also mask ±10 Å around the rest-frame [O III] emission line at 5007 Å to mitigate contamination from planetary nebulae and diffuse gas within the Milky Way.

2.2.7. Principal Component Analysis

Despite the extensive masking described above, the spectra still contain sky emission residuals. We use a principal component analysis (PCA) to remove these sky emission residuals from the spectra.

We perform the PCA on the Fall, Spring, and NEP fields separately. We first subtract the mean spectrum of each field, $\langle I(\lambda)\rangle = N_{\rm IFU}^{-1}\sum_{i=1}^{N_{\rm IFU}} I_i(\lambda)$, and normalize

$$X_i(\lambda) = \frac{I_i(\lambda) - \langle I(\lambda)\rangle}{\sigma_I(\lambda)}, \tag{2}$$

where the subscript i refers to the IFU, and $\sigma_I(\lambda)$ is the standard deviation of the intensity within each wavelength slice. $N_{\rm IFU}$ is the number of IFU spectra used for the analysis, which is 132,051 in the Fall field, 241,969 in the Spring field, and 34,239 in the NEP field. Of these IFU spectra, 5882 (Fall), 13,851 (Spring), and 1803 (NEP) are fully masked in the steps described in Sections 2.2.2, 2.2.4, and 2.2.6. In the remaining fibers, $\simeq 7\%$ of the voxels are masked. We then set all masked intensity values to zero.

We define a vector of length N_λ as

$$\boldsymbol{X}_i = \sum_{j=1}^{N_\lambda} X_i(\lambda_j)\hat{\boldsymbol{e}}_j^{\lambda} = (X_i(\lambda_1), X_i(\lambda_2), \ldots, X_i(\lambda_{N_\lambda}))^{\top}, \tag{3}$$

where $N_\lambda = 1036$ is the number of wavelength bins per spectrum. The jth value of the vector $\boldsymbol{X}_i$ is the normalized intensity fluctuation in the jth wavelength bin, $X_i(\lambda_j)$. The value of the basis vector, $\hat{\boldsymbol{e}}_j^{\lambda}$, is equal to 1 in the jth wavelength bin and zero otherwise.

We define an $N_{\rm IFU}$-by-N_λ matrix, $\boldsymbol{X}$, which contains the vectors $\boldsymbol{X}_i$ as its rows. PCA then calculates an N_λ-by-N_λ covariance matrix, $\boldsymbol{X}^{\top}\boldsymbol{X}$. The eigenvectors of the covariance matrix are the weight vectors. These form an orthonormal basis $\hat{\mathbf{e}}_j^{\rm PCA}$, such that

$$\boldsymbol{X}_i = \sum_{j=1}^{N_\lambda} y_{ij}\hat{\boldsymbol{e}}_j^{\rm PCA}. \tag{4}$$

PCA arranges the eigenvectors $\hat{\mathbf{e}}_j^{\rm PCA}$ in order of decreasing variance of y_{ij}, with the variance being highest for $j = 1$. The variance of y_{ij} is proportional to the jth eigenvalue.

Figure 3 shows the first seven PCA weight vectors, $\hat{\mathbf{e}}_j^{\rm PCA}$ with $j = 1$–7, of the "total" spectra in the fall field; an example "full-field" estimated sky spectrum; and the eigenvalues corresponding to the first 200 weight vectors. The first weight vectors contain features of the sky spectrum, such as the K and H lines at 3934 and 3968 Å.

The high-variance weight vectors are dominated by systematics such as sky emission residuals. We therefore set the first $N_{\rm PC} \in \{10, 20, \ldots, 200\}$ of the principal components (PCs) to zero and work with

$$\boldsymbol{X}_i^{\rm PCA} = \sum_{j=N_{\rm PC}+1}^{N_\lambda} y_{ij}\hat{\boldsymbol{e}}_j^{\rm PCA}. \tag{5}$$

To obtain clean spectra, we perform an inverse coordinate transformation,

$$\boldsymbol{X}_i^{\rm PCA} = \sum_{k=1}^{N_\lambda}\ \sum_{j=N_{\rm PC}+1}^{N_\lambda} y_{ij}(\hat{\boldsymbol{e}}_k^{\lambda\top}\hat{\boldsymbol{e}}_j^{\rm PCA})\hat{\boldsymbol{e}}_k^{\lambda}, \tag{6}$$

so that

$$X_i^{\rm PCA}(\lambda_k) = \sum_{j=N_{\rm PC}+1}^{N_\lambda} y_{ik}(\hat{\boldsymbol{e}}_k^{\lambda\top}\hat{\boldsymbol{e}}_j^{\rm PCA}), \tag{7}$$

and undo the normalization:

$$I_i^{\rm PCA}(\lambda) = X_i^{\rm PCA}(\lambda)\sigma_I(\lambda). \tag{8}$$

We then re-apply the mask.

We calculate the weight vectors, $\hat{\mathbf{e}}_j^{\rm PCA}$, from the covariance matrix $\boldsymbol{X}^{\top}\boldsymbol{X}$ of the normalized "total" spectra. We then use these vectors to clean the "total," "detected," and "undetected" spectra. This ensures that the sum of the "detected" and "undetected" spectra equals the "total" spectra. Since most spectral elements do not contain a detected LAE, the "total" and "undetected" spectra are similar.

The PC removal also removes cosmological signal. We model this loss in the mocks in Sections 3.2 and 4.3.

2.2.8. Comparison with Noise

Figure 4 compares the measured and expected standard deviations of the IFU spectra in the absence of astrophysical sources, as a function of wavelength. We translate the measured 5σ sky and read noise per fiber per spectral resolution element (FWHM = 5.6 Å) of Figure 18 in G. J. Hill et al. (2021) to the intensity noise per fiber by dividing it by $(5 \times \pi(0\farcs75)^2 \times 5.6\text{Å})$. This assumes that the noise of the pixels within a resolution element is fully correlated, as expected if the noise is dominated by sky photons. The noise is

---

[22] https://github.com/sncosmo/extinction/.

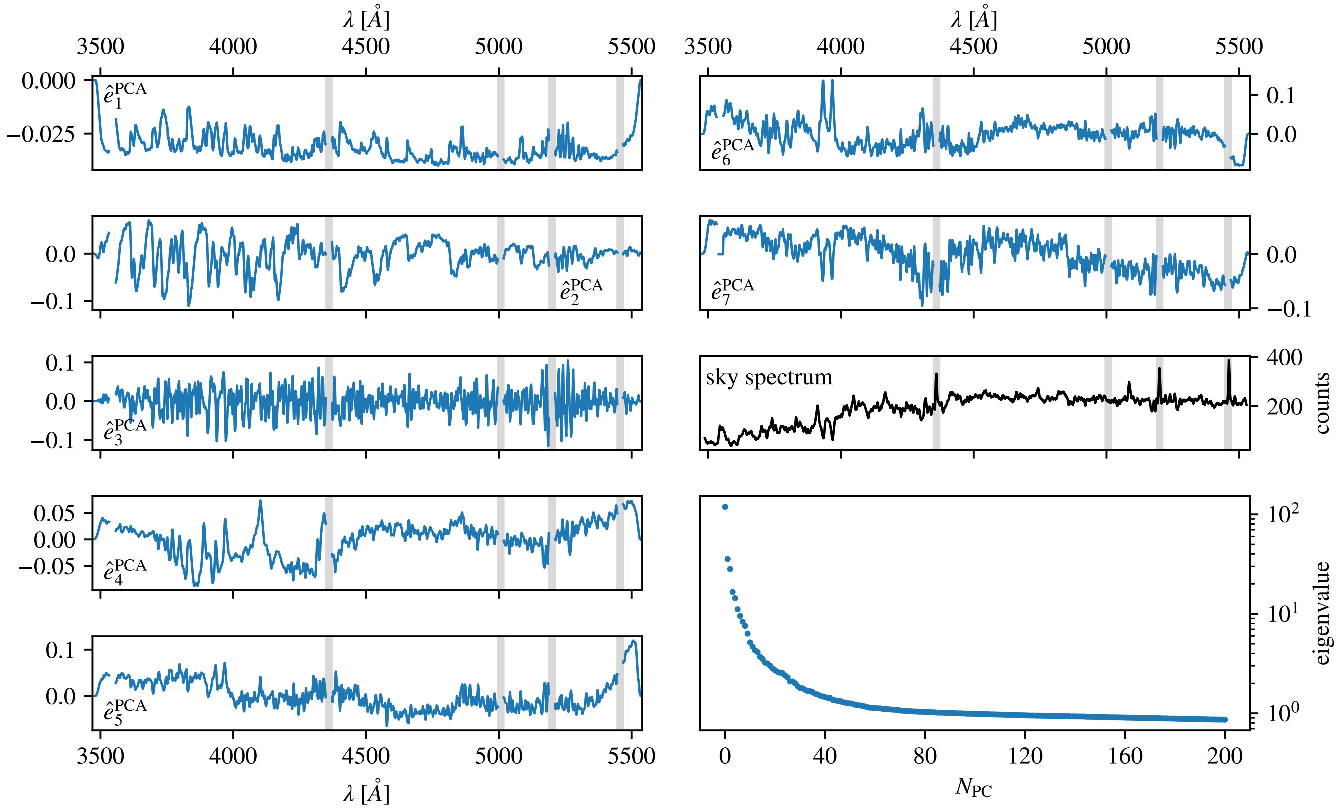

Figure 3. First seven weight vectors, $\hat{\boldsymbol{e}}_j^{\rm PCA}$ with $j = 1$–7, of the Fall field are displayed as functions of wavelength. All $\hat{\boldsymbol{e}}_j^{\rm PCA}$ are normalized vectors. For comparison, we also show an example "full-field" sky spectrum in counts (third panel on the right). Masked wavelength regions are shown in gray. The bottom-right panel shows the first 200 eigenvalues, which are proportional to the variance along the corresponding weight vectors.

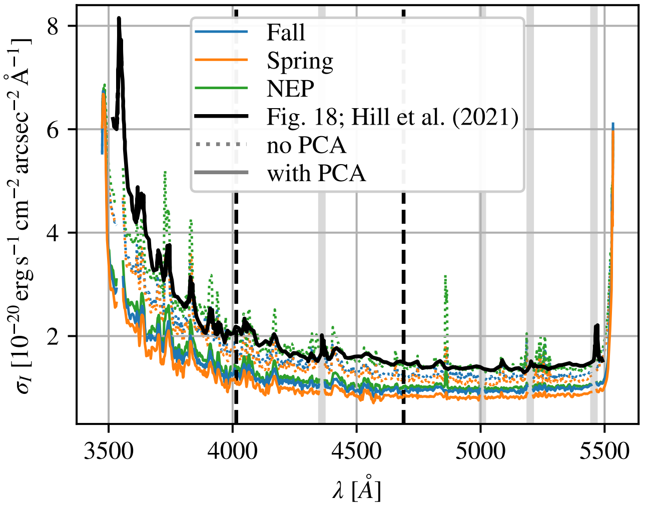

Figure 4. Standard deviation of the IFU spectra in each wavelength slice in the Fall (blue), Spring (orange), and NEP (green) fields compared to the pure-noise expectation inferred from Figure 18 of G. J. Hill et al. (2021; black). The dashed colored lines show the standard deviation of the IFU spectra without PCA ($N_{\rm PC} = 0$). The solid colored lines are the standard deviation of the IFU spectra after PCA cleaning, where $N_{\rm PC} = 170$ in the low-z bin, $N_{\rm PC} = 130$ in the medium-z bin, and $N_{\rm PC} = 100$ in the high-z bin. The vertical black dashed lines indicate the boundaries of the three redshift bins used for power spectrum measurements. The gray shaded areas show masked wavelength values.

indeed dominated by sky photons at $\lambda \gtrsim 4000$ Å (G. J. Hill et al. 2021). At shorter wavelengths, it is read-noise dominated, which is not correlated within a spectral resolution element. The expected intensity noise per IFU spectrum is obtained by dividing this value by $\sqrt{3 \times 448}$ for three dithers and 448 fibers per IFU per dither. This measured standard deviation of the IFU spectra without PCA cleaning is similar to the expectation inferred from G. J. Hill et al. (2021) in the NEP field, but slightly smaller at most wavelengths in the Fall and Spring fields. This discrepancy could be due to masking outlier pixels and high-variance spectra. After PCA cleaning, the standard deviation is smaller at most wavelengths. This is a natural consequence of our removal of PCs with the highest variance.

2.3. Creation of Intensity Maps

Due to the loss of power from sky subtraction (see M. Lujan Niemeyer et al. 2023), we measure the power spectrum at wavenumbers $k > 0.08\,h\,{\rm Mpc}^{-1}$. To accommodate the required high angular resolution, we split the three fields into smaller regions with a side length of 5°.6 for the maps.

We divide the redshift range into three redshift bins from 1.85–2.3 ($\bar{z} = 2.07$), from 2.3–2.86 ($\bar{z} = 2.58$), and from 2.86–3.56 ($\bar{z} = 3.21$). This binning yields 23 × 3 = 69 maps for the Spring field and 6 × 3 = 18 maps for the Fall field. The NEP field fits within one such map in each of the three redshift bins. Figure 1 shows the map division in the three fields.

Each map is contained within a cubic box with a side length of $432\,h^{-1}$ Mpc for the low-z and medium-z bins and $463\,h^{-1}$ Mpc for the high-z bin. The fundamental frequencies are $k_{\rm F} = 0.015\,h\,{\rm Mpc}^{-1}$ and $0.014\,h\,{\rm Mpc}^{-1}$, respectively. We create cubic voxels that are $2\,h^{-1}$Mpc long, corresponding to a Nyquist frequency of $k_{\rm Ny} = 1.57\,h\,{\rm Mpc}^{-1}$.

