The Astrophysical Journal, 936:131 (23pp), 2022 September 8 https://doi.org/10.3847/1538-4357/ac8546

Stellar Populations of Lyα-emitting Galaxies in the HETDEX Survey. I. An Analysis of LAEs in the GOODS-N Field

Adam P. McCarron[1], Steven L. Finkelstein[1], Oscar A. Chavez Ortiz[1], Dustin Davis[1], Erin Mentuch Cooper[1,2], Intae Jung[3,4,5], Delaney R. White[1], Gene C. K. Leung[1], Karl Gebhardt[1], Viviana Acquaviva[6,7], William P. Bowman[8,9], Robin Ciardullo[8,9], Eric Gawiser[10], Caryl Gronwall[8,9], Gary J. Hill[1,2], Wolfram Kollatschny[11], Martin Landriau[12], Chenxu Liu[1], Daniel N. Mock[13], and Ariel G. Sánchez[14]
[1] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA; apm.astro@utexas.edu
[2] McDonald Observatory, University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA
[3] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[4] Department of Physics, The Catholic University of America, Washington, DC 20064, USA
[5] Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA
[6] Physics Department, NYC College of Technology, 300 Jay Street, Brooklyn, NY 11201, USA
[7] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[8] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[9] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[10] Department of Physics and Astronomy, Rutgers, The State University, Piscataway, NJ 08854, USA
[11] Institut für Astrophysik, Universität Göttingen, Friedrich-Hund Platz 1, D-37077 Göttingen, Germany
[12] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[13] Department of Physics, Florida State University, Tallahassee, FL 32306, USA
[14] Max-Planck-Institut für extraterrestrische Physik, Postfach 1312, Giessenbachstr., D-85748 Garching, Germany
Received 2022 January 26; revised 2022 July 11; accepted 2022 July 27; published 2022 September 8

Abstract

We present the results of a stellar population analysis of 72 Lyα-emitting galaxies (LAEs) in GOODS-N at $1.9 < z < 3.5$ spectroscopically identified by the Hobby−Eberly Telescope Dark Energy Experiment (HETDEX). We provide a method for connecting emission-line detections from the blind spectroscopic survey to imaging counterparts, a crucial tool needed as HETDEX builds a massive database of ∼1 million Lyα detections. Using photometric data spanning as many as 11 filters covering $0.4 < \lambda\,(\mu\mathrm{m}) < 4.5$ from the Hubble Space Telescope and Spitzer Space Telescope, we study the objects' global properties and explore which properties impact the strength of Lyα emission. We measure a median stellar mass of $0.8^{+2.9}_{-0.5} \times 10^9\ M_\odot$ and conclude that the physical properties of HETDEX spectroscopically selected LAEs are comparable to LAEs selected by previous deep narrowband studies. We find that stellar mass and star formation rate correlate strongly with the Lyα equivalent width. We then use a known sample of $z > 7$ LAEs to perform a protostudy of predicting Lyα emission from galaxies in the epoch of reionization, finding agreement at the 1σ level between prediction and observation for the majority of strong emitters.

Unified Astronomy Thesaurus concepts: Galaxy evolution (594); Galaxy formation (595); Lyman-alpha galaxies (978); High-redshift galaxies (734)

1. Introduction

Lyα-emitting galaxies (hereafter LAEs) have fascinated astronomers for decades, from when Partridge & Peebles (1967) first predicted that primitive galaxies in formation could emit a detectable Lyα line through their discovery by Cowie & Hu (1998) and Rhoads et al. (2000). These objects exhibit strong emission of the Lyα photon corresponding to the $n = 2$ to $n = 1$ resonant transition in hydrogen atoms. These photons are expected to face high optical depths from neutral hydrogen to escape the galaxies in which they are generated, and dust grains along their paths can absorb them. To date, despite enormous effort (see Ouchi et al. 2020 for review), the community has not formed a strong consensus on exactly how Lyα radiation escapes its host galaxy, and no reliable model exists to predict the Lyα luminosity or equivalent width, W_λ(Lyα), of a galaxy given its global physical properties, such as stellar mass, metallicity, age, star formation rate (SFR), and dust extinction.

Part of the problem arises from discrepant conclusions drawn from studying LAEs identified using different selection techniques. Locally ($z \ll 1$), the ultraviolet (UV) flux measured in broadband or narrowband (NB) filters often defines LAE samples, biasing studies to brighter, higher-mass systems than those found spectroscopically (Hayes et al. 2014). Observations in the nearby universe paint LAEs as low-mass galaxies with young stellar ages as determined from spectral energy distribution (SED) fitting, and many studies concur on trends showing an increase in Lyα luminosity with decreasing dust and metals (Hayes 2015). Nonetheless, many galaxies show stronger Lyα emission than models would predict based on dust extinction (e.g., Martin et al. 2015; Atek et al. 2014; Scarlata et al. 2009; Finkelstein et al. 2009), and a satisfactory explanation of this Lyα enhancement does not currently exist.

With NB-selected LAEs at higher redshift, discrepant results still persist. Finkelstein et al. (2009) found that LAEs at $z \sim 4.5$ represent a diverse population in terms of stellar age, mass, and dust extinction. Finkelstein et al. (2015) modeled the SEDs of IRAC-detected LAEs at $z \sim 5$ and found that a third have old

stellar populations, contrasting with the young populations found in the local universe, and Guaita et al. (2011) observed similar heterogeneous populations in an NB-selected sample at $z \simeq 2.1$. Moreover, Gawiser et al. (2007) found NB-selected LAEs at $z = 3.1$ to generally be low-mass, dust-free objects, but their model allowed for both young and more evolved stellar populations, and Acquaviva et al. (2012) found LAEs at $z = 3.1$ to be older than those at $z = 2.1$. Kornei et al. (2010) compiled a UV-continuum-selected sample of $z \sim 3$ galaxies, finding that those with strong Lyα emission had older stellar populations with lower SFRs and less dust. Recently, Santos et al. (2020) used SED fitting of nearly 4000 LAEs in the COSMOS field at $2 < z < 6$ to find that LAEs were younger and/or more dust-poor than other UV-selected objects based on their UV slopes.

Studies of LAE samples compiled using detection of the Lyα emission line itself in the high-redshift universe confound consensus as well. Hagen et al. (2016) used the Hobby–Eberly Telescope Dark Energy Experiment (HETDEX) pilot survey (Adams et al. 2011; Blanc et al. 2011) to compare properties of LAEs at $z \sim 2$ with optical emission-line galaxies (oELGs) and found no significant differences between the populations. Remarkably, even the UV slope did not differ in the two samples, implying either that diffuse dust in the interstellar medium (ISM) did not modulate Lyα emission or that oELGs strongly emit Lyα. Recently, spectroscopic surveys have also yielded confusing results about LAEs at $z > 2$. Using data from the VANDELS survey, Marchi et al. (2019) suggested that LAEs have low mass and low dust extinction but found no correlation with SFR. From the VIMOS Ultra-Deep Survey, Hathi et al. (2016) concurred with LAEs having lower mass and lower dust extinction, but they found that the objects have lower SFRs than non-LAEs. Approaching the problem from the other direction, Oyarzún et al. (2017) found from studying the spectra of stellar-mass-selected galaxies at $3 < z < 4.6$ that a negative correlation existed between Lyα equivalent width and both stellar mass and SFR. A review of the field's current knowledge of high-redshift Lyα emission can be found in Ouchi et al. (2020).

A deeper understanding of what makes LAEs unique from other star-forming galaxies (SFGs) tantalizes astronomers because of the profound implications for leveraging LAEs as sensitive probes of reionization at $z \gtrsim 6$. Whether the universe reionized rapidly at late times (e.g., Robertson et al. 2015) or gradually beginning very early in its history (e.g., Finkelstein et al. 2019) can determine whether massive, rare galaxies or low-mass, ubiquitous objects emitted the needed ionizing photons. Answering such a fundamental cosmological question hinges on our ability to detect neutral hydrogen in the universe's infancy. Crucially, the attenuation of Lyα photons can probe the presence of neutral hydrogen in the intergalactic medium (IGM; e.g., Miralda-Escudé 1998; Malhotra & Rhoads 2004; Dijkstra 2014), but the photons also undergo complicated resonant scattering within the galaxy, complicating our understanding of how much of the emission exits the ISM and circumgalactic medium (CGM) and enters the IGM in the first place. Recent attempts to use Lyα as a reionization probe have struggled to account for the intrinsic effects of host galaxy properties on the Lyα luminosity before the radiation encounters the IGM, leaving an unknown systematic uncertainty present in their results. The most detailed spectroscopic studies of post-reionization LAEs point to the covering fraction of optically thick neutral hydrogen (e.g., Reddy et al. 2022) as the key predictor of Lyα escape, but such observations remain expensive and time intensive. Finding correlations between Lyα emission and global properties such as mass and star formation activity, which photometry can reliably measure even at very high redshifts, could be a path forward to predicting galaxies' intrinsic Lyα output.

Small LAE sample sizes (<20) were typical a decade ago, and although recently large samples with >1000 objects have been amassed using NB surveys (e.g., Sobral et al. 2018; Ono et al. 2021), spectroscopically confirmed samples remain small. This has statistically hindered the efficacy of studies of global property correlations with Lyα emission. The HETDEX project (Hill et al. 2008, 2021; Gebhardt et al. 2021) is in the process of discovering a transformative sample of LAEs, clearing the way for the community to obtain a better understanding of this intriguing population. The untargeted (targets not preselected), spectroscopically selected HETDEX LAE sample at $1.9 < z < 3.5$ provides a unique vantage point of galaxy evolution, as these galaxies probe the lower-mass end of the galaxy distribution, making them analogous to typical galaxies discovered in the epoch of reionization (e.g., Finkelstein et al. 2010).

As the first step toward realizing HETDEX's ability to unlock LAEs as probes of reionization, we present an initial study detailing how to link detections from the survey to imaging counterparts, and we provide an SED-fitting analysis of their stellar population properties. Our modest sample of 72 LAEs in the GOODS-N field will pave the way for future large samples from HETDEX to obtain the best understanding of LAEs to date. In Section 2 we describe how we built our sample and selected imaging counterparts. In Section 3 we describe our SED-fitting procedure. We present our results in Section 4, comparing them to other studies, and we discuss our interpretations in Section 5. Finally, we attempt to predict the Lyα emission from a sample of epoch of reionization (EoR) galaxies in Section 6 and summarize this study in Section 7. In our analysis, we adopt a flat ΛCDM cosmology with $H_0 = 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $\Omega_m = 0.30$.

2. Methodology

In order to explore how Lyα emission from galaxies depends on stellar population properties, we built a sample of LAEs using emission-line detections from the HETDEX survey, carefully identifying them as Lyα or other contaminant features, such as [O II] $\lambda\lambda$3726, 3729, which is unresolved at HETDEX resolution. We then created a procedure for assigning the line detections to imaging counterparts in HST data so that we could proceed with fitting their SEDs.