To create the intensity map, we collect all of the IFU spectral elements contained within each region and redshift bin. Then, we transform the sky and redshift coordinates into Cartesian coordinates and collect the IFU spectral elements in each voxel.

For each voxel, we calculate the mean intensity $I(\boldsymbol{x})$ of the IFU spectral elements within. We then calculate the intensity fluctuation, $\delta I(\boldsymbol{x})$, in each voxel,

$$\delta I(\boldsymbol{x}) = I(\boldsymbol{x}) - \langle I \rangle (x_0), \tag{9}$$

where $\boldsymbol{x} = (x_0, x_1, x_2)$, x_0 is the position along the LOS axis of the map, x_1 and x_2 are the coordinates perpendicular to the LOS, and $\langle I \rangle (x_0)$ is the mean intensity in each slice along the LOS direction of the map. We calculate $\langle I \rangle (x_0)$ as

$$\langle I \rangle (x_0) = \frac{1}{N_{\rm vox}^2} \sum_{i=1}^{N_{\rm vox}} \sum_{j=1}^{N_{\rm vox}} I(x_0, x_1^i, x_2^j), \tag{10}$$

where x_1^i and x_2^i are the coordinates of the ith voxel, and $N_{\rm vox} = L_{\rm box}/L_{\rm vox}$ is the number of voxels in each direction of $\boldsymbol{x}$. Here, $L_{\rm box}$ is the length of the box, and $L_{\rm vox}$ is the length of a voxel.

2.4. Mean Expected Number of LAEs per Voxel

We define the galaxy overdensity as

$$\delta_{\rm g}(\boldsymbol{x}) = \frac{N_{\rm LAE}(\boldsymbol{x}) - \bar{N}(\boldsymbol{x})}{\bar{N}(\boldsymbol{x})}. \tag{11}$$

The number of LAEs detected per voxel is $N_{\rm LAE}(\boldsymbol{x})$, and the expected number of LAEs detected per voxel at location $\boldsymbol{x}$ in the absence of clustering is $\bar{N}(\boldsymbol{x})$. This parameter can be calculated as

$$\bar{N}(\boldsymbol{x}) = \delta V(\boldsymbol{x}) \int_{L_{\rm min}}^{L_{\rm max}} dL \, \frac{dn}{dL} C\!\left(\frac{L}{4\pi D_L^2(z)}, f_{50}(\boldsymbol{x})\right), \tag{12}$$

where $\delta V(\boldsymbol{x})$ is the volume covered by spectra within the voxel at position $\boldsymbol{x}$, $\frac{dn}{dL}$ is the luminosity function of LAEs, D_L is the luminosity distance, and f_{50} is the flux limit at which 50% of LAEs will be detected. The completeness factor, C, is a function of LAE line flux and is characterized by the 50% completeness, f_{50}. Examples of completeness curves as a function of flux are shown in Figure 28 of K. Gebhardt et al. (2021).

We use the completeness model v4 of the HETDEX API[23] (K. Gebhardt et al. 2021; D. Farrow et al. 2026, in preparation). This completeness function depends on the ratio of the LAE flux to the flux at 50% completeness; $C(f, f_{50}) = C_{\rm v4}(f/f_{50})$, which is a nonparametric model to map f/f_{50} to detection completeness based on inserting simulated LAEs into the data. The value of f_{50} is obtained by multiplying the flux noise, $\sigma_f(\boldsymbol{\Theta}, \lambda)$, at the sky coordinate $\boldsymbol{\Theta}$ and wavelength λ by the minimum SNR for an emission-line detection (here 5.5, as described in Section 2.1), a polynomial dependence on the wavelength of the detection $a_\lambda(\lambda)$, and a line-width dependent adjustment (D. Farrow et al. 2026, in preparation). The value of $\sigma_f(\boldsymbol{\Theta}, \lambda)$ is calculated as a PSF-weighted sum of the spectral errors within an aperture 3.″5 in diameter and 14 Å deep.

We collect the values of f_{50} in each IFU spectral element. We extract the flux noise $\sigma_f(\boldsymbol{\Theta}, \lambda)$ on a 2″ × 2″ × 2 Å grid for each IFU using the HETDEX API. The redshift bins correspond to the 2 Å grid of wavelengths between 3470 and 5540 Å of the spectra. Because the API also provides a mask on the grid, we obtain the volume covered per IFU spectral element, $\delta V_i(z)$, for line detections. We transform these values to $f_{50}(\boldsymbol{\Theta}, \lambda) = 5.5\sigma_f(\boldsymbol{\Theta}, \lambda) a_\lambda(\lambda)$, not accounting for the line-width adjustment. For each IFU and wavelength, we calculate a histogram of f_{50} values in 5001 logarithmically spaced bins between 10^{-18} and 10^{-13} erg s^{-1} cm^{-2}.

We calculate the number density, $\bar{n}(f_{50}, z)$, for each bin center of the f_{50} histogram and for each redshift bin corresponding to the 2 Å wavelength grid. We use the best-fit Schechter function component (P. Schechter 1976) of the $z = 2.2$ Lyα luminosity function measured by H. Umeda et al. (2025); with $L^* = 10^{42.8}$ erg s^{-1}, $\phi^* = 10^{-3.16}$ Mpc^{-3}, and $\alpha = -1.53$ for $z \leqslant 2.3$, corresponding to the low-z bin for the power spectrum measurement. In the medium-z and high-z bins ($z > 2.3$), we adopt the $z = 3.3$ luminosity function of H. Umeda et al. (2025) with $L^* = 10^{42.29}$ erg s^{-1}, $\phi^* = 10^{-2.13}$ Mpc^{-3}, and $\alpha = -1.19$.

We evaluate Equation (12) by translating the f_{50} histogram, $N_{f_{50}}^{(i)}(z)$, in the ith IFU into the number of expected LAEs, $\bar{N}_i(z)$, in the absence of clustering:

$$\bar{N}_i(z) = \sum_{f_{50}} \delta V_i(z) \bar{n}(f_{50}, z) N_{f_{50}}^{(i)}(z). \tag{13}$$

We save $\bar{N}_i(z)$, the covered volume $\delta V_i(z)$, and the mean value of f_{50},

$$\bar{f}_{50}^{(i)}(z) = \frac{\sum_{f_{50}} f_{50} N_{f_{50}}^{(i)}(z)}{\sum_{f_{50}} N_{f_{50}}^{(i)}(z)}. \tag{14}$$

The value of $\bar{N}_i(\lambda)$ predominantly depends on the observing conditions and the sky spectrum. It therefore varies from IFU to IFU, although it is similar for different IFUs within the same observation.

Equation (13) overestimates the number of detected LAEs by a factor of ≃2. This could be due to an overestimated luminosity function, especially at the faint end, or an imperfect model for f_{50}. To avoid a systematic power spectrum signal resulting from this discrepancy, we ensure that the redshift distribution of $\bar{N}(z)$, averaged over the IFUs, matches that of the detected LAEs in each field. However, we still need $\bar{N}_i(z)$ to model the IFU-to-IFU variations. We therefore multiply $\bar{N}_i(z)$ by the ratio, $f_N(z) = N_{\rm LAE}(z)/\bar{N}(z)$. We calculate this ratio separately for the Spring and Fall fields. We use $f_N(z)$ from the Fall field for the NEP field due to the low number of LAEs. The values of $f_N(z)$ are mostly around 0.5, with individual spikes up to ≃3. After this correction, 84% (95%) of the values of $\bar{n}_i(z) = \bar{N}_i(z)/\delta V_i(z)$ are below 4, 6, and 4 (7, 10, and 6) $\times 10^{-4}$ Mpc^{-3} h^3 in the Fall, Spring, and NEP fields, respectively.

To create a map of the LAE overdensity, we first transform the number of LAEs, $N_{\rm LAE}$, and the expected $\bar{N}$ into the same format as the spectra. If the mean total spectrum at an IFU spectral element is masked, then all other values, including $N_{\rm LAE}$ and $\bar{N}$, are also masked at this IFU spectral element. We then calculate the total number of detected LAEs $N_{\rm LAE}(\boldsymbol{x})$, the expected number of detected LAEs in the absence of clustering $\bar{N}(\boldsymbol{x})$, in each map voxel to create a map of $\delta_{\rm g}$. We also save the mean f_{50} in each voxel and use it for the mocks.

---

[23] https://github.com/HETDEX/hetdex_api

3. Mock Data

We generate mock data to compare the measured LIM power spectra with a theoretical expectation. These mocks allow us to quantify the effect of the data cleaning, and to calculate the covariance matrix of the power spectra.

3.1. General Setup

To create a model for the power spectra and estimate their covariance matrices, we generate mock data using the SIMPLE code[24] (M. Lujan Niemeyer et al. 2023). This package generates galaxy and intensity maps with nonlinear redshift-space distortions (RSD) due to peculiar velocities (A. Agrawal et al. 2017).

We use the same box size, voxel size, and redshifts as in the HETDEX maps (see Section 2.3). The redshifts used to calculate the intensity of the mock data vary along one axis. We generate 100 mocks for each redshift bin and each region shown in Figure 1, henceforth called boxes, i.e., 6900 for the Spring field, 1800 for the Fall field, and 300 for the NEP field. We input the nonlinear matter power spectrum calculated at the three mean redshifts using the public CLASS code[25] (D. Blas et al. 2011) with the default Halofit model parameters (R. E. Smith et al. 2003; S. Bird et al. 2012; R. Takahashi et al. 2012). We compare the mock power spectra from the nonlinear and linear matter power spectra in Appendix A.

We set the LAE bias to $b_{\rm mock} = 2$, which is consistent with the measurements of $b_{\rm LAE} = 1.7^{+0.3}_{-0.4}$ at $z = 3.1$ (E. Gawiser et al. 2007), 1.8 ± 0.3 at $z = 2.1$ (L. Guaita et al. 2010), and $1.72^{+0.26}_{-0.27}$ and $2.01^{+0.26}_{-0.29}$ at $z = 2.4$ and 3.1, respectively (D. Herrera et al. 2025). In this setup, detected and undetected LAEs in the mock have the same bias parameter; thus, the LAE bias is $b_{\rm g} = 2$, and the intensity bias is $b_I = 2$. Our simulations use the same luminosity functions of H. Umeda et al. (2025) as we did for the $\bar{N}$ calculation described in Section 2.4, and we set the minimum Lyα luminosity to $4 \times 10^{40}\,{\rm erg\,s^{-1}}$. This minimum luminosity mainly determines the number of simulated LAEs and the mean intensity. However, a smaller value barely changes the mean intensity. We apply a $\sigma_\lambda = 2.38$ Å LOS Gaussian smoothing algorithm to the intensity maps in order to imitate the VIRUS spectral resolution. Since the VIRUS PSF is much smaller than the voxel size, we do not smooth in the angular direction.

To determine which simulated LAEs are detected, we randomly draw from the simulated LAEs. The probability that a simulated LAE with flux f at position $\boldsymbol{x}$ will be detected is given by the completeness function, $C(f, f_{50}(\boldsymbol{x}))$, described in Section 2.4, where $f_{50}(\boldsymbol{x})$ is the mean of the values of f_{50} within the voxel.

3.2. Inserting the Observed Data into the Mock Data

To create the same mask for the data and mocks, we convert the mock galaxy and intensity maps to the same format as the IFU spectra. For each IFU spectral element, we select the corresponding map and voxel in the mock. We save the mock "total," "detected," and "undetected" intensities and the number of detected galaxies in the voxel.

The volume of a voxel, $V_{\rm vox}$, is larger than the volume covered by one IFU spectral element, $V_{{\rm IFU}\times\Delta\lambda}$. Therefore, the number of galaxies that can be detected in the spectra is less than the number calculated for the voxels by a factor of $f_{\rm cov} = V_{{\rm IFU}\times\Delta\lambda}/V_{\rm vox}$. Ideally, we could account for this effect through binomial sampling with a probability of $p = f_{\rm cov}$. However, the redshift distribution of the detected mock galaxies differs from the actual LAE distribution similarly to $\bar{N}$ (see Section 2.4). We therefore sample from the detected mock LAEs using a binomial distribution with a probability of $p = f_{\rm cov} f_N(z)$. We use the redshift of the voxel for z.

We imitate the sky subtraction process by subtracting the mean mock intensity spectrum of each observation from the mock intensity. We calculate the mean mock "sky" spectrum for the mock "total," "detected," and "undetected" intensity maps separately, and subtract it from the same mock intensity map. We refer to the mock intensity spectra as $I^{\rm m}(\lambda)$ and the actual data intensity spectra as $I^{\rm d}(\lambda)$. We add the mock intensity to the data intensity ($I^{\rm m+d}_{\rm undet}(\lambda) = I^{\rm m}_{\rm undet}(\lambda) + I^{\rm d}_{\rm undet}(\lambda)$, etc.), where the data have been processed until, and including, the step described in Section 2.2.5. The intensity noise is thus naturally included in the mock spectra, $I^{\rm m+d}_{\rm undet}(\lambda)$. Then, we perform the steps detailed in Section 2.2.6 (masking outlier pixels and high-variance spectra) and Section 2.2.7 (PCA) on $I^{\rm m}(\lambda)$ and $I^{\rm m+d}(\lambda)$. Finally, we calculate the PCA weight vectors of the $I^{\rm m+d}_{\rm tot}(\lambda)$ spectra and apply the PCA cleaning to the $I^{\rm m}_{\rm undet}(\lambda)$ and $I^{\rm m+d}_{\rm undet}(\lambda)$ spectra.

4. Power Spectrum Measurement

4.1. Power Spectrum Estimator

We calculate the cross-power spectrum of the detected LAE overdensity, $\delta_{\rm g}$, and the Lyα intensity fluctuation, δI, in each box. We use a fast Fourier transform to calculate $\widetilde{\delta I}(\boldsymbol{k})$ and $\tilde{\delta}_{\rm g}(\boldsymbol{k})$ and compute the power spectrum,

$$\hat{P}^{(i)}(\boldsymbol{k}) = V_{\rm box}\,\tilde{\delta}^{*}_{\rm g}(\boldsymbol{k})\,\widetilde{\delta I}(\boldsymbol{k}), \tag{15}$$

where $V_{\rm box}$ is the volume of the box, the asterisk denotes the complex conjugate, and the superscript i identifies each box, as shown in Figure 1.

We calculate the power spectrum monopole, $\hat{P}^{(i)}_0(k)$, by averaging $\hat{P}^{(i)}(\boldsymbol{k})$ over angles,

$$\hat{P}^{(i)}_0(k) = \frac{1}{4\pi}\int d^2\hat{\boldsymbol{k}}\,\hat{P}^{(i)}(\boldsymbol{k}), \tag{16}$$

in 23 linearly spaced k bins from $k_{\rm min} = 0.08\,h{\rm Mpc}^{-1}$ to $k_{\rm max} = 1\,h{\rm Mpc}^{-1}$ with $\Delta k = 0.04\,h\,{\rm Mpc}^{-1}$. We calculate $\hat{P}^{(i)}_0(k)$ for the cross-power spectrum of LAEs with the intensity of undetected sources for $N_{\rm PC} \in \{0, 10, \ldots, 200\}$. We also calculate $\hat{P}^{(i)}_0(k)$ of detected mock galaxies and the mock intensity of undetected sources, $\delta I^{\rm m+d}_{\rm undet}$ and $\delta I^{\rm m}_{\rm undet}$.

We do not use weights for the intensity because we find that the generated inverse-variance weights do not improve statistical significance of the power spectrum and complicate the analysis.

$\bar{N}$ naturally includes duplicate detections in repeated observations. The effective mask, $m(\boldsymbol{x})$, for both the galaxy and intensity maps is therefore equal to 1 in the voxels that contain IFU spectral elements and equal to zero otherwise ($m \in \{0, 1\}$). We adjust the power spectra of the boxes by the fraction of observed voxels, $f_V = \langle m\rangle = \langle m^2\rangle$, and calculate the

---

[24] https://github.com/mlujnie/simple

[25] http://class-code.net

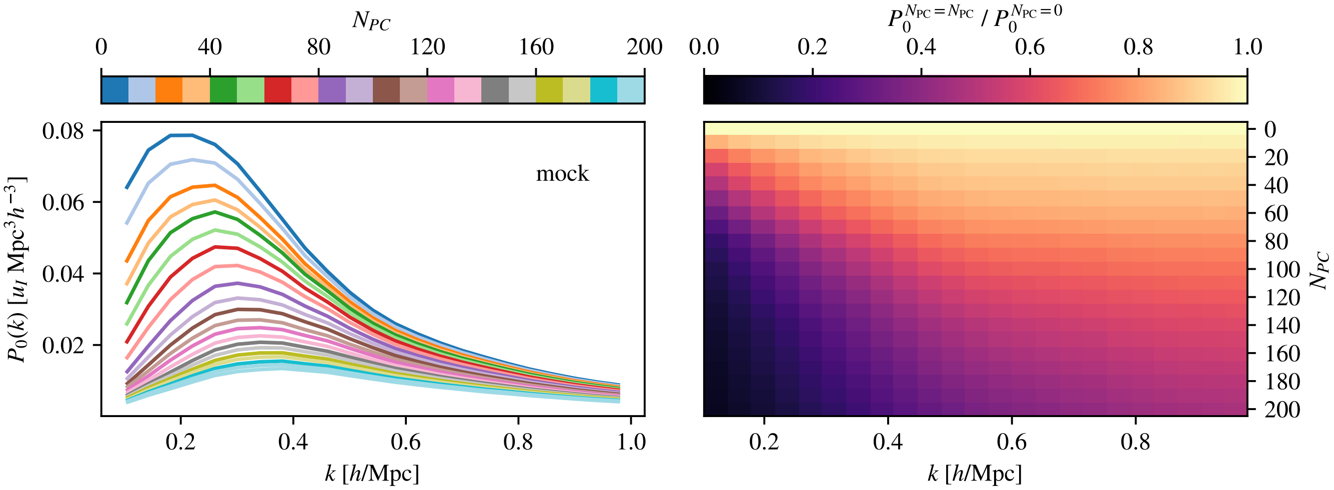

Figure 5. Left panel: suppression of the power spectrum monopole due to PCA cleaning. We show the mean power spectrum monopole obtained from the fiducial mocks in the Spring field in the medium-z bin. The different colors correspond to different values of $N_{\rm PC}$, as indicated in the color bar. Here, $u_I = 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} Å^{-1}. Right panel: the ratio of the power spectrum monopole with and without $N_{\rm PC}$ PCs removed as a function of k and $N_{\rm PC}$.

weighted mean in each redshift bin:

$$\hat{P}_0(k) = \frac{\sum_{i=1}^{N_{\rm boxes}} \hat{P}_0^{(i)}(k) f_{V,i}^{-1} w_i(k)}{\sum_{i=1}^{N_{\rm boxes}} w_i(k)}, \tag{17}$$

where $w_i(k)$ are the weights defined in the next section.

4.2. Covariance Matrix, Weights, and Masking

We calculate the covariance matrix of $\hat{P}_0(k)$ as the sum of two matrices: the first from shuffling the data ($C_{\rm shuffle}$), and the second from the mocks ($C_{\rm mock}$). $C_{\rm shuffle}$ quantifies the statistical covariance arising from noise in the data, while $C_{\rm mock}$ quantifies cosmic variance. The total covariance matrix is given by $C = C_{\rm shuffle} + C_{\rm mock}$.

For $C_{\rm shuffle}$ and the weights, $w_i(k)$, we randomly shuffle intensities of the IFU spectral elements in each of the three fields while keeping the redshift fixed, then calculate the power spectra, $\hat{P}_0^{s,(i)}(k)$. We repeat this process 100 times, and compute weights as the inverse variance of the power spectrum monopole of each box as a function of wavenumber,

$$w_i(k) = \frac{1}{{\rm var}(\hat{P}_0^{s,(i)}(k) f_{V,i}^{-1})}. \tag{18}$$

For each shuffling instance, we calculate the weighted mean, $\hat{P}_0^{\rm s}(k)$, of the shuffled box power spectra, $\hat{P}_0^{s,(i)}(k)$. Finally, we construct the covariance matrix of the weighted mean power spectrum of each field in each redshift bin.

For $C_{\rm mock}$, we calculate the covariance matrix using the mock-only intensity, $I_{\rm undet}^{\rm m}(z)$, of 100 mocks per box. This covariance quantifies cosmic variance of the intensity and the LAE catalog.

The power spectra of some boxes exhibit a much higher variance along the wavenumbers than the rest, indicating the presence of residual systematic errors. To remain cautious, we exclude these boxes from the analysis. We also mask one box in the Spring field due to its small number of observations. The masked boxes are marked by the hatched regions in Figure 1.

4.3. Loss of Power from PCA

Figure 5 displays the effect of removing PCs on the cross-power spectrum of galaxies and intensity from the mocks. The left panel shows the cross-power spectrum of mock galaxies with $I_{\rm undet}^{\rm m}(z)$ after sky subtraction and removal of $N_{\rm PC} \in \{0, 10, \ldots, 200\}$ PCs in the Spring field in the medium-z bin. The suppression of power at $k \lesssim 0.4\,h\,{\rm Mpc}^{-1}$ is caused by sky subtraction. The right panel shows the ratio of the sky-subtracted mock power spectrum monopole with and without PCA.

Removing PCs suppresses power, but does not completely eliminate it. Mocks are useful for capturing the effect of PCA on the signal power spectrum. We will use this information when we compare theoretical predictions with observed power spectra.

5. Fitting the Power Spectrum Model

5.1. Model

We use mock power spectra to calculate theoretical predictions and compare them with observed power spectra. The mock power spectrum can be written as

$$\begin{aligned} P_{\rm mock}(\boldsymbol{k}) = {} & b_{\rm g} b_I \langle I \rangle \mathcal{T}(\boldsymbol{k}) \\ & \times \int \frac{d^3\boldsymbol{k}'}{(2\pi)^3} P_{\rm m}(\boldsymbol{k}') F_{\rm RSD}(\boldsymbol{k}') \tilde{D}(\boldsymbol{k}') \\ & \times \widetilde{W}_I^*(\boldsymbol{k} - \boldsymbol{k}') \widetilde{W}_g(\boldsymbol{k} - \boldsymbol{k}'), \end{aligned} \tag{19}$$

where $b_{\rm g}$ and b_I are the linear bias parameters of LAEs and intensity, respectively, $\langle I \rangle$ is the mean intensity of the mock, and $P_{\rm m}(\boldsymbol{k})$ is the matter power spectrum computed for the fiducial cosmological model.

The factor $F_{\rm RSD}(\boldsymbol{k})$ accounts for the effects of nonlinear RSD and a potential Lyα RT (M. Lujan Niemeyer 2025). In the mock, we account for nonlinearity in the Jacobian of the coordinate transformation from real to redshift space non-perturbatively (A. Agrawal et al. 2017). In a linear model, it would be $F_{\rm RSD}^{\rm lin} = (1 + f\mu^2 b_{\rm g}^{-1})(1 + f\mu^2 b_I^{-1})$ (N. Kaiser 1987), where $\mu = \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}_\parallel$ is the cosine of the angle between the unit wavevector $\hat{\boldsymbol{k}}$ and the LOS $\hat{k}_\parallel$.

In Appendix B we show a linear model for Lyα RT in $F_{\rm RSD}$. We do not, however, account for this effect in this paper because we do not detect the quadrupole power spectrum, and the RT effect is degenerate with the bias parameters.

The factor $\tilde{D}(\boldsymbol{k})$ accounts for the LOS damping due to limited spectral resolution (see Section 3.1). The transfer function, $\mathcal{T}(\boldsymbol{k})$, accounts for the loss of power due to data processing, especially sky subtraction and PCA (see the right panel of Figure 5 for the loss of power due to PCA). Note that we include the transfer function in our forward model and do not attempt to reconstruct a PCA-free power spectrum from the data. The window functions, $\widetilde{W}_I^*(\boldsymbol{k})$ and $\widetilde{W}_g(\boldsymbol{k})$, account for the observing mask and weights. Since the mask and weights are the same for the mocks and the data, we assume that the window functions are fully modeled by the mocks.

In our model for the monopole power spectrum, the only free parameter is the overall factor, $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}$, where $\bar{F}_{\rm RSD}$ is the effective nonlinear RSD factor from the integral given in Equation (19) and the angular averaging given in Equation (16).

In the absence of a mask, a linear model gives $\bar{F}_{\rm RSD}^{\rm lin} = 1 + \frac{f}{3}(b_g^{-1} + b_I^{-1}) + \frac{f^2}{5} b_g^{-1} b_I^{-1} \simeq 1.3$. Therefore, a slight mismatch in the values of b_g and b_I between the mock and the data only gives a minor correction to this factor.

In practice, we define the fitting parameter, $A = b_g b_I \langle I \rangle \bar{F}_{\rm RSD} / (b_g^{\rm fid} b_I^{\rm fid} \langle I^{\rm fid} \rangle \bar{F}_{\rm RSD}^{\rm fid})$, where $b_g^{\rm fid}$, $b_I^{\rm fid}$, and $\langle I^{\rm fid} \rangle$ are the fiducial LAE bias, the intensity bias, and the mean intensity of undetected sources assumed in the mocks. $\bar{F}_{\rm RSD}^{\rm fid}$ is the effective nonlinear RSD factor in the mocks. If the mocks were in perfect agreement with the data, we would find that $A = 1$.

Due to nonlinearity, $\bar{F}_{\rm RSD}$ depends on the wavenumber k. However, the ratio, $\bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$, depends on k only weakly if the mock and the data are in reasonable agreement, which is the case (see Section 6). Therefore, we neglect the scale dependence of $\bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$ and treat it as a constant factor. This is a good approximation given our limited ability in determining the shape of the power spectrum due to the large statistical uncertainty.

5.2. Parameter Inference

To determine A, we fit the fiducial mock power spectrum multiplied by an amplitude A to the data. We define $\chi^2 = \boldsymbol{Y}^\top \boldsymbol{C}^{-1} \boldsymbol{Y}$, where $\boldsymbol{Y} = \boldsymbol{Y}_{\rm data} - A \boldsymbol{Y}_{\rm fid}$ is the difference between the measured power spectrum and the model power spectrum,

$$\begin{aligned} \boldsymbol{Y}_{\rm data}^\top &= (P_0^{\rm data}(k_1), P_0^{\rm data}(k_2), \ldots, P_0^{\rm data}(k_{23})), \\ \boldsymbol{Y}_{\rm fid}^\top &= (P_0^{\rm fid}(k_1), P_0^{\rm fid}(k_2), \ldots, P_0^{\rm fid}(k_{23})), \end{aligned} \tag{20}$$

and $P_0^{\rm fid}(k)$ is the monopole of the mock power spectrum $P_{\rm mock}(\boldsymbol{k})$ (Equation (19)), calculated in the same 23 k bins as the data.