2.1. The HETDEX Survey

With HETDEX the upgraded Hobby–Eberly Telescope (Ramsey et al. 1994; Hill et al. 2021) is observing an area of 540 deg^2 in the north Galactic cap and on the celestial equator using up to 78 pairs of integral field spectrographs that span 350–550 nm at $R \sim 800$. Each spectrograph pair is fed by an integral field unit (IFU) of 448 1.″5-diameter fibers that cover a $51'' \times 51''$ region on the sky with 1/3 fill factor (Kelz et al. 2014; Hill et al. 2021). Each HETDEX observation consists of three 6-minute dithered exposures to fill in the area between fibers, each with >30,000 individual fibers. The majority of

these fibers just contain blank sky, but some subset contain continuum sources such as stars or emission lines from both nearby and distant galaxies.

Gebhardt et al. (2021) describe the data reduction and calibrations needed to convert the raw observations into a three-dimensional spectroscopic data set, as well as the methods used to detect emission lines contained in the millions of observed spectra. As a brief summary, HETDEX reductions involve three types of calibration frames: biases (taken nightly), pixel flats (taken yearly using a laser-driven light source), and twilight sky flats (taken nightly and averaged monthly), which are used for bias subtraction, bad pixel masking, fiber profile tracing, wavelength calibration, scattered light removal, spectral extraction, fiber normalization, spectral masking, and sky subtraction. These frames, combined with sky background on science images, produce a wavelength-calibrated, sky-subtracted spectrum for each fiber in the array.

Astrometric calibrations are achieved by measuring the centroid of each field star and comparing their positions on the IFUs to the stars' equatorial coordinates in the Sloan Digital Sky Survey (SDSS; York et al. 2000; Abazajian et al. 2009) and Gaia (Gaia Collaboration et al. 2018) catalogs. This process typically results in global solutions that are good to $\sim 0\farcs2$ and no worse than $\sim 0\farcs5$, with the exact precision of a measurement dependent on the number of IFUs in operation at the time of the observation.

To find emission lines, the data pipeline searched every spatial and spectral resolution element in the internal HETDEX data release 2 (HDR2) to look for a peak in signal. Regions of enhanced signal were fit with a single Gaussian model with a constant continuum level, a model found adequate for potentially asymmetric line profiles by Gebhardt et al. (2021) because of the low resolution of the VIRUS spectrographs and low signal-to-noise ratio (S/N) of typical sources. The exact location was determined by rastering on a grid and maximizing the line's S/N. An internal catalog of high-quality emission lines was generated by Mentuch Cooper et al. (in preparation), and we drew our initial sample from the HDR2 version of that catalog. The catalog reduced the raw detected line emission sources as described in Gebhardt et al. (2021) into a more robust sample by passing the observations through a quality assessment pipeline and limiting various fitted line parameters. Specifically, emission lines were required to have a quality of fit $\chi^2 < 1.2$ and a line width, σ, in the Gaussian model between 1.7 and 8 Å. The full HETDEX survey will eventually detect $\sim$1 million LAEs, providing an incredible opportunity to study such objects, but our analysis is focused on LAEs discovered in 2018–2020 data from a HETDEX science verification field in GOODS-N, a roughly $10' \times 16'$ field centered at (J2000) $12^h36^m55^s$, $62°14^m15^s$ (Giavalisco et al. 2004; Grogin et al. 2011; Koekemoer et al. 2011) because we required deep, multiband imaging to study each galaxy's stellar populations.

2.2. Sample Selection

We visually inspected HETDEX detections in GOODS-N to obtain a clean sample of LAEs. To get initial candidates, we applied various quality cuts to the curated catalog for data release HDR 2.1.2 (Mentuch Cooper at al., in preparation). We restricted emission-line detections to those with $S/N > 5.5$ to limit the fraction of spurious detections from noise fluctuations to less than 5% (see Gebhardt et al. 2021), as well as $\chi^2 < 1.6$ for the Gaussian model fit, which was a value tuned to remove the most obvious artifacts while retaining the largest sample for inspection. We required emission-line FWHM between 3.4 and 24 Å, where the lower bound removed exceedingly narrow peaks arising from unidentified cosmic rays and the upper bound removed emission generated by broad-line active galactic nuclei (AGNs), which we considered contaminants in this study (see Section 3.1). We further only included observations with throughput >0.07 for reliable flux measurements minimally affected by cloud cover and seeing below $2\farcs8$ to enable continuum counterpart identification. We did not remove "repeat" detections coincident spatially and spectrally resulting from the survey revisiting the field multiple times in order to ensure that we found as many Lyα detections as possible. We excluded data in GOODS-N taken prior to 2018, as they included significant artifacts from early CCDs that had been replaced by 2018.

Finally, we did not initially remove any detections based on the Bayesian probability values used to help determine the identity of an emission line as Lyα versus [O II] $\lambda\lambda$3726, 3729, such as P(Lyα). These probabilities, which are calculated by the HETDEX team based on the work of Leung et al. (2017) and Farrow et al. (2021), leverage the inherent differences between the emission-line luminosity and equivalent width (EW), W_λ, distribution functions of LAEs and [O II] emitters to identify single emission-line detections using information about the line flux and continuum emission, when available. During the process of visual inspection, we used the statistic to guide our identifications, and we make recommendations for using quality cuts based on this statistic at the end of this section. We do not believe that keeping this statistic visible to the classifier biased our results because we implemented an independent procedure (see Section 2.3) to distinguish LAEs from low-redshift counterparts that relied on SED fitting.

After applying quality cuts, we began with 842 detections (of which $\sim$500 were "unique" in the sense that there were no other emission-line detections within 3″ spatially and 6 Å spectrally). To inspect each detection, we used the HETDEX Emission Line eXplorer tool ELiXer (Davis et al., in preparation), which shows measured quantities for the emission lines such as S/N, line width, line fit χ^2, the continuum estimate, and the Bayesian probability for Lyα emission described above, as well as useful visual information, such as cutouts of the 2D spectra for several fibers containing the feature, the Gaussian model fit to the feature, the full 1D spectrum, and any imaging and catalog data uploaded in the HETDEX pipeline.

We rated our confidence in a detection on a scale of 0–5 using a customized widget tool that allows interactive classification of detected sources based on the ELiXer report (see Figure 1). Additionally, other classifications include "artifact," a false detection caused by a malfunction in the instrument or the reduction pipeline, low-redshift sources, and "other" for miscellaneous objects like meteors. To qualify for a classification of 4 or 5 (a high-confidence LAE by our definition), a detection had to meet the following criteria:

1. A clear emission line in at least one fiber in the unsmoothed 2D spectrum, or a probable emission line in at least two fibers. Since each point-spread function (PSF) covers multiple fibers (due to the dithering pattern), we expected strong emission to be seen in more than one fiber, increasing the likelihood of a real detection.

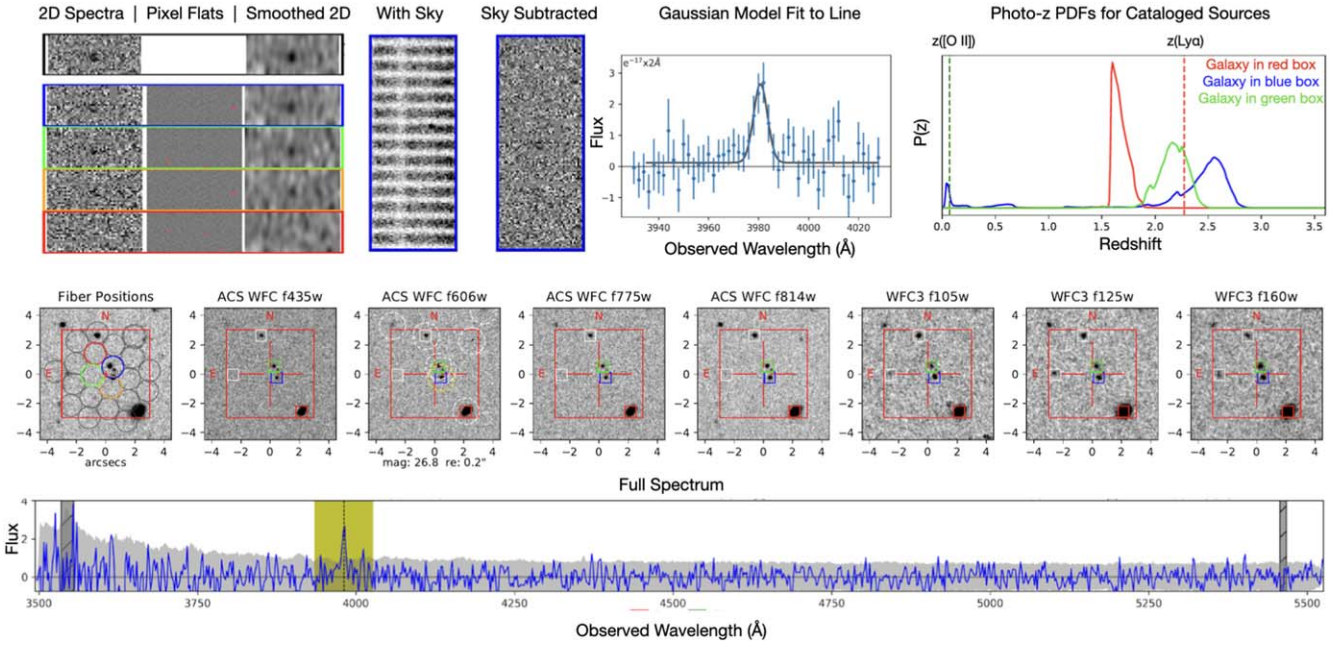


Figure 1. A section of an ELiXer report for detection ID 2100325857, the line from an LAE in our sample. The report contains information about the detected line and imaging at the detection position. The 2D spectra from the four fibers contributing the highest S/N to the detection are in the upper left corner; the stacked signal is shown on the top row, outlined in black. The pixel flats and smoothed 2D spectra are displayed in the right two columns. Plots to the right show the sky subtraction and the emission-line model fit (black line). The middle row contains 10″ imaging cutouts, with the fiber positions shown in gray in the first image and the locations of cataloged sources marked with colorful boxes in all subsequent images. White boxes indicate sources too far to be considered, and white circles show the aperture for brightness measurement in F606W. The AB magnitude of the nearest likely source is reported. The upper right corner shows the photo-z probability distributions calculated by the CANDELS team (B. Andrews et al., in preparation) for sources of matching color in the imaging, and the [O II] and Lyα redshifts are shown as vertical dashed green and red lines. Finally, the bottom row shows the full 1D spectrum. Flux densities have units of erg s^{-1} cm^{-2} 2 Å^{-1}.

2. No obvious defects at the emission-line location in the pixel flat or sky subtraction cutouts. This eliminated hot pixels, sky model residuals, charge traps, and other artifacts from the sample.
3. A Gaussian plus constant continuum model fit that adequately matched the data and did not have an FWHM far below the spectral resolution of ∼6 Å.
4. A line peak that exceeded the typical noise level in multiple pixels in the 1D spectrum.
5. No source at the line's detection position brighter than roughly $m_{\rm AB} = 24$ in the imaging cutouts, if available. The high equivalent widths of sources fainter than this threshold drastically decrease the likelihood of contamination by [O II] emitters (see Figure 6 in Leung et al. 2017), though a few low equivalent width, luminous LAEs can be missed with this requirement.

As the [O II] $\lambda\lambda$3726, 3729 emission feature falls into the 3500 Å $< \lambda <$ 5500 Å spectral range for $z < 0.5$, the imaging proved crucial in choosing between high-redshift LAEs and interloping [O II]-emitting galaxies.

Figure 1 shows an example ELiXer report for a source classified as a high-redshift LAE. Note that for readability, tabulated numeric information such as P(Lyα), line flux, line model χ^2, and more was cropped out of this visualization but was visible to the classifier. In Figure 1, A clear emission feature is present as a black signal in three out of the four 2D unsmoothed fiber spectra, the sky subtraction looks clean, the model fit accurately represents the data, and the image stamps show a number of faint sources with photometric redshift estimates reasonably close to the Lyα redshift (shown by the vertical red dashed line). Figure 2 shows a clear example of a low-redshift object detected by its [O II] emission line. As in Figure 1, the line appears strong in multiple fibers, and the sky subtraction and model fit present no concerns. Characteristically of a brighter low-redshift galaxy, continuum emission is visible as a horizontal black trace in the fiber spectra, and a large, bright object appears in the HST image stamps. In this case, the object is in fact a cataloged [O II] emitter, but even without such information this would be a clear low-redshift classification. In both of these cases, no other emission lines are detected, or would be expected to be detectable, across the observed wavelength range.

After classifying each detection, we obtained ∼200 detections categorized as high-confidence LAEs (scores of 4–5) and almost three times as many classified as low-z sources (Figure 3). Note that we did not include detections with scores of 3 or below for initial study, as we want the cleanest sample possible. To assess the HETDEX collaboration's built-in Bayesian classification probability, P(Lyα), we plotted that statistic for all of our detections classified as either low-z galaxies or LAEs. Figure 4 shows that true LAE detections rarely score low in the P(Lyα) statistic, but a few low-z sources can score in the intermediate range. For this reason, we suggest that future studies can dramatically reduce the amount of visual inspections needed by adopting a cutoff of P(Lyα) $\gtrsim 0.6$ for LAE candidates.

To finalize our sample, we removed detections of the same source (since the GOODS-N field was observed multiple times between 2018 and 2020), by selecting the highest S/N measurement of all detections grouped within 2″ and one spectral resolution element (6 Å). Our final emission-line sample consisted of 94 high-confidence Lyα detections (with classification scores of 4–5).

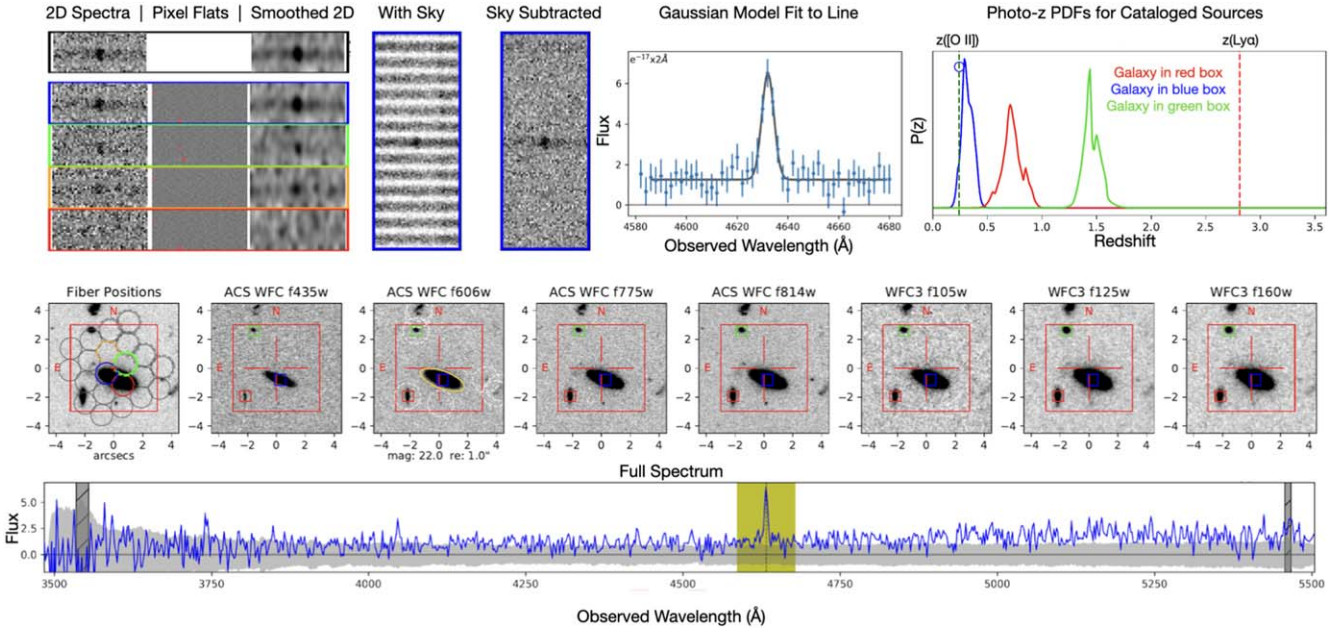


Figure 2. A section of an ELiXer report, in the same format as Figure 1, for detection ID 2100037191, corresponding to the [O II] $\lambda\lambda$3726, 3729 feature in a galaxy at $z \approx 0.24$. The black trace in the 2D fiber spectrum (blue rectangle) indicates a clear detection of continuum emission, which is also evident in the 1D spectrum in at the bottom of the figure. The imaging shows a large, bright source ($m_{\rm AB} = 22.0$) centered on the detection position, and the source has a cataloged spectroscopic redshift consistent with [O II] indicated as an open blue circle in the photometric redshift plot in the upper right corner.

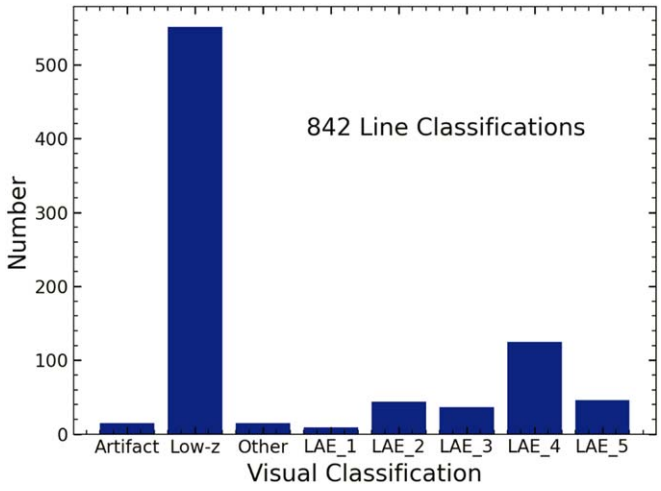


Figure 3. The distribution of visual classifications of candidate detections. Making cuts based on $P(\mathrm{Ly}\alpha)$ to remove low-z sources can dramatically reduce the visual inspection workload.

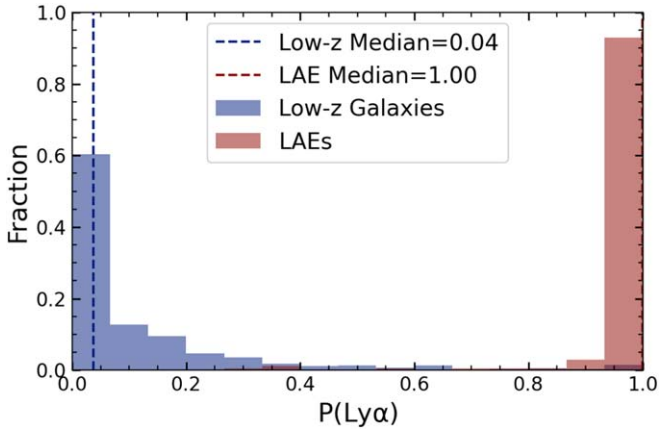


Figure 4. The distribution of $P(\mathrm{Ly}\alpha)$ for detections visually classified as low-redshift galaxies (light blue) and high-confidence (scores of 4 and 5) LAEs (light red). Adopting a minimum threshold for $P(\mathrm{Ly}\alpha)$ can remove a large fraction of low-redshift interlopers without eliminating very many LAEs.

2.3. Counterpart Identification

In order to study the stellar populations of the LAEs in our sample, we developed a method to match the untargeted spectroscopic detections to counterparts in HST imaging of the GOODS-N field.

The overall astrometric precision of a HETDEX observation is $\sim 0\farcs2$. However, due to the $1\farcs5$ diameters of the fibers, the typical seeing, and the three-dither pattern, the position of an individual (faint) LAE is known to no better than $\sim 0\farcs5$. Since the HST images have a resolution that is $\sim$20 times higher than this, great care is needed to ensure that an emission-line source is matched with the correct counterpart.

We used the imaging obtained by the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004) with the optical ACS camera and the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) with the WFC3/IR infrared camera, using the internal CANDELS team's reduced mosaics for each filter. This data set consists of imaging in nine filters (F435W, F606W, F775W, F814W, and F850LP with ACS, and F105W, F125W, F140W, and F160W with WFC3/IR). We made use of the photometric catalogs derived by Finkelstein et al. (2022), which used SOURCEEXTRACTOR (Bertin & Arnouts 1996) in two-image mode to create an F160W-selected catalog, coupled with using the Tractor (Lang et al. 2016) to perform deblended photometry on the deep S-CANDELS (Ashby et al. 2015) Spitzer/IRAC 3.6 and 4.5 μm imaging. Further details on the cataloging process are available in Finkelstein et al. (2022). Similar to the widget used to classify detections as Lyα, we created a visual inspection tool that provided information about the distance between the centroid

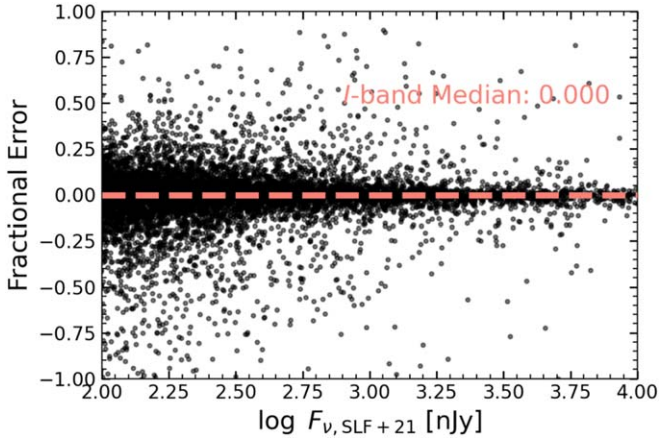


Figure 5. The fractional error between *I*-band (F775W) fluxes of sources measured in our derived stacked-detection catalog matched to sources in the original catalog from Finkelstein et al. (2022) as a function of flux in the latter catalog. The median offset is indicated with the dashed pink line, and its value is given with text, showing good agreement between these two catalogs.

of the HETDEX emission location and a given imaging source, the HETDEX emission-line strength when re-extracted centered at the imaging counterpart position, and the goodness of an SED fit assuming the Lyα redshift, $z_{\mathrm{Ly}\alpha}$.