We also combine the power spectra of the fields and fit a single amplitude A to all of the fields within each redshift bin simultaneously. Specifically, we use

$$\begin{aligned} \boldsymbol{Y}_{\rm data,all}^\top = (&P_0^{\rm data,Fall}(k_1), \ldots, P_0^{\rm data,Fall}(k_{23}), \\ &P_0^{\rm data,Spring}(k_1), \ldots, P_0^{\rm data,Spring}(k_{23}), \\ &P_0^{\rm data,NEP}(k_1), \ldots, P_0^{\rm data,NEP}(k_{23})) \end{aligned} \tag{21}$$

and the equivalent vector for the model. For the covariance matrix, we combine the individual covariance matrices:

$$C_{\rm all} = \begin{pmatrix} C_{\rm Fall} & 0 & 0 \\ 0 & C_{\rm Spring} & 0 \\ 0 & 0 & C_{\rm NEP} \end{pmatrix}. \tag{22}$$

The best-fit amplitude that minimizes χ^2 for the individual fields is given by

$$A_{\rm best} = \frac{\boldsymbol{Y}_{\rm fid}^\top C^{-1} \boldsymbol{Y}_{\rm data}}{\boldsymbol{Y}_{\rm fid}^\top C^{-1} \boldsymbol{Y}_{\rm fid}}. \tag{23}$$

The best-fit amplitude that minimizes χ^2 for the combined fit is given by

$$A_{\rm best}^{\rm all} = \frac{\boldsymbol{Y}_{\rm fid,all}^\top C_{\rm all}^{-1} \boldsymbol{Y}_{\rm data,all}}{\boldsymbol{Y}_{\rm fid,all}^\top C_{\rm all}^{-1} \boldsymbol{Y}_{\rm fid,all}}. \tag{24}$$

5.3. Error Estimation and Goodness-of-fit

The uncertainty of the best-fit amplitude can be estimated as

$$\sigma_{\rm ls} = \frac{1}{\sqrt{\boldsymbol{Y}_{\rm fid}^\top C^{-1} \boldsymbol{Y}_{\rm fid}}}. \tag{25}$$

We calculate $A_{\rm best}$ and $\sigma_{\rm ls}$ for each field, as well as $A_{\rm best}^{\rm all}$ and $\sigma_{\rm ls}^{\rm all}$, for each $N_{\rm PC}$ and redshift bin.

To ensure that we do not underestimate the uncertainty of the measured power spectra, we empirically estimate the uncertainty, referred to as $\sigma_{\rm emp}$. This method assumes that the best-fit model is correct, but that the data contain an unaccounted-for systematic error that increases the point-to-point variance. We generate fake power spectra by randomly sampling from a Gaussian probability density function (PDF) at each wavenumber bin. The mean of the Gaussian is given by the power spectrum of the best-fit model, $\boldsymbol{Y}_{\rm model} = A_{\rm best} \boldsymbol{Y}_{\rm fid}$; the standard deviation of the point in the jth k bin is given by $\sqrt{N_{\rm rms}^{-1} \sum_{i=j-4}^{j+4} (Y_{{\rm data},i} - Y_{{\rm model},i})^2}$, where we sum over the $N_{\rm rms} = 9$ nearest wavenumber bins. We fit the model to the data using Equations (23) and (24) and repeat this process 1000 times. The standard deviation $\sigma_{\rm emp}$ of the 1000 best-fit amplitudes is typically larger than $\sigma_{\rm ls}$ with the mean ratio $\langle \sigma_{\rm emp}/\sigma_{\rm ls} \rangle \simeq 1.6$. To remain conservative, we select the larger of the two estimated uncertainties,

$$\sigma_A = \max\{\sigma_{\rm emp}, \sigma_{\rm ls}\}. \tag{26}$$

To determine the goodness-of-fit, we calculate the reduced χ^2, $\chi_\nu^2 = \chi_{\rm min}^2/\nu$, where $\chi_{\rm min}^2$ is the χ^2 minimum, and $\nu = N_k - N_{\rm fit}$ is the number of degrees of freedom, assuming that the covariance matrix is diagonal (see Appendix C for the correlation matrices). Each individual field has $N_k = 23$ wavenumber bins per power spectrum. Since we are fitting a single parameter, $N_{\rm fit} = 1$, there are $\nu = 22$ degrees of freedom for the fit to individual fields. The combined fit has $\nu = 3 \times 23-1 = 68$ degrees of freedom per redshift bin.

We also calculate the probability-to-exceed (PTE, also called p-value), from the integral of a χ^2 PDF, $\mathcal{P}_\nu(\chi^2)$, from $\chi_{\rm min}^2$ of the best fit to infinity:

$${\rm PTE}(\chi_{\rm min}^2) = \int_{\chi_{\rm min}^2}^{\infty} \mathcal{P}_\nu(\hat{\chi}^2) d\hat{\chi}^2. \tag{27}$$

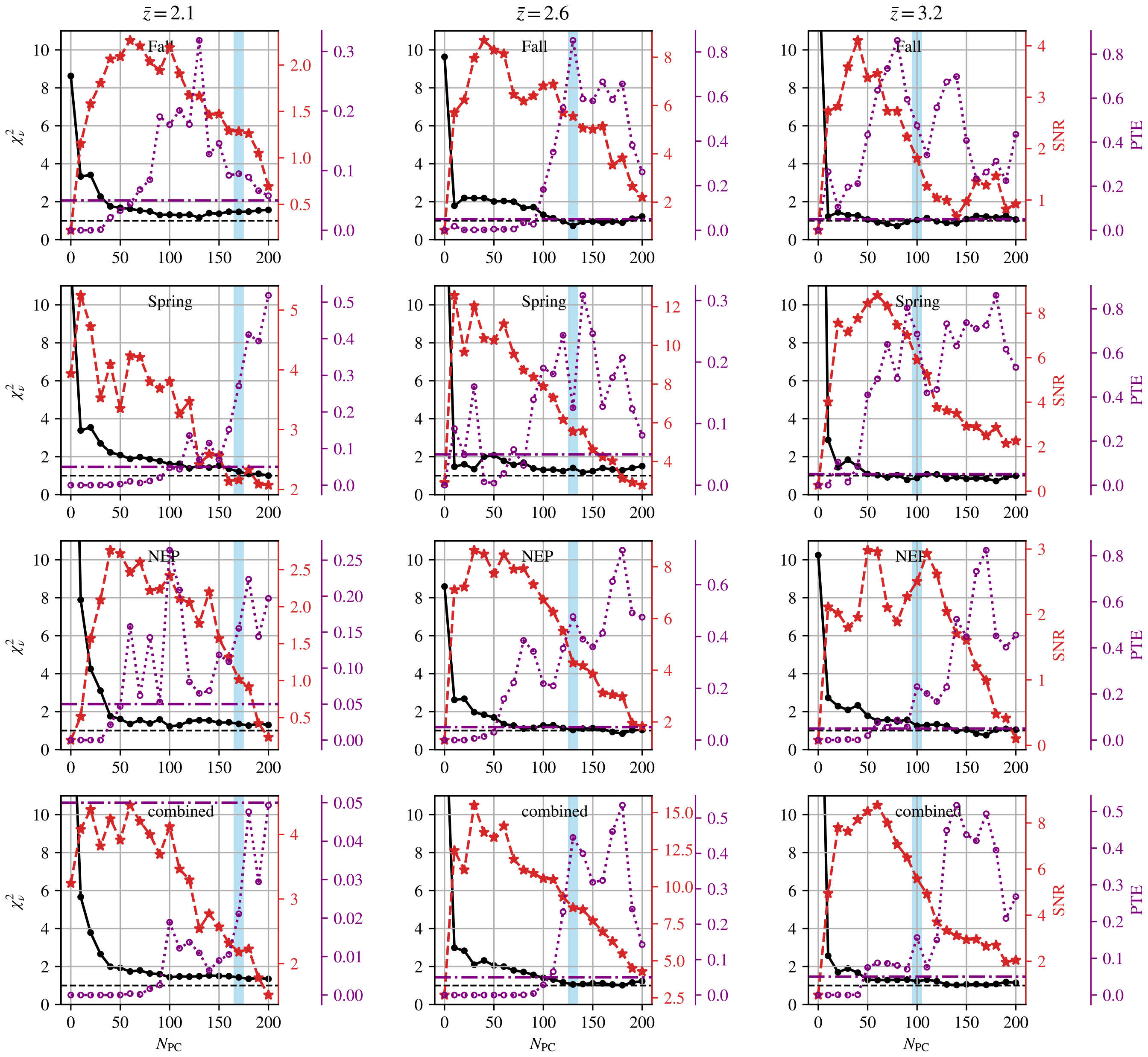

Figure 6. Reduced χ^2_ν (black dots), SNR (red stars), and PTE (purple empty circles) of the fit as a function of $N_{\rm PC}$ in the Fall (first row), Spring (second row), and NEP (third row) fields, and that combining the three fields (fourth row). The columns show the three redshift bins. The horizontal black dashed lines indicate $\chi^2_\nu = 1$. The horizontal purple dotted–dashed lines indicate PTE = 0.05.

If a fit yields PTE > 0.05, the probability of obtaining a higher value of $\chi^2_{\rm min}$ than this fit is >5% (assuming the model is correct), indicating consistency with the model.

6. Results

Choosing the optimal $N_{\rm PC}$ is a crucial decision in the analysis. We want $N_{\rm PC}$ to be high enough that most systematics are removed. We use goodness-of-fit as a guide: good agreement between the model and the data implies a low value of $\chi^2_\nu \simeq 1$ and a high value of PTE > 0.05. However, removal of PCs also reduces the signal. We therefore want $N_{\rm PC}$ to be small enough to retain a high SNR. Finally, we want the amplitudes inferred in the three fields to be consistent with each other at the chosen $N_{\rm PC}$.

Figure 6 shows χ^2_ν, the PTE, and the $SNR(A) = A_{\rm best}/\sigma_A$ as a function of $N_{\rm PC}$ for the three fields and the combined fit, for the three redshift bins. χ^2_ν decreases with increasing $N_{\rm PC}$, indicating a better agreement between the measured and model power spectra. In the low-z bin of the combined fit, χ^2_ν approaches $\simeq$1.5 at $N_{\rm PC} \geqslant 100$; in the medium- and high-z bins of the combined fit, χ^2_ν is close to 1 at $N_{\rm PC} \geqslant 120$. The PTE generally also improves (increases) with increasing $N_{\rm PC}$. The SNR, however, decreases with $N_{\rm PC}$ above a threshold value of $N_{\rm PC} \simeq 50$.

The goodness-of-fit is acceptable in all cases except for the fit of the combined power spectrum in the low-z bin, which yields a low PTE < 0.05 at all $N_{\rm PC}$. This result could indicate the presence of systematic errors that are not quantified by the covariance matrix or an underestimation of the statistical

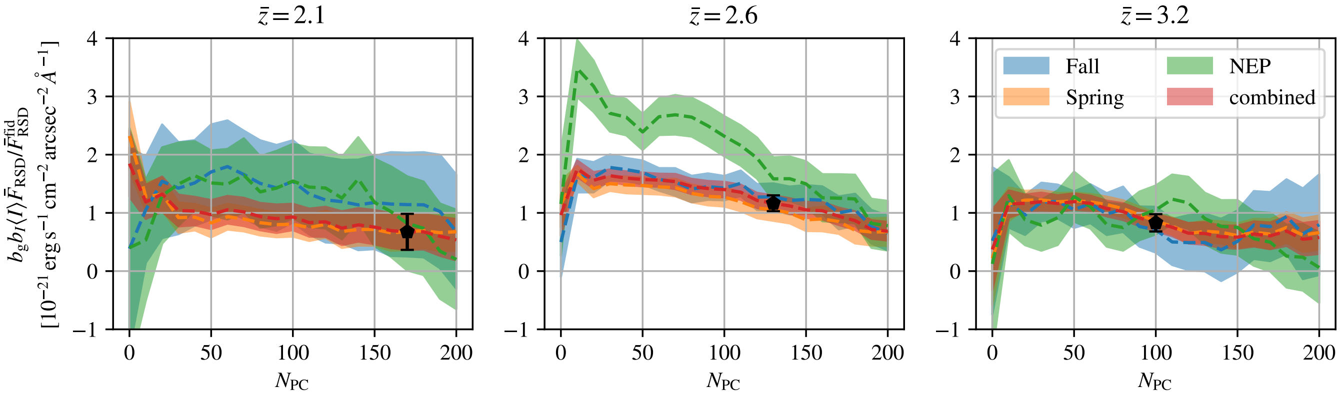

Figure 7. Constraints on $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}^{\rm fid}_{\rm RSD}$ inferred from the best-fit amplitude as a function of $N_{\rm PC}$ in the Fall (blue), Spring (orange), and NEP (green) fields, along with the combination of all three fields (red). The dashed lines show the mean, while the shaded areas show the 1σ uncertainties. The black polygons indicate the constraints from the combined fit, as determined from the best combination of χ^2_ν, SNR, and PTE values.

covariance matrix of the power spectrum. Otherwise, the fit is excellent.

We also want to choose $N_{\rm PC}$ so that the inferred amplitudes in the three fields agree. Figure 7 displays the best-fit $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}^{\rm fid}_{\rm RSD} = A_{\rm best} b_g^{\rm fid} b_I^{\rm fid} \langle I^{\rm fid} \rangle$ and its uncertainty for the three fields and the combined fit as a function of $N_{\rm PC}$. The fiducial bias values used for the mocks are $b_I^{\rm fid} = b_g^{\rm fid} = 2$ (Section 3.1). The fiducial mean intensity of undetected sources obtained from the mocks is $\langle I^{\rm fid} \rangle = 1.5$, 2.2, and $1.4 \times 10^{-22}\,{\rm erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}\,Å^{-1}}$ in the redshift bins centered around $\bar{z} = 2.1$, 2.6, and 3.2, respectively.

In principle, we could choose different values of $N_{\rm PC}$ for each field and redshift bin because the PCA is performed independently in each field. However, for simplicity, we choose the same $N_{\rm PC}$ for all fields in each redshift bin. Based on the small χ^2_ν and high SNR shown in Figure 6, and the requirement that the best-fit amplitudes be consistent across fields, we choose $N_{\rm PC} = 170$, 130, and 100 in the low-, medium-, and high-z bins, respectively. The best-fit values of $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}^{\rm fid}_{\rm RSD}$ at these $N_{\rm PC}$ are $b_g b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}^{\rm fid}_{\rm RSD}$=(6.7 ± 3.1), (11.7 ± 1.4), and (8.3 ± 1.5) $\times 10^{-22}\,{\rm erg\,s^{-1}\,cm^{-2}\,arcsec^{-2}\,Å^{-1}}$ in the low-, medium-, and high-z bins, respectively. Our main result is that we significantly detect the LAE-Lyα intensity cross-power spectrum.