Before selecting counterpart candidates, we optimized our search by developing a deep photometric catalog using a stacked image across all HST filters in GOODS-N. Each pixel value in this image and its error were computed using an inverse-variance-weighted average across $N = 9$ filters with pixel value p_i and rms error σ_i given by

$$\bar{p} = \frac{\sum_i^N p_i \sigma_i^{-2}}{\sum_i^N \sigma_i^{-2}},\ \sigma_{\bar{p}} = \left(\sum_i^N \sigma_i^{-2}\right)^{-1/2}. \quad (1)$$

Since LAEs are often low-mass, faint systems, this stacked image improved our chances of identifying the continuum source corresponding to the detected emission line.

We then used SOURCEEXTRACTOR (Bertin & Arnouts 1996) to detect the faintest possible sources in the stacked image, requiring a source to have 5 contiguous pixels with $\mathrm{S/N} > 1.6$. Following the procedures outlined in Finkelstein et al. (2022), we used the same software in two-image mode to measure the flux in each filter and applied the appropriate aperture correction obtained from simulations. We performed extinction corrections using a Cardelli extinction law with $R_V = 3.1$ for the Milky Way (Cardelli et al. 1989). We then compared the fluxes measured in this catalog to the F160W-selected catalog of Finkelstein et al. (2022) and found the flux measurements to have no systematic offset and minimal scatter. Figure 5 shows the fractional error of the stacked catalog photometry compared to the Finkelstein et al. (2022) photometry as a function of source brightness in the *I* band. The median offset is zero with scatter of roughly 25% for fluxes near 100 nJy, in agreement with the typical error bars for such sources, providing confidence in the fidelity of the stacked catalog. In all subsequent analysis, we defaulted to using measurements from the Finkelstein et al. (2022) catalog for sources detected in both, and we only used photometry from the stacked catalog for five LAEs in our sample unique to it.

After generating the catalog from stacked imaging, we identified all imaging sources within 3″ of the HETDEX detection position as possible LAE counterparts. Since the typical image quality of the HETDEX observations used here has a PSF of $\sim$1.″7, the 3″ annulus served as a generous aperture around the Lyα centroid to encompass all possible counterparts for the detected emission.

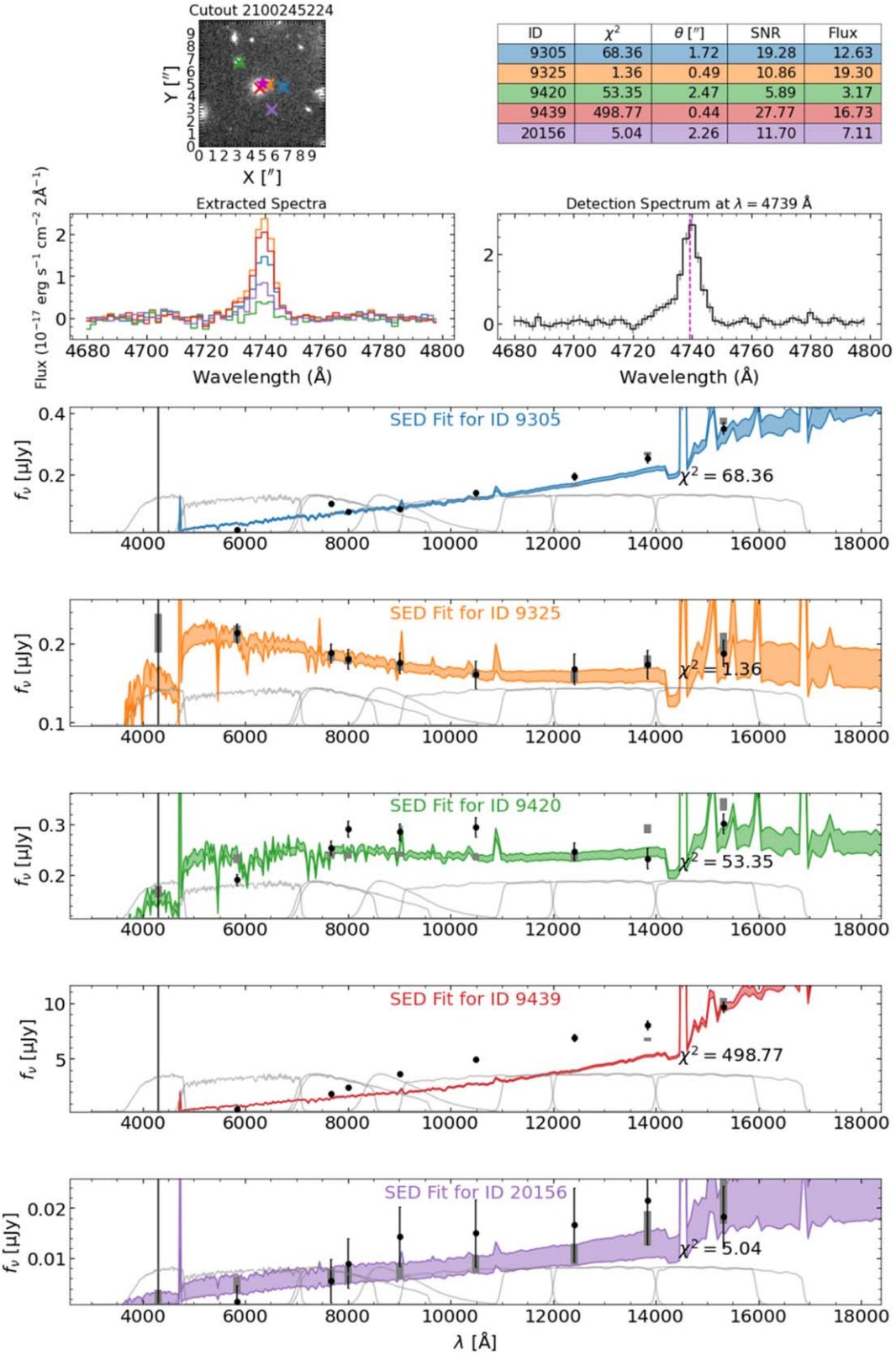
We selected imaging counterparts based on the neighboring sources' angular distances from the detection, significance of emission extracted at the source positions, and goodness of SED fits performed by fixing the redshift assuming an Lyα detection. First, we measured the on-sky angular separation from the detection position to the position of each possible source in the photometric catalog (labeled θ in Figure 6). Then, for each source, we used the HETDEX API script, get_spectrum.py,[15] to perform an aperture-weighted optimized spectral extraction (following Horne 1986) at the source position to obtain a 1D spectrum. We created a Markov Chain Monte Carlo (MCMC) line-fitting code using EMCEE (Foreman-Mackey et al. 2013) to fit a model to the feature to estimate its flux and significance. Our model consisted of two components: a linear trend with slope m and intercept b, which captured any underlying continuum, and a Gaussian with total flux F and standard deviation σ to fit the line profile,

$$f_\lambda = m(\lambda - \lambda_0) + b + \frac{F}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}\right]. \quad (2)$$

In the model, λ_0, the wavelength of the emission line, was allowed to vary by ± 1 pixel (2 Å) from the detection wavelength reported by HETDEX. For each fit, we measured an effective S/N (labeled S/N in Figure 6) by comparing the median value of the line flux to the standard deviation of the line flux for the last 20% of the MCMC sampling chain, which had converged at that stage of sampling. To limit computation time for future counterpart identification steps, we ruled out any counterpart candidates that did not have any indication ($\mathrm{S/N} > 1$) of an emission feature at the pixel corresponding to the detected wavelength. Finally, for those sources with significant emission, we performed SED fitting with BAGPIPES (see Section 3.1 for a full description of this procedure), fixing the redshift as $z_{\mathrm{Ly}\alpha}$. Our simple SED model for counterpart identification included free parameters for stellar mass, metallicity, dust extinction, and star formation history (SFH), and we adopted the Calzetti et al. (1994) dust attenuation law, the Chabrier (2003) initial mass function (IMF), and a delayed-τ SFH. At this stage, we did not include any IRAC fluxes in our fits since those fluxes depend sensitively on deblending, which is unreliable when sources are crowded. Furthermore, between $1.9 < z < 3.5$, there are no strong spectral features at the rest-frame wavelengths probed by IRAC, and redshift-sensitive features such as the 4000 Å break are adequately covered by HST. We then visually inspected the separations, spectral extractions, and SED fits of all candidate counterparts to choose the one most likely to be the detected LAE.

Figure 6 shows an example of our approach. Separate sources are marked with a cross, and the color of the mark corresponds to the color of the table row, spectrum, and SED in the subsequent plots. In this case, the red and orange sources within 0.″5 of the detection position (magenta star) show similar extracted emission-line flux at the detection wavelength. Crucially, the SED fit for the red object poorly matches the data when fixing the redshift as $z_{\mathrm{Ly}\alpha}$, but the orange object has a fit in excellent agreement with its observed SED based on

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



ID	χ^2	θ [″]	SNR	Flux
9305	68.36	1.72	19.28	12.63
9325	1.36	0.49	10.86	19.30
9420	53.35	2.47	5.89	3.17
9439	498.77	0.44	27.77	16.73
20156	5.04	2.26	11.70	7.11

Figure 6. An example of the visualizations used to select imaging counterparts for Lyα detections. The figure shows a 10″ image cutout with the detection positioned marked with a magenta star and sources marked with crosses of various colors, 1D spectra extracted at each source position, the original detection 1D spectrum, and SED fits (with redshift fixed assuming Lyα emission) of all sources with significant measured fluxes. The table in the upper right corner contains the χ^2 value of each SED fit, the separation between each source and the emission-line detection position (labeled θ), and the significance (labeled S/N) and the line flux in erg s^{-1} cm^{-2} of the emission line extracted at the source position. The colors are consistent across all plots and tables, so each source corresponds to a unique color. The SED plots also contain normalized filter response curves as gray lines. In this case, while the red and orange sources have similarly small distances from the detection position and similar line fluxes, the SED fit χ^2 strongly favors the orange source to be an LAE at $z = 2.90$.