Finally, Figure 8 shows the quality of the fits to the power spectrum monopoles from each field and redshift bin using the selected $N_{\rm PC}$. We compare the measured power spectra to the fiducial mock power spectrum and to the mock power spectrum multiplied by the best-fit amplitude. The error bars of the power spectra are given by the square root of the diagonal elements of the covariance matrix (see Section 4.2). The mocks accurately describe the data.

Figure 9 shows the weighted mean power spectrum monopole of the three fields and the three redshift bins shown in Figure 8. We use inverse-variance weights,

$$\bar{P}_0(k_i) = \frac{\sum_{j=1}^{9} \hat{P}_0^{(j)}(k)\,\sigma^{-2}_{\hat{P}_0^{(j)}}(k_i)}{\sum_{j=1}^{9} \sigma^{-2}_{\hat{P}_0^{(j)}}(k_i)}, \tag{28}$$

where $\sigma_{\hat{P}_0^{(j)}}(k_i)$ is the square root of the diagonal element i of the covariance matrix of the power spectrum monopole $\hat{P}_0^{(j)}$, and the index j denotes a field and redshift bin. The propagated uncertainty of the weighted mean power spectrum monopole, $\bar{P}_0(k)$, is given by

$$\sigma_{\bar{P}_0}(k_i) = \frac{1}{\sqrt{\sum_{j=1}^{9} \sigma^{-2}_{\hat{P}_0^{(j)}}(k_i)}}. \tag{29}$$

We calculate the weighted mean of the best-fit mock power spectra using the same inverse weights of the data. We use the same $N_{\rm PC}$ as in Figure 8 for the data and mocks. As a test, we fit an amplitude times the weighted mean of the best-fit mock power spectrum monopole to the weighted mean power spectrum monopole. We find a best-fit amplitude of 0.97 ± 0.07, which is consistent with the expectation of 1. With $\nu = 22$ degrees of freedom, we find $\chi^2 = 21$, $\chi^2_\nu = 0.94$, and PTE = 0.54.

Although all quadrupole power spectra are consistent with zero, we do not use them to extract physical constraints, such as the Lyα RT effect (M. Lujan Niemeyer 2025), as the shape and amplitude of the quadrupole power spectra are strongly affected by sky subtraction and PCA, and are therefore not yet stable. We leave detailed investigations of the quadrupole power spectra for future study.

7. Comparison to Previous Constraints

To compare our results to previous work, we convert our measurements to $b_I \langle I \rangle$ assuming a fiducial LAE bias of $b_g = 2$ (Section 3.1) and $\bar{F}_{\rm RSD} = \bar{F}^{\rm fid}_{\rm RSD}$.

Figure 10 compares our combined best-fit values of $b_I \langle I \rangle$, i.e., the product of the linear intensity bias and the intensity of only undetected sources, with those obtained from cross-correlations between QSOs and Lyα intensity (R. A. C. Croft et al. 2016, 2018; X. Lin et al. 2022), and the 95% confidence upper limit on $b_I \langle I \rangle$ inferred from Lyα forest–Lyα intensity cross-correlation (R. A. C. Croft et al. 2018). These studies are all performed at redshifts from $z = 2$ to 3.5.

Since these analyses constrain the product of the intensity bias and total Lyα intensity, we also show $b_I^{\rm fid}(A_{\rm best} \langle I^{\rm fid}_{\rm undet} \rangle + \langle I^{\rm fid}_{\rm det} \rangle)$, where $A_{\rm best}$ is the best-fit amplitude, $b_I^{\rm fid} = 2$ is the fiducial intensity bias, and $\langle I^{\rm fid}_{\rm undet} \rangle$ and $\langle I^{\rm fid}_{\rm det} \rangle$ are the mean intensities of only undetected and only detected sources from the fiducial mocks, respectively. The luminosity function of the mocks does not include QSOs (see Section 3). These values are only slightly larger than $b_I \langle I_{\rm undet} \rangle$.

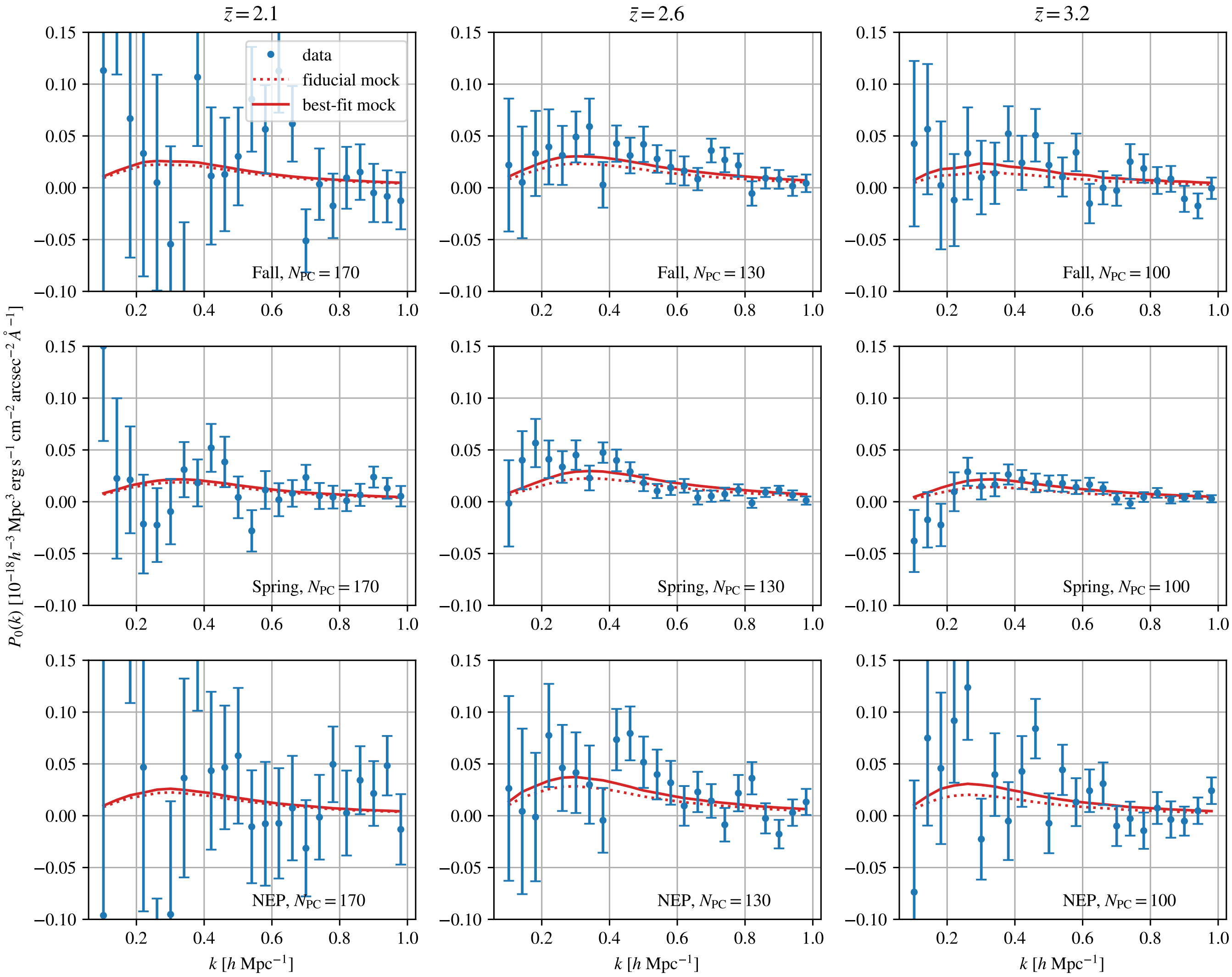

Figure 8. Power spectrum monopoles in the Fall (top row), Spring (middle row), and NEP (bottom row) fields and in the three redshift bins (different columns). Each panel shows the mean power spectrum monopole (Equation (17)) in each field and redshift bin. We have removed $N_{\rm PC} = 170$, 130, and 100 PCs in the low-, medium-, and high-z bins, respectively. The dotted red lines indicate the fiducial mock power spectra using the same $N_{\rm PC}$. The solid red lines are the mock power spectra multiplied by the best-fit amplitude.

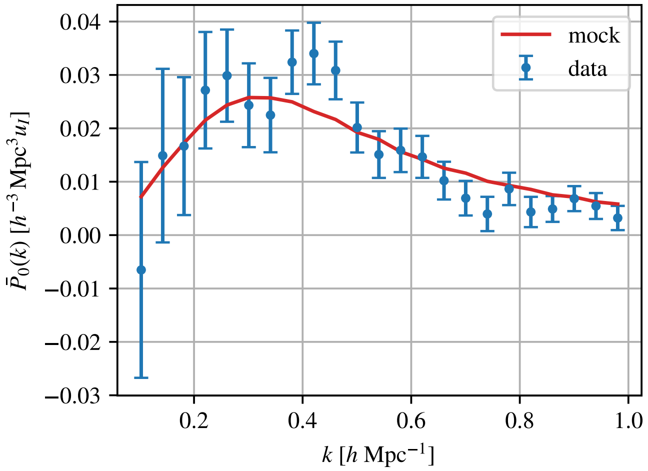

Figure 9. Weighted mean power spectrum monopole of the three fields and three redshift bins of the data (blue) and the best-fit mocks (red). Here, $u_I = 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} Å^{-1}.

Both $b_I \langle I_{\rm undet} \rangle$ and $b_I^{\rm fid} (A_{\rm best} \langle I_{\rm undet}^{\rm fid} \rangle + \langle I_{\rm det}^{\rm fid} \rangle)$ are significantly smaller than the values inferred from QSO-Lyα intensity correlations. The low- and high-z best-fit values are consistent with the upper limit inferred from the Lyα forest–Lyα intensity cross-correlation, while the medium-z bin is 2.6σ higher.

This discrepancy could be due to the QSO-Lyα cross-correlations being dominated by small-scale physics that are affected by QSOs, such as an elevated radiation field surrounding them (e.g., J. S. A. Miller et al. 2021; C. Dong et al. 2023). Since our galaxy sample is dominated by star-forming galaxies, our measurement is less affected by QSO-related issues and may provide a more accurate representation of the Lyα radiation field in the IGM.

8. Origins of the Lyα Emission

In this work, we masked the spectra within $10''$ of the detected LAEs, which, in our redshift range, corresponds to a proper distance of $\simeq$75–85 kpc. Beyond this distance, the

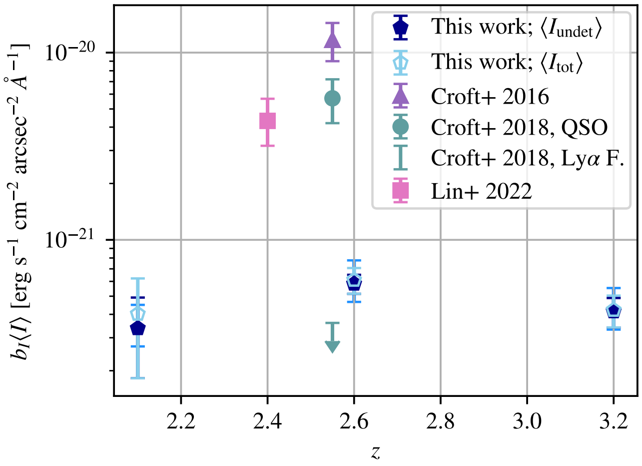

Figure 10. Constraints on $b_I\langle I\rangle$ obtained from this work (dark-blue polygons). We assume $b_{\rm g}^{\rm fid}=2$ and $\bar{F}_{\rm RSD}=\bar{F}_{\rm RSD}^{\rm fid}$. The medium blue error bars span the constraints when assuming $b_{\rm g}^{\rm fid}=2\pm0.5$. The light-blue polygons show $b_I\langle I_{\rm tot}\rangle\simeq b_I^{\rm fid}(A_{\rm best}\langle I_{\rm undet}^{\rm fid}\rangle+\langle I_{\rm det}^{\rm fid}\rangle)$ (see the text for more information). The purple triangle, olive circle, and pink square are the constraints from the QSO-Lyα intensity cross-correlations of R. A. C. Croft et al. (2016, 2018) and X. Lin et al. (2022), respectively. The upper limit from the Lyα forest–Lyα intensity cross-correlation (R. A. C. Croft et al. 2018) is displayed as an olive arrow.

photons contributing to the radial profiles of the simulated LAEs originate from outside the stacked LAEs' dark-matter halos (C. Byrohl et al. 2021; C. Byrohl & D. Nelson 2023). Therefore, the intensity should be dominated by photons unrelated to the detected LAEs. The intensity includes photons emitted within undetected galaxies, from detected LAEs that scatter many tens of kiloparsecs from the origin, and from the CGM and IGM. While we cannot directly constrain the origin and emission mechanism from our measurement, we can compare it to theoretical expectations.

We translate our constraint on $b_{\rm g}b_I\langle I\rangle\bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$ to a constraint on the comoving luminosity density

$$\rho_L=\frac{4\pi}{c}\langle I\rangle H(z)(1+z)^2\lambda_0, \tag{30}$$

where $\lambda_0=1215.67$ Å is the Lyα rest-frame wavelength. We assume $b_{\rm g}^{\rm fid}=b_I^{\rm fid}=2$, as we do not know the bias of the detected LAEs or of the intensity. We will discuss a possible range of the intensity bias from simulation work in Section 8.3. In addition, we assume that $\bar{F}_{\rm RSD}=\bar{F}_{\rm RSD}^{\rm fid}$. We compare our constraint on ρ_L to theoretical expectations in Figure 11.