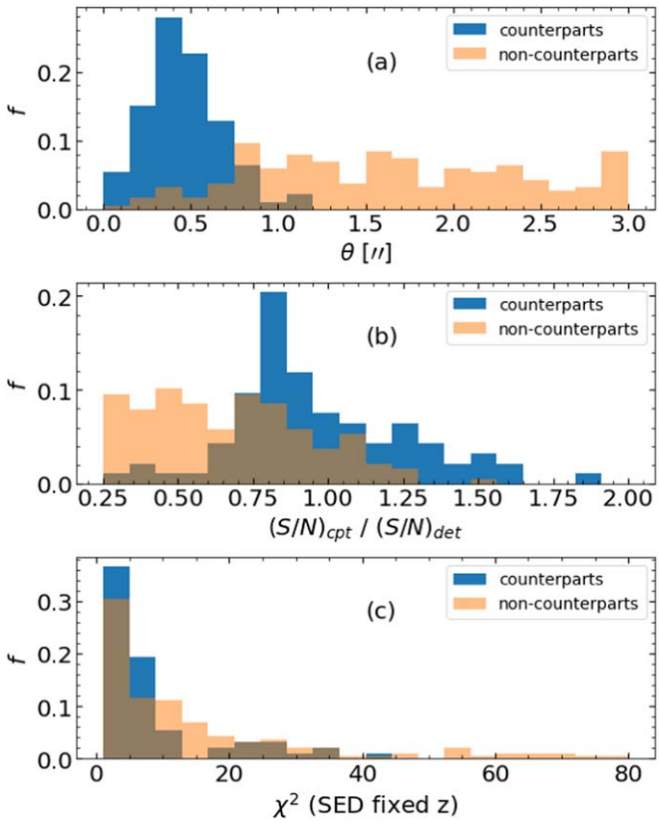


Figure 7. The distributions, expressed as the fraction of objects in a given bin, of (a) separation, (b) position-extracted emission-line S/N (relative to that of the detection), and (c) SED χ^2 assuming $z_{\mathrm{Ly}\alpha}$, for objects identified as the detection imaging counterparts and those that happened to be spatially coincident. Histograms are normalized to the population size. The top panel indicates that finding a counterpart with an imaging separation larger than 1″ from the detection position is exceedingly rare.

the χ^2 statistic. Therefore, in this example case we selected the orange object as the detected LAE at $z = 2.90$. We followed the same process to identify counterparts for the other Lyα lines in our sample.

By studying the distribution of our counterparts in the parameter space of separation, S/N of emission, and SED χ^2, we found no obvious way to select counterparts reliably based on these numbers alone, but we did find favorable regions. Figure 7(a) shows the distribution of separation from the detection positions for sources we identified as LAEs and sources that just happened to be nearby. Clearly, it was exceedingly unlikely that the true counterpart lay farther than 1″ away on-sky. For this reason, we could very reasonably shrink our selection criteria from all sources within 3″ to roughly 1″ without significant loss of LAEs. In terms of emission-line S/N (compared to the measured value of the detection itself), we found that while typically the identified counterparts had stronger emission, the HETDEX PSF caused the extracted flux to not depend sensitively enough on position to clearly identify the counterpart for sources separated by less than 1″. This is clearest in Figure 8(a), which shows that true counterparts and close neighbors show overlap in the S/N–separation plane. Note that the different on-sky centroids for emission-line extraction between the counterparts and the original detection allow for the values of the S/Ns in Figure 7(b) to be greater than unity. Finally, we note that while most of the LAEs in our sample had χ^2 values in good agreement with the $z_{\mathrm{Ly}\alpha}$ hypothesis, many neighboring galaxies also had low χ^2, as shown in Figure 7(c). We attribute the low χ^2 values for noncounterparts to our inclusion of such faint objects, which have large flux errors and are thus easily fit by a wide range of models.

After visually vetting all detections in our sample of 94 Lyα lines, we found six instances of detected emission with no continuum-detected counterpart. Since we could not study the properties of an LAE without photometry, we removed these objects from the final analysis. Furthermore, we removed 16 objects from the sample owing to the following quality concerns. We eliminated the LAE corresponding to HETDEX detection ID 2100245124 (R.A., decl. = 189.°346621, 62.°260662) from our sample because it was the only counterpart with an X-ray detection in the catalog of Xue et al. (2016), indicating that the galaxy hosted an AGN. Since our SED-fitting code did not have an AGN template, we could not reliably report the physical properties of this object. We also eliminated the detection for ID 2100171783, as the counterpart inspection revealed that the Lyα emission line came from two probable LAEs separated by less than 0.″5, meaning that we could not assign flux accurately to each source. Finally, we only analyzed objects detected in the *H* band (F160W) of the HST imaging and at least two bluer bands in order to span the rest-frame 4000 Å break at the sample redshift range. This serves as a crucial feature for constraining galaxy masses and ages with SED fitting (see, e.g., Shapley et al. 2003). These choices limited our final sample size to 72 LAEs in GOODS-N spanning $1.98 < z < 3.48$. For five of these objects, photometry was not present in the catalog of Finkelstein et al. (2022), so we used photometry from the stacked catalog described in Section 2.3. In Appendix B, Figure 21 shows all the HETDEX emission lines for the LAEs in the final sample, and Figure 20 shows HST imaging in F160W (*H* band) for all objects.

3. Analysis

After connecting HETDEX emission-line detections with HST imaging counterparts, we leveraged SED fitting to measure the galaxies' stellar population properties. From the SED fits and emission-line detections, we also inferred the UV slope and Lyα equivalent width.

3.1. SED Fitting with BAGPIPES

We fit all LAEs in our final sample with BAGPIPES (Carnall et al. 2018), a flexible PYTHON code that rapidly generates galaxy model spectra through stellar population synthesis using the 2016 version of the Bruzual & Charlot (2003) stellar spectral libraries. It explores the high-dimensional, multimodal, and degenerate (e.g., age–dust–metallicity) model parameter space using the MULTINEST algorithm (Feroz & Skilling 2013).

Our sample in GOODS-N had photometry across nine HST filters ranging from 0.4 to 1.6 μm, as well as two Spitzer/IRAC channels centered at 3.6 and 4.5 μm. Translating to the rest frames of the objects in our sample at $1.9 < z < 3.5$, these filters probed the UV, optical, and near-infrared (NIR) energy output of our objects.

The filter coverage of our sample of LAEs motivated our choice of SED modeling parameters. Table 1 shows the names and units of the free parameters in our model, as well as the prior probability distributions assumed in our Bayesian

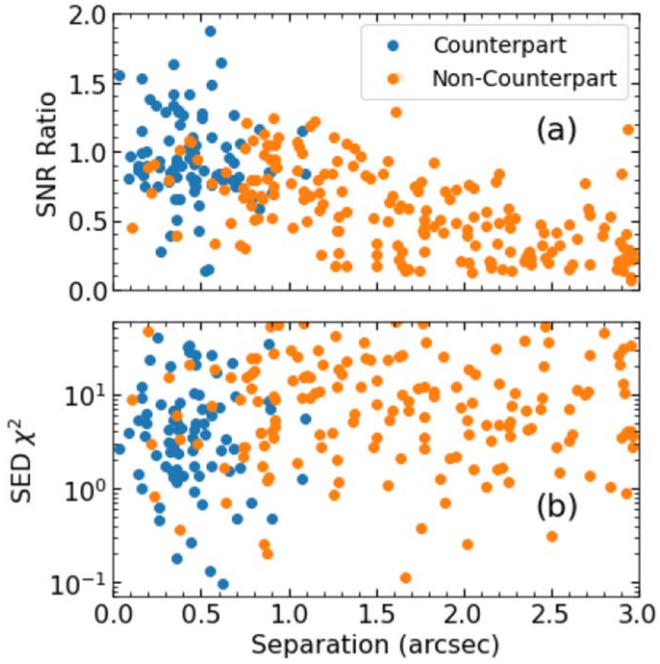


Figure 8. (a) The 2D distribution (in S/N and separation space) for objects identified as the imaging counterparts for emission-line detections and those that happened to be spatially coincident. (b) Same as panel (a), but in SED χ^2 and separation space. These two figures show substantial overlap in these parameter spaces for true LAEs and neighboring sources, motivating the benefits of detailed visual inspection shown in Figure 6.

Table 1
Free Parameters and Their Prior Probability Distributions for SED Fitting

Parameter	Prior	Bounds	Units
t_0	Uniform	0, $T(z)$	Gyr
τ	Uniform	0.3, 10.	Gyr
$M_{\rm form}$	Log Uniform	10^6, 10^{12}	$M_\odot$
Z	Log Uniform	10^{-5}, 2	$Z_\odot$
A_V	Uniform	0, 2	mag
$\log U$	Uniform	−4, −2	⋯

Note. In our galaxy models, the redshift, z, was fixed based on the observed wavelength of Lyα from HETDEX. $T(z)$ refers to the age of the universe at redshift z. Note that we fit the cumulative stellar mass formed, $M_{\rm form}$, from which the stellar mass (excluding remnants) at the object redshift was computed within the BAGPIPES (Carnall et al. 2018) code.

framework. We adopted a delayed-τ SFH, defined as

$$\mathrm{SFR}(t) \propto \begin{cases} (t - t_0)e^{-(t-t_0)/\tau} & t > t_0 \\ 0 & t < t_0 \end{cases}. \quad (3)$$

This flexible SFH allows for star formation to be either rising, peaking, or falling, as opposed to the common exponentially declining model that only allows for falling SFRs over time. For example, Lee et al. (2010) found that SED fitting that adopted rising SFHs matched the stellar masses and SFRs from semianalytic models for galaxies at $3 < z < 6$ better than exponentially declining models, while Papovich et al. (2011) found similar results favoring rising SFHs for real galaxies at $z = 4$–7. We fit the e-folding scale of the SFH, τ, the age of the universe at the onset of star formation, t_0, the stellar mass formed, $M_{\rm form}$, the global metallicity, Z, the dust extinction in the V band, A_V, and the ionization parameter, $\log U$, defined as the log of the ratio of the number densities of ionizing photons and hydrogen atoms. Though we fit the total stellar mass formed by a galaxy, $M_{\rm form}$, we report its stellar mass at the redshift of observation excluding remnants, and we denote that stellar mass $M_\star$. We note that some of the parameters (namely, Z and U) are not expected to be well constrained by our photometric data. Nonetheless, we allow them to vary within our imposed priors such that the uncertainties in the other parameters include the uncertainties in these parameters. We adopted the Calzetti et al. (1994) dust attenuation law for SFGs and the Chabrier (2003) IMF.

All 11 filters were not necessarily included for every galaxy SED fit in our sample. For example, due to the large PSF of the IRAC imager, modeling sources in crowded fields of view and deblending the flux contribution of each source is crucial to accurately measuring the NIR fluxes of our LAEs. Although the catalog we used performed deblended photometric modeling with the IRAC PSF, this process can fail in crowded regions. We thus visually inspected all IRAC residual maps for objects in our sample and removed the IRAC fluxes from our SED fitting if there were obvious problems in the deblending procedure. For the five objects not present in the catalog of Finkelstein et al. (2022), we did not have IRAC measurements. Furthermore, because the purpose of our analysis was to study the SED-derived properties of our LAEs in relation to their Lyα emission, we did not want BAGPIPES's modeling of Lyα emission or the IGM attenuation to bias our results. For this reason, we masked out all filters whose bandpass extended blueward of the observed Lyα line; thus, the B band (F435W) and sometimes the V band (F606W) were excluded, depending on redshift.

Figure 9 shows an example BAGPIPES SED fit for an LAE in our sample. We plotted the 1σ spread on the model photometry as rectangles, as well as the 1σ spread on the underlying model spectrum computed by evaluating the 16th and 84th percentiles of the posterior models. In this example, the fit did an excellent job matching salient features like the rest-frame 4000 Å break and nebular emission in the rest-frame optical region. We estimated galaxy properties using the posterior distributions for all free parameters explored by BAGPIPES. Figure 10 shows an example "corner" plot (produced via Foreman-Mackey 2016), where all free parameters are plotted against each other for easy assessment of constraints and correlations. Stellar mass, time since the onset of star formation, and dust extinction were constrained well, while metallicity, τ, and ionization parameter were not well constrained by our broadband photometry data. Figure 22 in Appendix B shows the SED fits for all 72 LAEs in our final sample.

3.2. Measuring W_λ(Lyα) and β

Emission-line strengths can be represented by the parameter equivalent width (W_λ), which represents the width of a rectangle drawn to the same height as the continuum needed for the rectangular area to match the area under the emission line. To estimate the equivalent width of HETDEX Lyα detections, we used the measured line flux and error from the internal HDR 2.1.3 catalog computed by optimally extracting flux from all fibers within a 3.″5-radius circular aperture (roughly 15–20 individual fiber spectra) contributing to the emission-line detection (following Horne 1986), weighted by the PSF of a point source. We approximated the continuum flux

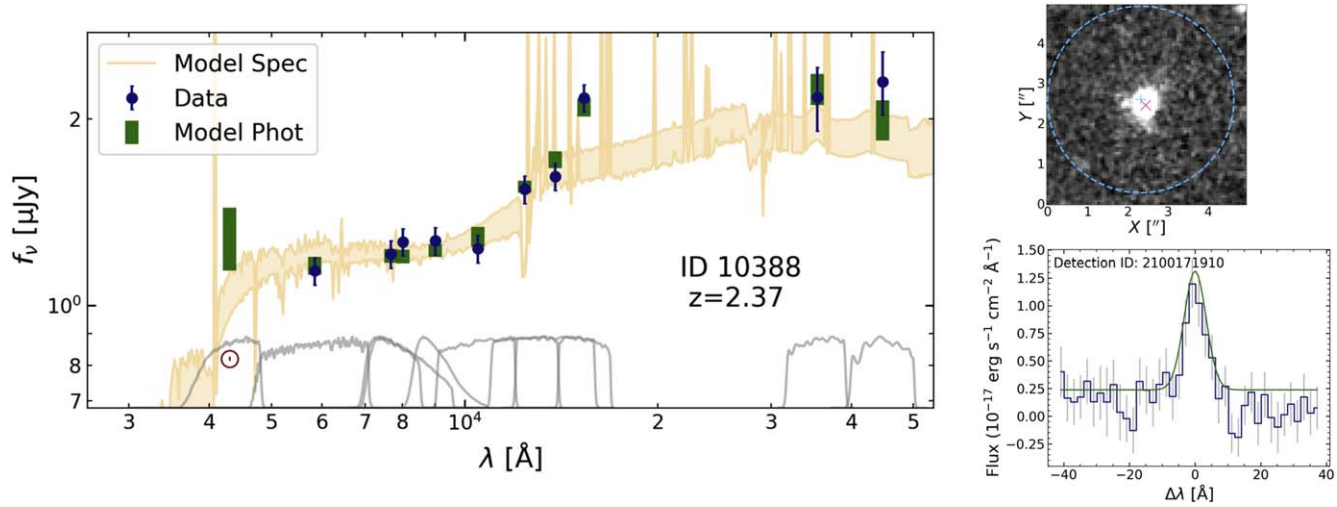


Figure 9. Left: an example fit to the SED of an LAE in the sample. Data are shown by blue circles, and the 68% spread for the posterior model photometry and spectrum are shown by the green rectangles and orange shaded lines, respectively. A maroon open circle indicates the measured flux for a filter masked during SED fitting. In this case, the *B* band was masked out since it includes the Lyα emission line. Our imaging data constrain the 4000 Å break and rest-optical colors. Top right: 5″ square image cutout for the source in the F160W HST filter. The pink cross indicates the source position, and the blue plus sign and dashed blue circle indicate the detection position and the FWHM of the HETDEX fiber PSF. Bottom right: 1D extracted spectrum for this source, centered on an 80 Å window around the Lyα emission line. The solid green line indicates the HETDEX Gaussian model fit to the data.

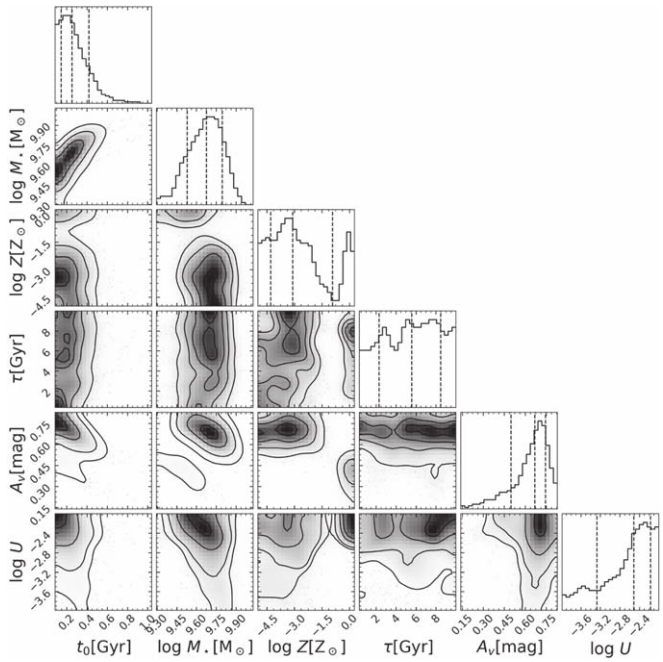


Figure 10. A "corner" plot of the fit in Figure 9 for object ID 10388. The 1D histograms are shown on the diagonal for the posterior distribution of each free parameter in our model (see Table 1). The 2D histograms show the correlations of all parameters with one another, where contour lines are drawn for each σ level. With our broadband photometry, we constrained ages, masses, and dust extinctions well.

density using the BAGPIPES-sampled model spectra from the SED fit. We took the continuum flux density to be the median value of all 500 sampled spectra averaged between 1250 and 1300 Å in the given object's rest frame, and we computed the 1σ error using half the spread between the 16th and 84th percentiles of those values. This method allowed us to take advantage of complex computations performed by BAGPIPES to get a statistically representative estimate of the continuum flux density instead of using a coarse approximation based on the flux in one of our photometric bands. We evaluated the Lyα flux and the continuum flux density in the observer frame and translated to the galaxy rest frame by dividing by a factor of $(1+z)$ using the detected wavelength of Lyα,

$$W_\lambda(\mathrm{Ly}\alpha) = \frac{F_{\mathrm{Ly}\alpha}}{f_\lambda}(1+z)^{-1}. \quad (4)$$

We measured β, the UV continuum slope (uncorrected for dust), using the model spectra for galaxies in our sample following the method described in Finkelstein et al. (2012). We masked the stellar and interstellar absorption features in the rest-frame UV using the windows provided by Calzetti et al. (1994), and we fit a linear model to the spectrum in log space ($\log f_\lambda = \beta \log \lambda + C$) using *polyfit* from the PYTHON package NUMPY (Harris et al. 2020). We determined 1σ uncertainties on β for each object by measuring the distribution of values fitted to 500 spectral models sampled from the posterior by BAGPIPES.

4. Results

We measured various physical properties of objects in our LAE sample using the posterior distributions returned by BAGPIPES' exploration of the parameter space. We took the 16th and 84th percentiles of the posterior distributions to represent the error bars on physical properties. Examples of such measurements are shown in Figure 10 for a representative LAE in our sample.

4.1. SED-derived Properties

Figure 11 shows the 1D distributions of posterior median values of stellar mass ($M_\star$), SFR, specific SFR (sSFR), dust extinction (A_V), mass-weighted age, and UV slope (β) for all objects in our final LAE sample. We found the median stellar mass of our HETDEX LAEs to be $0.8^{+2.9}_{-0.5} \times 10^9\ M_\odot$. This stellar mass value lies near the median masses of LAEs selected in NB imaging surveys covering redshifts similar to this study (e.g., Guaita et al. 2011; Gawiser et al. 2007; Vargas et al. 2014; Kusakabe et al. 2018; Santos et al. 2020) and well below typical

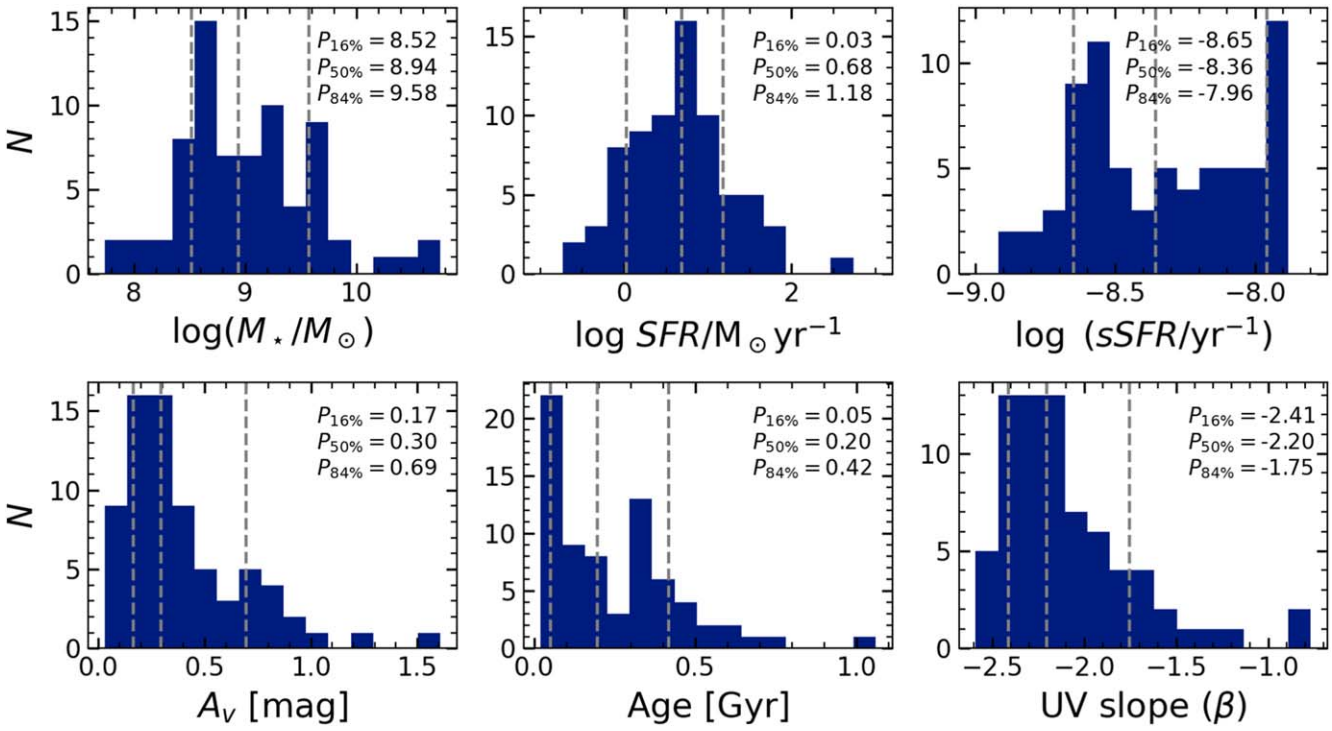


Figure 11. The distributions of posterior median values for (clockwise from top left) stellar mass, SFR, sSFR, dust extinction (in *V*-band mag), mass-weighted age, and UV slope for all objects in the sample. The 16th, 50th, and 84th percentiles are indicated by vertical dashed gray lines, and their values are indicated with text in the same units as the *x*-axis labels. The LAEs in our sample exhibit average properties similar to other LAE samples compiled at comparable redshifts using NB selection.

masses of Lyman-break-selected objects (e.g., Shapley et al. 2003; Papovich et al. 2001; Trainor et al. 2019), which often have minimum masses an order of magnitude larger owing to the depth of the broadband imaging used in their selection.