8.1. Expectations from Luminosity Functions

The mock is based on our fiducial model, which only includes the intensity from faint, undetected galaxies, as expected from extrapolating the LAE luminosity function measured by H. Umeda et al. (2025). The mean comoving luminosity density from undetected galaxies can be calculated by integrating the luminosity function up to the HETDEX detection limit:

$$\langle\rho_L^{\rm undet}\rangle=\int_{L_{\rm min}}^{L_{\rm det}}dL\frac{dn}{dL}L, \tag{31}$$

where $L_{\rm min}$ is given by the lowest Lyα luminosity of galaxies, and $L_{\rm det}$ is the minimum luminosity detectable by HETDEX. The corresponding mean specific intensity at redshift z is then given by

$$\langle I_{\rm undet}^{\rm gal}\rangle=\frac{c\langle\rho_L^{\rm undet}\rangle}{4\pi H(z)(1+z)^2\lambda_0}. \tag{32}$$

We integrate Equation (31) from $L_{\rm min}=4\times10^{40}$ erg s^{-1} to a typical HETDEX detection limit, $L_{\rm max}=4\times10^{42}$ erg s^{-1} (see Figure 2). To illustrate the uncertainty due to the unknown minimum luminosity, we also show the result for $L_{\rm min}=10^{36}$ erg s^{-1}. Using a smaller value for $L_{\rm min}$ barely changes the integral. Figure 11 shows $\langle\rho_L^{\rm undet}\rangle$ obtained by integrating various luminosity functions from the literature (M. Ouchi et al. 2008; A. Konno et al. 2016; D. Tornotti et al. 2025; H. Umeda et al. 2025). The luminosity densities of the different luminosity functions are similar to each other and slightly lower than that of our measurement. Note that the luminosity functions of M. Ouchi et al. (2008), A. Konno et al. (2016), and H. Umeda et al. (2025) are obtained from narrowband surveys, which have different LAE selection criteria from spectroscopic surveys like HETDEX or D. Tornotti et al. (2025). The value of ρ_L from integrating luminosity functions is also strongly dependent on the faint-end slope α of the luminosity function, which is degenerate with $L_\star$ and $\phi_\star$ (see Figure 3 of D. Tornotti et al. 2025).

8.2. Star Formation

The comoving star formation rate density at $z=2$–3.5 is $\simeq0.1\,M_\odot$ yr^{-1} Mpc^{-3} (M. Rowan-Robinson et al. 2016). We use the conversion of the star formation rate, $\dot{M}_\star$, to the intrinsic Lyα luminosity (M. Dijkstra 2019)

$$L_{{\rm Ly}\alpha}\simeq10^{42}\ {\rm erg\ s^{-1}}\left(\frac{\dot{M}_\star}{M_\odot\ {\rm yr}^{-1}}\right). \tag{33}$$

Note that this conversion factor strongly depends on the initial stellar mass function and the assumption of case-B recombination and can deviate by factors of 2–3 with other assumptions (A. Raiter et al. 2010).

Using Equation (33), we obtain a comoving intrinsic luminosity density of $\rho_L\simeq10^{41}$ erg s^{-1} Mpc^{-3}. The corresponding observed intensity at redshift $z=2.1$ (2.6/3.2) is $\langle I_{\rm SF}\rangle\simeq f_{\rm esc}\times2.4\,(1.4/0.8)\times10^{-21}$ erg s^{-1} cm^{-2} arcsec^{-2} Å^{-1}, where $f_{\rm esc}$ is the mean escape fraction of Lyα photons from star-forming regions.

Therefore, star formation can power the mean intensity inferred from our measurements if $b_I\langle I\rangle f_{\rm esc}/\langle I_{SF}\rangle=b_If_{\rm esc}\simeq0.1$, 0.4, and 0.5 in the low-, medium-, and high-z bins, respectively, assuming $b_{\rm g}=2$ and $\bar{F}_{\rm RSD}=\bar{F}_{\rm RSD}^{\rm fid}$. These numbers are reasonable: much higher ($b_If_{\rm esc}\gg1$) and lower ($b_If_{\rm esc}\ll0.1$) values would not support the interpretation of Lyα emission powered by star formation.

8.3. Comparison to an RT Simulation

C. Byrohl & D. Nelson (2023) post-processed a cosmological hydrodynamic simulation in a cosmological volume at $z=2$ with Lyα RT. The authors calibrated the escape fraction of Lyα emission from galaxies to an observed luminosity function, painted Lyα emission from recombination and collisional excitation into cells without star formation, and simulated the scattering process of the photons. In the simulation, 48% of Lyα photons are produced by collisional

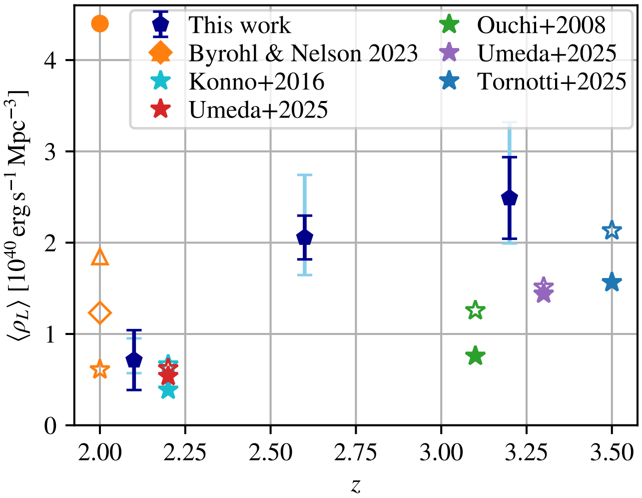

Figure 11. Constraints on the mean comoving luminosity density of undetected sources from this work, $\langle\rho_L\rangle$, assuming $b_{\rm g}^{\rm fid} = b_I^{\rm fid} = 2$ and $\bar{F}_{\rm RSD} = \bar{F}_{\rm RSD}^{\rm fid}$ (dark-blue polygons). The dark-blue error bars are the statistical uncertainty of the fit, whereas the light-blue error bars indicate the best-fit values with $b_I^{\rm fid} = 2.0 \pm 0.5$ and $b_{\rm g}^{\rm fid} = 2$. The filled stars show the luminosity density obtained by integrating the luminosity functions (M. Ouchi et al. 2008; A. Konno et al. 2016; H. Umeda et al. 2025; D. Tornotti et al. 2025) from $L_{\rm min} = 4 \times 10^{40}$ erg s^{-1} to a typical HETDEX detection limit of $L_{\rm max} = 4 \times 10^{42}$ erg s^{-1}. The empty stars of the same colors show the luminosity density obtained with $L_{\rm min} = 10^{36}$ erg s^{-1}, below which the integral barely changes. The orange symbols are predictions from cosmological simulations with Lyα RT (C. Byrohl & D. Nelson 2023): the filled circle shows the total mean luminosity density; the empty orange star represents photons from galaxies with luminosity $L < 10^{41.75}$ erg s^{-1}; and the empty diamond and triangle denote photons that last scattered in the IGM and CGM, respectively.

excitation, making it the dominant emission mechanism. Star formation produces 28%, and recombination outside the interstellar medium (ISM) contributes 23% of the photons. The vast majority, 98%, of Lyα photons are produced within halos, with more than half produced within the CGM. Only 2% are produced in the IGM.

Figure 11 compares our constraints on the luminosity density with the total luminosity density of the simulation of C. Byrohl & D. Nelson (2023) and its contributions from photons that last scattered in faint galaxies, in the CGM, and in the IGM.

The total mean luminosity density of the C. Byrohl & D. Nelson (2023) simulation is 4.4×10^{40} erg s^{-1} Mpc^{-3} (filled circle). After scattering, 30% of the luminosity density originates from galaxies within an $1\rlap{.}''5$ aperture, 42% of photons last scatter in the CGM (empty triangle), and 28% of Lyα photons reach the observer from the IGM (empty diamond).

Of the photons that reach us from galaxies, 6.1×10^{39} erg s^{-1} Mpc^{-3} (empty star) are due to galaxies with a Lyα luminosity of $L < 10^{41.75}$ erg s^{-1}, which is fainter than the typical HETDEX detection limit. Given this luminosity density, the photons that reach us from faint galaxies, the CGM, and the IGM would produce an intensity of $9\,(5/3) \times 10^{-22}$ erg s^{-1} cm^{-2} arcsec^{-2} Å^{-1} at redshift $z = 2.1$ (2.6/3.2). Assuming that this value is the mean intensity of undetected sources, the intensity bias necessary for our measurement is $b_I \simeq 0.4$ (1.1/1.3). Note, however, that the approximation that $\bar{F}_{\rm RSD} = \bar{F}_{\rm RSD}^{\rm fid}$ breaks for values of b_I that differ significantly from $b_I^{\rm fid} = 2$.

We masked spectra containing continuum emission in our measurement, which may include galaxies with nonzero Lyα emission. When we include only emission that last scattered in the CGM or IGM, the necessary intensity bias to match the simulation results is $b_I \simeq 0.5$, 1.3, and 1.6 at $z = 2.1$, 2.6, and 3.2, respectively.

We can approximate the intensity bias b_I from the relation between the luminosity density as a function of dark-matter halo mass. As shown in Figure 3 in C. Byrohl & D. Nelson (2023), their simulation predicts a tight relationship between the Lyα luminosity and halo mass. From this relation, the intensity bias b_I can be estimated as

$$b_I = \frac{\int_{M_{\rm h,min}}^{M_{\rm h,max}} dM_{\rm h}\; L dn/dM_{\rm h}\; b(M_{\rm h})}{\int_{M_{\rm h,min}}^{M_{\rm h,max}} dM_{\rm h}\; L dn/dM_{\rm h}}, \tag{34}$$

where $M_{\rm h}$ is the halo mass, and we adopt a fitting formula from J. L. Tinker et al. (2010) to compute the halo bias, $b(M_{\rm h})$. We obtain $b_I \simeq 1.4$ for $(\log_{10}(M_{\rm h,min}/M_\odot),\, \log_{10}(M_{\rm h,max}/M_\odot)) = (9,\, 12)$ and $b_I \simeq 1.6$ for $(\log_{10}(M_{\rm h,min}/M_\odot),\, \log_{10}(M_{\rm h,max}/M_\odot)) = (9,\, \infty)$. This is a consequence of the intensity bias being weighted more toward small-mass halos given the relation in the simulation, which can be compared with $b_I \simeq 3$ assumed in R. A. C. Croft et al. (2016). Given the 10%–20% accuracy of the fitting formula in J. L. Tinker et al. (2010), it is reasonable to estimate that $1.5 \lesssim b_I \lesssim 2$.

9. Discussion

9.1. Possible Data Processing Improvements

This work has demonstrated the feasibility of detecting the LAE-Lyα intensity cross-power spectrum in HETDEX data and provides a constraint on $b_{\rm g} b_I \langle I\rangle \bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$. We have used extensive masking and PCA to remove systematics from our results. This resulted in significant signal loss. This subsection describes three ideas for improving the HETDEX LIM measurement in the future.

1. Implementing a better full-frame sky subtraction technique would reduce systematics. This includes identifying low-level systematic contributions associated with individual fiber spectra or entire IFUs. With these improvements, we would not have to mask vast regions of the observed HETDEX volume or apply the PCA.
2. Using more robust weights for the intensity and galaxy catalogs could improve the statistical significance of the power spectrum detection.
3. The completeness model and luminosity function used in this work could be updated to better match the redshift distribution of LAEs.

9.2. Possible Modeling Improvements

Within the statistical uncertainty, the measured LAE-Lyα intensity cross-power spectra are consistent with our mock based on a lognormal model (A. Agrawal et al. 2017; M. Lujan Niemeyer et al. 2023). This allows us to constrain $b_{\rm g} b_I \langle I_{\rm undet}\rangle \bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$. Fast lognormal mocks are therefore essential because they allow us to study many systematic effects in the data and enable us to faithfully forward-model the data.

However, future higher-fidelity measurements with lower statistical uncertainty will require more detailed models with better treatment of the relevant physics. Although cosmological simulations with Lyα RT are still computationally expensive, they are essential for a more accurate modeling of physics (see, e.g., C. Behrens et al. 2018; C. Byrohl et al. 2019; S. Gurung-López et al. 2021; C. Byrohl & D. Nelson 2023; H. Khoraminezhad et al. 2025).

While our current model includes RSD, the LAE luminosity function, and the HETDEX sensitivity and masking function, it is otherwise simple. For example, all galaxies have the same bias, luminosity function, and emission-line width, regardless of their environment. The model only includes Lyα emission sourced by galaxies expected from extrapolating the luminosity function of H. Umeda et al. (2025) to faint luminosities. Future models could include diffuse Lyα emission of the CGM and IGM, and the scattering of Lyα photons into or out of the LOS. These emissions source Lyα halos around galaxies, which have been found ubiquitously at $z > 2$ (e.g., C. C. Steidel et al. 2011; L. Wisotzki et al. 2018; M. Lujan Niemeyer et al. 2022a, 2022b). They could also model the Lyα absorption around LAEs, which we discuss further in Section 9.4.

9.3. Contamination from [O II]-emitting Galaxies

Despite thoroughly masking spectra with continuum emission and detected low-redshift sources, some interloper contamination may remain. If present in both the galaxy catalog and the intensity map, the signal of these interlopers would contaminate the cross-power spectrum of the detected LAEs and the Lyα intensity measurement.

However, we believe that any contamination in the HETDEX LAE catalog is likely to be small. D. Davis et al. (2023a) estimated that [O II]-emitting galaxies make up only $\simeq$1.2% ± 0.1% of the HETDEX LAE catalog, and contamination from other sources is only $\simeq$0.8% ± 0.1%.