We used our SED-fitting procedure to obtain the attenuation in the *V* band of starlight due to dust for galaxies in the sample, and we obtained a median value of $A_V = 0.3^{+0.4}_{-0.1}$ mag. The presence of dust has been measured in many other samples of LAEs with values of A_V or $E(B-V)$ often falling within a factor of two of this study (e.g., Guaita et al. 2011; Finkelstein et al. 2009; Hathi et al. 2016; Kusakabe et al. 2018; Matthee et al. 2021).

Similar to dust reddening, our LAE sample has similar ages and SFRs to LAE samples in the literature compiled using NB or continuum selection methods. Our SED-derived mass-weighted ages, typically spanning 0.05–0.5 Gyr, broadly agree with the NB samples of Acquaviva et al. (2011), Finkelstein et al. (2009), Gawiser et al. (2007), and Vargas et al. (2014). Our median SFR, $4.8^{+10.4}_{-3.8} M_\odot\ yr^{-1}$, falls near values reported by Gawiser et al. (2007), Hathi et al. (2016), and Kusakabe et al. (2018) but falls 1 dex above the median SFR for LAEs found in the MUSE HUDF Survey (Feltre et al. 2020). This discrepancy does not surprise us since the MUSE HUDF LAE sample had a median mass roughly 0.5 dex lower than this study, and their sample spanned $2.9 < z < 4.6$, probing an era of lower star formation activity in the universe than the one studied here (see Madau & Dickinson 2014).

Our model included stellar metallicity and ionization parameter as free parameters, but our broadband photometric data could not constrain those values precisely (see Figure 10), since reliable estimates typically require sensitive emission-line diagnostics (e.g., Reddy et al. 2022), which were coarsely probed at best by our filter set. For this reason, we do not present or discuss our galaxies' metallicities or ISM ionization conditions, but we note that by letting these parameters vary, our posterior constraints on all other parameters include the uncertainties in these quantities.

4.2. $W_\lambda(Ly\alpha)$ Distribution

The equivalent width distribution of LAEs has been modeled by various authors as exponential with the form given by the following equation (e.g., Gronwall et al. 2007; Guaita et al. 2010; Wold et al. 2014; Jung et al. 2018):

$$\frac{dN}{dW_\lambda} \propto e^{-W_\lambda/W_0}. \tag{5}$$

We show our sample's rest-frame $W_\lambda(\mathrm{Ly}\alpha)$ distribution in Figure 12 with an *e*-folding scale $W_0 = 100$ Å drawn for comparison. We cannot measure the underlying distribution for LAEs from our sample since we have not measured the completeness as a function of equivalent width (which is complex owing to our method of sample creation, and not crucial for our study of stellar population properties). Various other studies have precisely measured the Lyα equivalent width distribution, such as Gronwall et al. (2007), who found an *e*-folding scale of 76^{+11}_{-8} Å for a deep, NB-selected LAE sample at $z = 3.1$, Guaita et al. (2010), who measured $W_0 = 50 \pm 7$ Å for an NB sample at $z = 2.1$, and recently Santos et al. (2020), who measured $W_0 = 129 \pm 11$ Å for the full SC4K sample at $2 < z < 6$. We plot some of these measured distributions in Figure 12 for comparison. It is apparent that our sample becomes increasingly incomplete at EW $\lesssim 50$ Å, due to a combination of the HETDEX flux limit, the emission-line identification process, and our counterpart selection process.

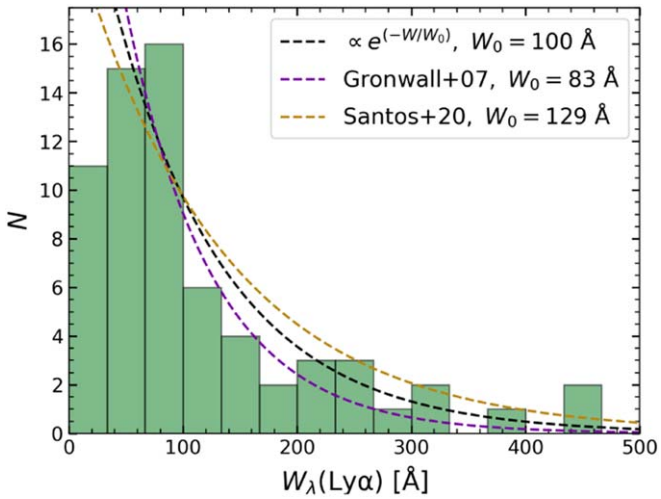


Figure 12. The equivalent width distribution of LAEs in the sample. An exponential distribution with $W_0 = 100$ Å is drawn in red for comparison, as well as models fit by Gronwall et al. (2007) and Santos et al. (2020). Our data favor models with larger values of W_0 to best match the number of high-EW sources.

4.3. Correlations between W_λ(Lyα) and Galaxy Properties

We combined our SED-derived galaxy properties with the W_λ(Lyα) measurements described above in order to assess correlations between Lyα emission and global galaxy properties. We used W_λ(Lyα) as a proxy for the fraction of photons emitted as Lyα as opposed to $L_{\mathrm{Ly}\alpha}$, for example, because the equivalent width more closely probes the physics governing Lyα escape, whereas the flux also includes physics related to the Lyα production rate. Figure 13 shows $M_\star$, sSFR, SFR, dust extinction (A_V), mass-weighted stellar population age, and UV slope (β) plotted against each galaxy's W_λ(Lyα) measurement. In the figure, error bars denote the 16th to 84th percentile range, and we indicate Pearson's linear correlation coefficient, r_p, and its significance (p-value) with text.

Stellar mass and SFR both correlate strongly with W_λ(Lyα), with low-mass, low-SFR systems achieving larger W_λ(Lyα) than higher-mass systems. The correlation with mass has been established in the literature from studies of a wide variety of galaxies such as LBGs, oELGs, and LAEs. It was noticed early by Ando et al. (2006) and measured recently by many works such as Du et al. (2018), Marchi et al. (2019), Oyarzún et al. (2017), and Shimakawa et al. (2017). Specifically, Weiss et al. (2021) found a negative correlation between the Lyα escape fraction, $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$, and stellar mass using data from the HETDEX survey. Additionally, Khostovan et al. (2021) found an intrinsic, negative correlation between Hα equivalent width and galaxy stellar mass from an NB survey at $z \sim 5$. While the lack of low-EW, low-mass systems can be driven by selection incompleteness, we should be complete to high-mass, high-EW systems, yet these are seemingly rare.

Notably, our results show no significant anticorrelation between W_λ(Lyα) and dust extinction (A_V), whereas numerous other studies of Lyα emission measured a clear relationship that indicates that dust hinders the ability of the Lyα photon to escape the galaxy. For example, Shapley et al. (2003), Guaita et al. (2011), Du et al. (2018), Hathi et al. (2016), Huang et al. (2021), Marchi et al. (2019), Matthee et al. (2016), Reddy et al. (2022), Trainor et al. (2019), and Weiss et al. (2021) all showed that dustier galaxies exhibit weaker Lyα emission measured as W_λ(Lyα) or have smaller $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha}$. However, the lack of a significant anticorrelation may be due to our limited sample size and small dynamic range in dust attenuation. Moreover, objects with significant amounts of dust that suppress their Lyα fluxes would not become members of our science sample in the first place. The majority of our sample has $A_V < 0.3$. We do observe multiple galaxies with $A_V > 0.5$, and interestingly these do not all have low W_λ(Lyα), implying that Lyα can escape even from modestly dusty galaxies, which could indicate enhanced escape due to outflows (e.g., Steidel et al. 2010; Erb et al. 2012) or a multiphase ISM (e.g., Finkelstein et al. 2009; Neufeld 1991).

Our Pearson correlation coefficient suggests a moderate correlation between W_λ(Lyα) and galaxy stellar-mass-weighted age ($r_p = 0.32$) in the sense that older galaxies exhibit larger W_λ(Lyα). Marchi et al. (2019) found a similar result, obtaining a Spearman rank correlation coefficient of 0.40. This contrasts with Pentericci et al. (2009) and Pentericci et al. (2010), who found no strong dependence of Lyα equivalent width on age for LAEs and LBGs, as well as Reddy et al. (2022), who found a weak negative correlation between the two measurements for SFGs in the same redshift range probed by this study.

Finally, a moderate negative correlation exists between sSFR and W_λ(Lyα), though the large error bars for our measurements of sSFR weaken the reliability of the correlation. For comparison, Hathi et al. (2016) found no significant correlation between the two properties for a sample including Lyα in absorption and emission.

We also plot SFR against $M_\star$ for all objects in our sample in Figure 14 to see how our galaxies compare to other objects at similar redshift in relation to the star-forming main sequence (SFMS). We include the best-fit line found by Sanders et al. (2018) for SFGs in the MOSDEF survey at $z \sim 2.3$. Note that masses derived for that study used the Chabrier (2003) IMF and Calzetti et al. (2000) dust curve but stellar population synthesis models from Conroy et al. (2009). We also use a color bar to show the value of W_λ(Lyα) for each galaxy. The position of LAEs on the SFMS remains somewhat controversial. Studies such as Vargas et al. (2014), Finkelstein et al. (2015), Hagen et al. (2016), and Santos et al. (2020) found LAEs to lie above the relation, while other studies have interpreted them as lying directly on the low-mass end of the relation (e.g., Kusakabe et al. 2018). Figure 14 shows that the LAEs in our sample lie largely on the SFMS, though a significant fraction lie below the relation of Sanders et al. (2018) for $M_\star < 10^9\,M_\odot$.

In Appendix A, we explore the model dependence of our measured galaxy properties, since the parameters derived from SED fitting can be systematically different using different models (see Conroy 2013). We conclude that our results, including the median physical properties and the correlations with W_λ(Lyα), are not driven by our specific choice of model.