9.4. Lyα Absorption around LAEs

By stacking the spectra of tens of thousands of LAEs detected in HETDEX, D. Davis et al. (2023b) and L. H. Weiss et al. (2024) found broad, negative absorption troughs around the stacked Lyα emission lines.

L. H. Weiss et al. (2025) argued that these troughs are due to the extragalactic background light (EBL) that is absorbed in the CGM surrounding the LAEs. If HETDEX detects LAEs preferentially on the near side of overdensities—for example, because less Lyα emission is absorbed on the way to Earth—then the EBL is enhanced behind the LAEs and absorbed around the Lyα wavelength. The sky subtraction removes the mean EBL spectrum from all spectra, including those where the EBL is completely absorbed. Consequently, the absorption troughs become negative, explaining the data.

M. M. Khanlari et al. (2025) stacked the spectra as a function of the angular distance from the LAEs and find Lyα absorption out to $\simeq$350 kpc (proper). The Lyα absorption at LAE positions and in their surroundings is stronger for sources with a smaller detection significance, and is not observed for LAEs with SNR > 6.

The expected effect of this absorption on our measurement is twofold. First, the absorption around LAEs can be modeled as part of the emission-line profile. This profile can be written as $\phi(\lambda) = \phi_{\rm em}(\lambda) + \phi_{\rm abs}(\lambda)$, where $\phi_{\rm em}(\lambda)$ is the shape of the emission line, and $\phi_{\rm abs}(\lambda)$ is the shape of the absorption feature (e.g., a Voigt profile). One can then model an intensity map from delta functions at the locations of point sources and convolve this map with $\phi(\lambda)$ (see, e.g., M. Lujan Niemeyer et al. 2023). This convolution dampens the power spectrum, $P(\boldsymbol{k})$, along the LOS. In this work, we have neglected $\phi_{\rm abs}(\lambda)$ and modeled $\phi(\lambda) \simeq \phi_{\rm em}(\lambda)$, including the LOS resolution of HETDEX. We have also not included continuum emission or the EBL in the model. The measured cross-power spectrum of detected LAEs with undetected Lyα intensity is consistent with the model without the absorption feature, suggesting that Lyα absorption found in D. Davis et al. (2023b) and L. H. Weiss et al. (2024) does not dominate the cross-correlation signal.

Second, we only considered sources with SNR > 5.5 as detected; thus, galaxies with lower detection significance and undetected galaxies contribute to the Lyα intensity map. If the intensity map were dominated by a continuum background that is absorbed in overdensities, the cross-correlation between the detected LAEs and the Lyα intensity map excluding detected LAEs would be negative. However, since we detect a positive signal, we conclude that Lyα emission dominates the intensity map.

More sophisticated modeling of and exploration of the absorption feature will be left for future work.

10. Summary and Conclusions

In this paper, we presented a Lyα LIM power spectrum measurement in HETDEX. We reported the detection of the LAE-Lyα intensity cross-power spectrum of LAEs with SNR > 5.5 and the intensity of undetected sources in three redshift bins centered around $\bar{z} = 2.1$, 2.6, and 3.2.

To accomplish this measurement, we thoroughly cleaned the spectral data to remove systematic contributions. We created self-consistent lognormal mocks for the LAEs and the Lyα intensity, including RSD, using the HETDEX window function, sensitivity, and noise (M. Lujan Niemeyer et al. 2023; M. Lujan Niemeyer 2025). We estimated the covariance matrix of the power spectra from the mocks and from the shuffled data.

By fitting the fiducial mock prediction times the overall amplitude to the measured power spectra, we constrained the product of the detected LAE bias, the intensity bias, the mean intensity of undetected sources, and the ratio of the real and fiducial RSD factors, $b_{\rm g} b_I \langle I \rangle \bar{F}_{\rm RSD}/\bar{F}_{\rm RSD}^{\rm fid}$.

Assuming a fiducial LAE bias $b_{\rm g}^{\rm fid} = 2$ and $\bar{F}_{\rm RSD} = \bar{F}_{\rm RSD}^{\rm fid}$, we inferred lower values of $b_I \langle I \rangle$ than those from the QSO-Lyα cross-correlations (R. A. C. Croft et al. 2016, 2018; X. Lin et al. 2022). Our constraint, however, is slightly higher than the upper limit inferred from the Lyα forest–Lyα intensity cross-correlation measurement (R. A. C. Croft et al. 2018).

Assuming fiducial LAE and intensity biases of $b_{\rm g}^{\rm fid} = 2$ and $b_I^{\rm fid} = 2$, and $\bar{F}_{\rm RSD} = \bar{F}_{\rm RSD}^{\rm fid}$, our constraints on the Lyα luminosity density are slightly larger than those obtained by integrating extrapolated LAE luminosity functions. Our constraints are on the same order of magnitude as the intensity of faint galaxies and of the CGM and IGM emission in the cosmological simulations with Lyα RT at $z = 2$ (C. Byrohl & D. Nelson 2023). Our measured intensity is smaller than the total Lyα intensity expected from the star formation rate density at $z \simeq 2$–3, and is consistent with an escape fraction of Lyα photons from the ISM being $f_{\rm esc} < 1$. These results will be useful for constraining models of galaxy formation and evolution.

This work represents initial LIM results from the HETDEX data. We will improve the data processing and modeling of the signal in the future.

Acknowledgments

HETDEX is led by the University of Texas at Austin McDonald Observatory and Department of Astronomy with participation from the Ludwig-Maximilians-Universiät München, Max-Planck-Institut für Extraterrestrische Physik (MPE), Leibniz-Institut für Astrophysik Potsdam (AIP), Texas A&M University, Pennsylvania State University, Institut für Astrophysik Göttingen, The University of Oxford, Max-Planck-Institut für Astrophysik (MPA), The University of Tokyo, and Missouri University of Science and Technology. In addition to Institutional support, HETDEX is funded by the National Science Foundation (grant AST-0926815), the State of Texas, the US Air Force (AFRL FA9451-04-2- 0355), and generous support from private individuals and foundations.

The observations were obtained with the Hobby-Eberly Telescope (HET), which is a joint project of the University of Texas at Austin, the Pennsylvania State University, Ludwig-Maximilians-Universität München, and Georg-August-Universität Göttingen. The HET is named in honor of its principal benefactors, William P. Hobby and Robert E. Eberly.

VIRUS is a joint project of the University of Texas at Austin, Leibniz-Institut für Astrophysik Potsdam (AIP), Texas A&M University (TAMU), Max-Planck-Institut für Extraterrestrische Physik (MPE), Ludwig-Maximilians-Universität Muenchen, Pennsylvania State University, Institut für Astrophysik Göttingen, University of Oxford, and the Max-Planck-Institut für Astrophysik (MPA). In addition to Institutional support, VIRUS was partially funded by the National Science Foundation, the State of Texas, and generous support from private individuals and foundations.

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high-performance computing, visualization, and storage resources that have contributed to the research results reported within this paper. URL: http://www.tacc.utexas.edu.

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2094—390783311.

J.L.B. acknowledges funding from the Ramón y Cajal grant RYC2021-033191-I, financed by MCIN/AEI/10.13039/501100011033 and by the European Union “NextGenerationEU”/PRTR, as well as the project UC-LIME (PID2022-140670NA-I00), financed by MCIN/AEI/10.13039/501100011033/FEDER, UE.

M.J.J. acknowledges the support of a UKRI Frontiers Research grant [EP/X026639/1], which was selected by the European Research Council, and the STFC consolidated grants [ST/S000488/1] and [ST/W000903/1]. M.J.J. also acknowledges support from the Oxford Hintze Centre for Astrophysical Surveys, which is funded through generous support from the Hintze Family Charitable Foundation.

The Institute for Gravitation and the Cosmos is supported by the Eberly College of Science and the Office of the Senior Vice President for Research at the Pennsylvania State University.

R.C. acknowledges support from the National Science Foundation under grant AST-2408358. S.S. acknowledges support from the National Science Foundation under grants NSF-2219212 and NSF-2511145.

Software: astropy (Astropy Collaboration et al. 2013, 2018), dustmaps (G. Green 2018), extinction https://extinction.readthedocs.io/en/latest/ matplotlib (J. D. Hunter 2007), numpy (C. R. Harris et al. 2020), scipy (P. Virtanen et al. 2020).

Appendix A
Effect of Nonlinearity on the Power Spectrum

To see the effects of nonlinear corrections to the matter power spectrum, we generate a new set of mocks with a linear input matter power spectrum generated by CLASS, i.e., not using Halofit. Otherwise, we keep the same settings as the fiducial mocks described in Section 3. In both cases, nonlinearity in RSD due to the Jacobian of the coordinate transformation from real space to redshift space is included (A. Agrawal et al. 2017). We also process the mock maps in the same way, including sky subtraction.

Figure 12 shows the LAE-Lyα intensity cross-power spectrum of the linear mock compared to that of the fiducial nonlinear mock in the Spring field. The power spectra are similar even at $k > 0.5\,h\,{\rm Mpc}^{-1}$ across all explored redshift bins from $z = 1.88$ to 3.52, as nonlinearity due to structure formation is mild at such high redshift (D. Jeong & E. Komatsu 2006).

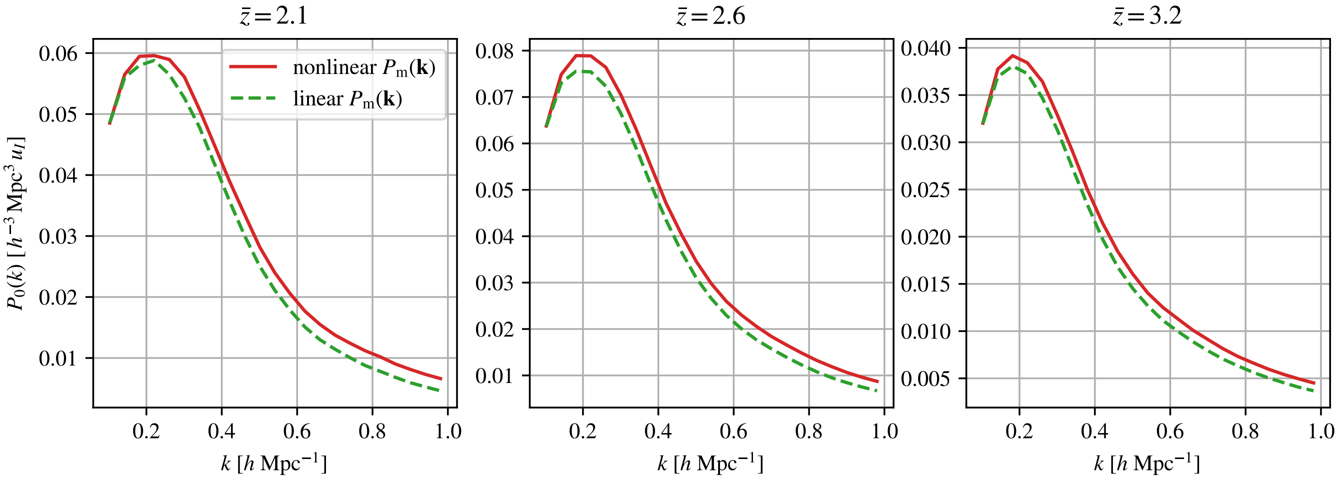

Figure 12. The red solid lines show the fiducial nonlinear mock cross-power spectra, while the green dashed lines show the linear mock cross-power spectra. These two power spectra are compared in the Spring field across three redshift bins. Here, $u_I = 10^{-18}$ erg s^{-1} cm^{-2} arcsec^{-2} Å^{-1}.

Appendix B
$F_{\rm RSD}$ Including the Lyα RT Effect

When Lyα photons travel from a galaxy to an observer on Earth, they can be scattered out of the LOS by hydrogen atoms in the IGM, which can be modeled as an effective absorption. This reduces the observed Lyα luminosity of a galaxy, making it less likely to be detected, and the observed Lyα intensity in the intensity map. This effect depends on the density and velocity gradient of the gas between the observed galaxy and us. If a significant fraction of Lyα photons are subject to scattering in the IGM, this can affect the observed clustering statistics of LAEs and Lyα LIM (e.g., Z. Zheng et al. 2011; C. Behrens & J. Niemeyer 2013; C. Behrens et al. 2018). Specifically, it can change the anisotropy of the power spectrum in redshift space. In a linear model for RSD and Lyα RT following J. S. B. Wyithe & M. Dijkstra (2011) and M. Lujan Niemeyer (2025), the RSD factor reads

$$F_{\rm RSD}(\boldsymbol{k}) = [b_I + b_{\rm ion}K_\lambda(k)C - c_\gamma C + (1 - C)f\mu^2] \\ [b_{\rm g} + b_{\rm ion}K_\lambda(k)C^{{\rm g}_\alpha} - c_\gamma C^{{\rm g}_\alpha} + (1 - C^{{\rm g}_\alpha})f\mu^2], \tag{B1}$$

where b_I and b_g are the intensity and LAE bias, respectively, $b_{\rm ion}$ is the bias of ionizing sources, and $c_\gamma \simeq 1.72$. The smoothing kernel $K_\lambda(k) = \arctan(k\lambda_{\rm mfp})/(k\lambda_{\rm mfp})$, with the mean free path of ionizing photons $\lambda_{\rm mfp}$, translates the locations of ionizing sources to a map of ionization rate fluctuations. The constants C and $C^{{\rm g}_\alpha}$ quantify the effect of the Lyα absorption in the IGM on the intensity fluctuation and the detected galaxy overdensity, respectively:

$$C = \frac{F_{\rm abs}\tau_0 e^{-\tau_0}}{1 - F_{\rm abs} + F_{\rm abs}e^{-\tau_0}}; \tag{B2}$$

$$C^{{\rm g}_\alpha} = (\beta_\phi - 1)\frac{F_{\rm abs}\tau_0 e^{-\tau_0}}{1 - F_{\rm abs} + F_{\rm abs}e^{-\tau_0}}. \tag{B3}$$

Here, $F_{\rm abs} \in [0, 1]$ is the fraction of Lyα photons that can be absorbed in the IGM. The fraction $1 - F_{\rm abs}$ travels to the observer unobstructed, for example, because they have redshifted out of resonance before escaping the galaxy. τ_0 is the mean effective optical depth of the IGM; $\beta_\phi = -\alpha$ is -1 times the faint-end slope of the Lyα luminosity function. Note that the linear model for Lyα RT is only valid for small overdensities and breaks down in the nonlinear environment around galaxies.

Appendix C
Correlation Matrices

We calculate the correlation matrices M_{ij} of the power spectrum monopoles shown in Figure 8 as $M_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$, where C_{ij} is the total covariance matrix described in Section 4.2. Figure 13 shows the correlation matrices. They are mostly diagonal.

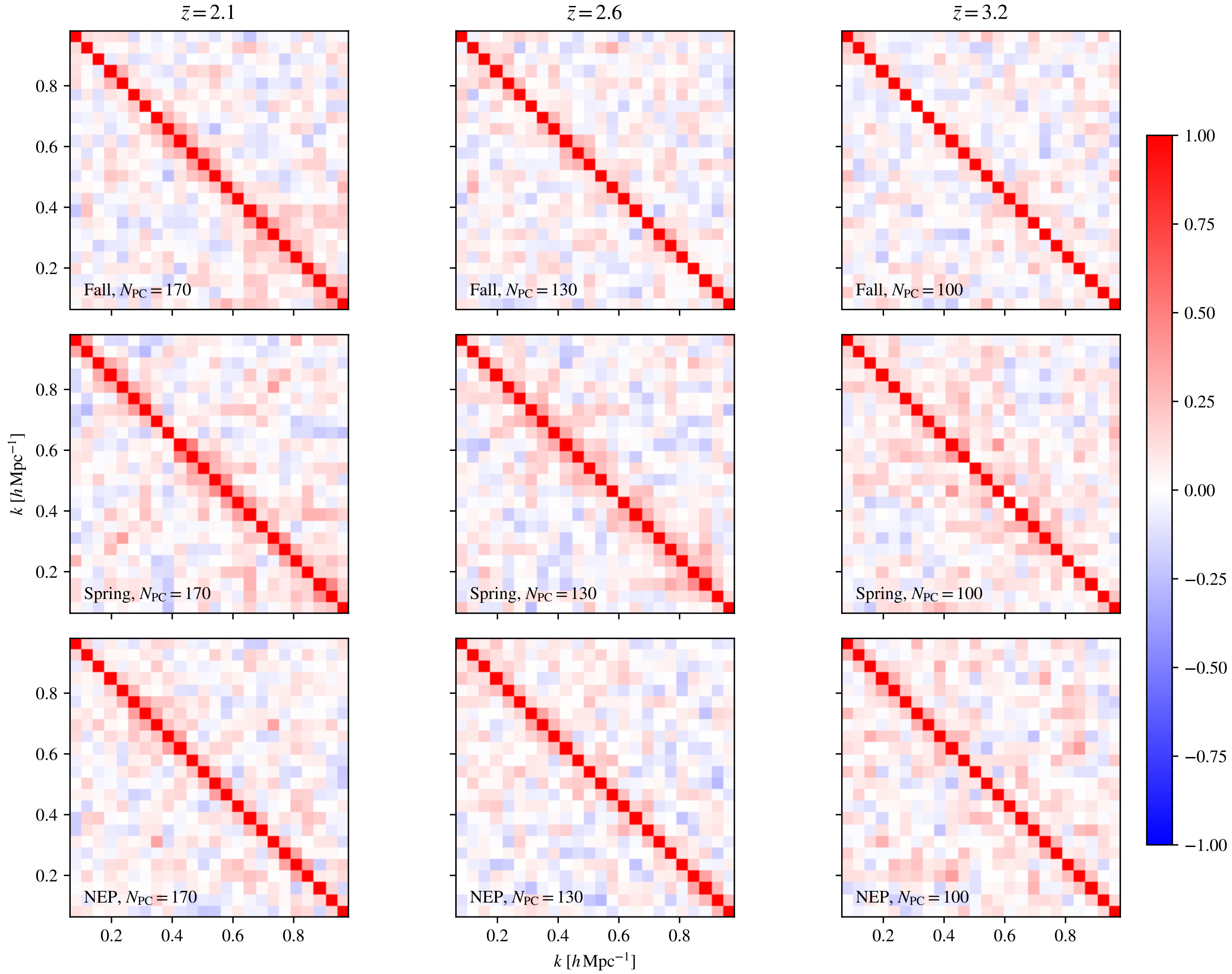

Figure 13. Correlation matrices of the power spectrum monopoles of Figure 8, in the same order.

ORCID iDs

Maja Lujan Niemeyer https://orcid.org/0000-0002-6907-8370
Eiichiro Komatsu https://orcid.org/0000-0002-0136-2404
José Luis Bernal https://orcid.org/0000-0002-0961-4653
Chris Byrohl https://orcid.org/0000-0002-0885-8090
Robin Ciardullo https://orcid.org/0000-0002-1328-0211
Olivia Curtis https://orcid.org/0000-0002-0212-4563
Daniel J. Farrow https://orcid.org/0000-0003-2575-0652
Steven L. Finkelstein https://orcid.org/0000-0001-8519-1130
Karl Gebhardt https://orcid.org/0000-0002-8433-8185
Caryl Gronwall https://orcid.org/0000-0001-6842-2371
Gary J. Hill https://orcid.org/0000-0001-6717-7685
Matt J. Jarvis https://orcid.org/0000-0001-7039-9078
Donghui Jeong https://orcid.org/0000-0002-8434-979X
Erin Mentuch Cooper https://orcid.org/0000-0002-2307-0146
Deeshani Mitra https://orcid.org/0000-0002-1350-019X
Shiro Mukae https://orcid.org/0000-0003-3823-8279
Julian B. Muñoz https://orcid.org/0000-0002-8984-0465
Masami Ouchi https://orcid.org/0000-0002-1049-6658
Shun Saito https://orcid.org/0000-0002-6186-5476
Donald P. Schneider https://orcid.org/0000-0001-7240-7449
Lutz Wisotzki https://orcid.org/0000-0003-2977-423X

References

Agrawal, A., Makiya, R., Chiang, C.-T., et al. 2017, JCAP, 2017, 003
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Behrens, C., Byrohl, C., Saito, S., & Niemeyer, J. C. 2018, A&A, 614, A31
Behrens, C., & Niemeyer, J. 2013, A&A, 556, A5
Bernal, J. L., & Kovetz, E. D. 2022, A&ARv, 30, 5
Bird, S., Viel, M., & Haehnelt, M. G. 2012, MNRAS, 420, 2551
Blas, D., Lesgourgues, J., & Tram, T. 2011, JCAP, 7, 034
Byrohl, C., & Nelson, D. 2023, MNRAS, 523, 5248
Byrohl, C., Nelson, D., Behrens, C., et al. 2021, MNRAS, 506, 5129
Byrohl, C., Saito, S., & Behrens, C. 2019, MNRAS, 489, 3472
CCAT-Prime Collaboration, Aravena, M., Austermann, J. E., et al. 2023, ApJS, 264, 7
Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Natur, 466, 463
Chávez Ortiz, Ó. A., Finkelstein, S. L., Davis, D., et al. 2023, ApJ, 952, 110
Chen, Z., Cunnington, S., Pourtsidou, A., et al. 2025, ApJS, 279, 19
Cleary, K. A., Borowska, J., Breysse, P. C., et al. 2022, ApJ, 933, 182

CONCERTO Collaboration, Ade, P., Aravena, M., et al. 2020, A&A, 642, A60
Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320
Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541
Cunnington, S., Li, Y., Santos, M. G., et al. 2023, MNRAS, 518, 6262
Davis, D., Gebhardt, K., Cooper, E. M., et al. 2023a, ApJ, 946, 86
Davis, D., Gebhardt, K., Cooper, E. M., et al. 2023b, ApJ, 954, 209
DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001
Dijkstra, M. 2019, SAAS, 46, 1
Dong, C., Lee, K.-G., Ata, M., Horowitz, B., & Momose, R. 2023, ApJL, 945, L28
Doré, O., Bock, J., Ashby, M., et al. 2014, arXiv:1412.4872
Dunne, D. A., Cleary, K. A., Breysse, P. C., et al. 2024, ApJ, 965, 7
Dunne, D. A., Cleary, K. A., Breysse, P. C., et al. 2025, A&A, 702, A247
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948
Gawiser, E., Francke, H., Lai, K., et al. 2007, ApJ, 671, 278
Gebhardt, K., Mentuch Cooper, E., Ciardullo, R., et al. 2021, ApJ, 923, 217
Green, G. 2018, JOSS, 3, 695
Guaita, L., Gawiser, E., Padilla, N., et al. 2010, ApJ, 714, 255
Gurung-López, S., Saito, S., Baugh, C. M., et al. 2021, MNRAS, 500, 603
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Herrera, D., Gawiser, E., Benda, B., et al. 2025, ApJL, 988, L57
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
House, L. R., Gebhardt, K., Finkelstein, K., et al. 2023, ApJ, 950, 82
House, L. R., Gebhardt, K., Finkelstein, K., et al. 2024, ApJ, 975, 172
Hunter, J. D. 2007, CSE, 9, 90
Jeong, D., & Komatsu, E. 2006, ApJ, 651, 619
Kaiser, N. 1987, MNRAS, 227, 1
Kakuma, R., Ouchi, M., Harikane, Y., et al. 2021, ApJ, 916, 22
Karkare, K. S., Anderson, A. J., Barry, P. S., et al. 2022, JLTP, 209, 758
Keating, G. K., Marrone, D. P., Bower, G. C., & Keenan, R. P. 2020, ApJ, 901, 141
Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34
Keenan, R. P., Keating, G. K., & Marrone, D. P. 2022, ApJ, 927, 161
Khanlari, M. M., Gebhardt, K., Weiss, L. H., et al. 2025, ApJ, 989, 169
Khoraminezhad, H., Saito, S., Gronke, M., & Byrohl, C. 2025, OJAp, 8
Kikuchihara, S., Harikane, Y., Ouchi, M., et al. 2022, ApJ, 931, 97
Kikuta, S., Matsuda, Y., Inoue, S., et al. 2023, ApJ, 947, 75
Konno, A., Ouchi, M., Nakajima, K., et al. 2016, ApJ, 823, 20
Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, arXiv:1709.09066
Kusakabe, H., Verhamme, A., Blaizot, J., et al. 2022, A&A, 660, A44
Lin, X., Zheng, Z., & Cai, Z. 2022, ApJS, 262, 38
Lujan Niemeyer, M. 2025, ApJ, 980, 250
Lujan Niemeyer, M., Bernal, J. L., & Komatsu, E. 2023, ApJ, 958, 4
Lujan Niemeyer, M., Bowman, W. P., Ciardullo, R., et al. 2022b, ApJL, 934, L26
Lujan Niemeyer, M., Komatsu, E., Byrohl, C., et al. 2022a, ApJ, 929, 90
Mentuch Cooper, E., Gebhardt, K., Davis, D., et al. 2023, ApJ, 943, 177
Miller, J. S. A., Bolton, J. S., & Hatch, N. A. 2021, MNRAS, 506, 6001
Ouchi, M., Ono, Y., & Shibuya, T. 2020, ARA&A, 58, 617
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
Paul, S., Santos, M. G., Chen, Z., & Wolz, L. 2023, arXiv:2301.11943
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Raiter, A., Schaerer, D., & Fosbury, R. A. E. 2010, A&A, 523, A64
Ramsey, L. W., Adams, M. T., Barnes, T. G., et al. 1998, SPIE, 3352, 34
Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, MNRAS, 461, 1100
Renard, P., Spinoso, D., Montero-Camacho, P., et al. 2024, Monthly Notices of the Royal Astronomical Society, 535, 826
Renard, P., Spinoso, D., Sun, Z., et al. 2025, in Proc. of the XVI Scientific Meeting of the Spanish Astronomical Society, ed. M. Manteiga, F. González-Galindo, & A. Labiano-Ortega (Granada, Spain: Spanish Astronomical Society)
Santos, M., Bull, P., Camera, S., et al. 2016, in MeerKAT Science: On the Pathway to the SKA (Trieste: SISSA), 32
Schechter, P. 1976, ApJ, 203, 297
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2011, ApJ, 736, 160
Sun, G., Chang, T. C., Uzgil, B. D., et al. 2021, ApJ, 915, 33
Switzer, E. R., Barrentine, E. M., Cataldo, G., et al. 2021, JATIS, 7, 044004
Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
Tornotti, D., Fumagalli, M., Fossati, M., et al. 2025, ApJL, 980, L43
Trainor, R. F., Lamb, N. R., Steidel, C. C., et al. 2025, arXiv:2505.15881
Umeda, H., Ouchi, M., Kikuta, S., et al. 2025, ApJS, 277, 37
Vieira, J., Aguirre, J., Bradford, C. M., et al. 2020, arXiv:2009.14340
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Weiss, L. H., Davis, D., Gebhardt, K., et al. 2024, ApJ, 962, 102
Weiss, L. H., Gebhardt, K., Davis, D., et al. 2025, ApJ, 983, 72
Wisotzki, L., Bacon, R., Brinchmann, J., et al. 2018, Natur, 562, 229
Wyithe, J. S. B., & Dijkstra, M. 2011, MNRAS, 415, 3929
Zheng, Z., Cen, R., Trac, H., & Miralda-Escudé, J. 2011, ApJ, 726, 38