5. Discussion

5.1. Are HETDEX LAEs Special?

The question "What is a HETDEX LAE?" holds particular importance for astronomers studying galaxy science with this survey. A vast sample of HETDEX LAEs is upcoming, and samples of such objects selected by emission-line detection from a blind spectroscopic survey remain rare in the literature (with the exception of the HETDEX Pilot Survey (Adams et al. 2011;

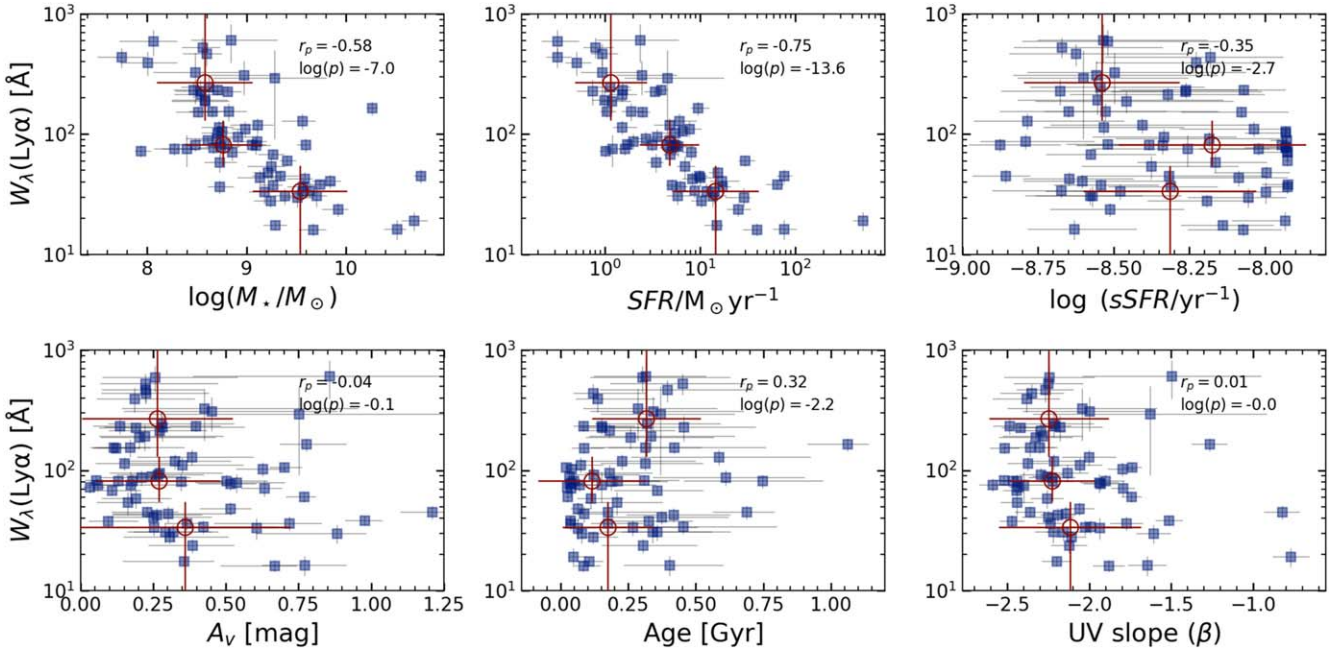


Figure 13. The relationship between W_λ(Lyα) and (clockwise from top left) stellar mass, SFR, sSFR, dust extinction (A_V), mass-weighted age, and UV slope (β) for all objects in the sample. Gray lines indicate the $\pm 1\sigma$ error bars on physical properties and equivalent width. Open red circles show medians and standard deviations for properties in equivalent width bins having equal numbers of objects. Pearson correlation coefficients, r_p, and p-values are indicated for each plot. Stellar mass and SFR exhibit the strongest correlations with W_λ(Lyα), while age and sSFR correlate moderately. Surprisingly, no strong correlation exists with dust extinction.

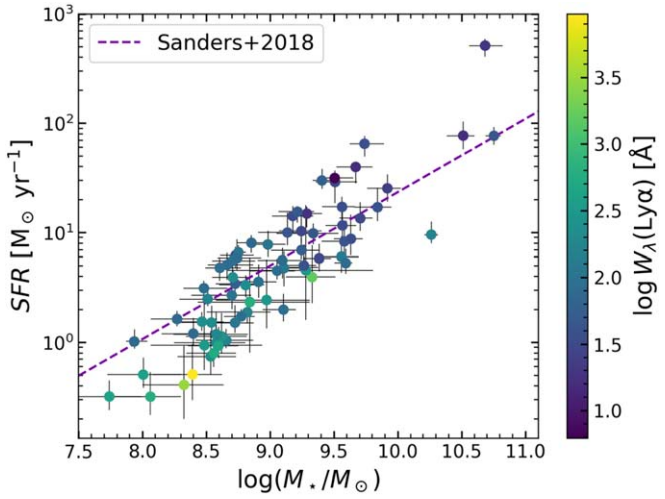


Figure 14. The stellar mass–SFR correlation for LAEs in our sample. The trend fit by Sanders et al. (2018) for $z \sim 2.3$ SFGs is drawn in dashed purple for comparison. LAEs in our sample largely fall on the SFMS, though the lowest-mass sources ($M_\star < 10^9\,M_\odot$) tend to fall below the relation.

Blanc et al. 2011), which probed a smaller area to a brighter flux limit, and MUSE surveys, which probe much smaller areas to fainter flux limits with only a small overlap in redshift with HETDEX). Characterizing any idiosyncrasies in the HETDEX LAE population will put these objects in context relative to the numerous LAEs found by previous studies, and it will aid the interpretation of future blind spectroscopic surveys for these objects in the EoR.

As described above, in our $f_{\mathrm{Ly}\alpha} \gtrsim 6 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ flux-limited sample (Gebhardt et al. 2021), the median galaxy mass of $0.8^{+2.9}_{-0.5} \times 10^9\,M_\odot$ lies very close to many LAE samples selected through NB imaging. For example, Gawiser et al. (2007) found a median mass of $1^{+0.6}_{-0.4} \times 10^9\,M_\odot$ with a flux limit of $1.5 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ at $z = 3.1$. Guaita et al. (2011) pushed to an even lower median mass of $\sim 4 \times 10^8\,M_\odot$, roughly a factor of two less massive than this sample's median, with a flux limit of $2.0 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ at $z = 2.1$. The MUSE HUDF went even deeper, finding sources at $z > 3$ with Lyα line fluxes as small as $\sim 2 \times 10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ and obtaining a median sample mass of $\sim 2.5 \times 10^8\,M_\odot$. The sample of Santos et al. (2020) was limited by medium-band line flux limits spanning $(3.0\text{–}4.8) \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ over $2 < z < 6$ (Sobral et al. 2018) and measured a median LAE mass of $2 \times 10^9\,M_\odot$, consistent with this study. Of course, the mass range probed by HETDEX falls far below samples selected using the Lyman/Lyα break (e.g., the lowest mass probed by Papovich et al. 2001 was $10^{10}\,M_\odot$ at $2.0 < z < 3.5$). Thus, the HETDEX flux limit explores an LAE mass range comparable to NB surveys, yet slightly more massive than the deepest NB and spectroscopic surveys. At the expense of sensitivity, the HETDEX survey can find fairly low mass LAEs over a large continuous redshift interval, reducing the effects of cosmic variance compared to NB observations.

As mentioned in Section 4.1, the LAEs in this sample do not stand out from NB samples at similar redshift in terms of age, SFR, and dust extinction. Thus, we can conclude that the HETDEX survey selects a typical LAE having properties consistent with the general NB-selected population, but it may have slightly higher stellar mass based on the line flux limit of the survey.

Nonetheless, our sample may stand out in its relation to the SFR–$M_\star$ relation shown in Figure 14. Compared to the relation measured in Sanders et al. (2018), LAEs in the sample with $M_\star \lesssim 10^9\,M_\odot$ appear to lie below the trend. This contrasts markedly with the work of Hagen et al. (2016), who compiled their sample using the HETDEX Pilot survey (Adams et al. 2011; Blanc et al. 2011) and found their LAEs to lie above the SFMS. Interestingly, the LAEs lying below the SFMS in

Figure 14 have very high W_λ(Lyα), which correlates with lower $M_\star$ and SFR in Figure 13. We are not surprised that the lowest-mass systems in our sample have the highest values of W_λ(Lyα) given the negative correlation with $M_\star$ and the fact that low-mass objects need large W_λ(Lyα) to be detected by HETDEX, but their position below the SFMS is peculiar. It could be related to the weak negative correlation we found between W_λ(Lyα) and sSFR, or could simply be an artifact of our small sample size. This motivates further study of the positions of LAEs on the SFMS with larger samples.

5.2. Which Properties Drive Lyα Emission?

While the size of the sample analyzed in this study is small, we were still able to extract important information linking galaxy stellar population properties to Lyα emission strength. As the number of LAEs detected by HETDEX grows in fields with rich photometric data, such as the Spitzer-HETDEX Exploratory Large-Area Survey (SHELA; Papovich et al. 2016), the number of LAEs with measured galaxy properties will grow by many orders of magnitude. This will provide a trove of useful data for explaining why some galaxies shine brightly in Lyα while others do not, as well as exploring the effects of galaxy environment on Lyα emission.

We found a significant, strong negative correlation between W_λ(Lyα) and stellar mass in our sample (see the top left panel of Figure 13). This trend is often theoretically attributed to low-mass SFGs having less neutral gas to resonantly scatter the Lyα photon (as well as less dust), leading to a shorter total path length to exit the galaxy without absorption by dust (see Ando et al. 2006). In this sample, W_λ(Lyα) also negatively correlated (even more strongly) with SFR, and the fact that stellar mass and SFR correlate strongly with each other complicates the interpretation of this result. Weiss et al. (2021) addressed this issue by binning their sample of [O III]-emitting galaxies with Lyα line flux measurements from HETDEX according to stellar mass and SFR. They found mass to better predict $f_{\rm esc}^{\rm Ly\alpha}$ at fixed SFR than SFR did at fixed mass.

Fascinatingly, we did not find even a weak correlation between dust extinction and W_λ(Lyα). This seems surprising given that many authors have noted such a correlation and that the theoretical explanation is inarguable: resonantly scattered Lyα photons can get absorbed readily in the presence of even a small amount of dust. A partial explanation for our sample's behavior with A_V could be that it consists of systems exhibiting strong Lyα emission, not absorption. For example, Reddy et al. (2022) studied systems with Lyα in net absorption or emission and found a strong correlation between W_λ(Lyα) and $E(B-V)$. If our sample contained objects with negative W_λ(Lyα), perhaps those objects would reveal the correlation. Nevertheless, other studies of only emitters (W_λ(Lyα) >0) have also noted a trend with dust extinction, such as Marchi et al. (2019), though a close examination of their Figure 7 shows that the negative correlation is largely driven by weak emitters with W_λ(Lyα) <10 Å. Our small dynamic range in W_λ(Lyα) may obfuscate a correlation with dust extinction. This interpretation may also be complicated by the Lyα photon's ability to escape the galaxy even in the presence of large amounts of dust. Given a clumpy ISM geometry, clumps of gas and dust can act as mirrors to Lyα photons, which "bounce" off the surfaces of these clumps through resonant scattering by neutral gas, while continuum photons pass through and thus experience extinction. Gronke et al. (2016) found that simulated Lyα emission lines agreed well with observations for models with clumpy ISM geometries, and Finkelstein et al. (2009) found that clumpy ISM models better fit the SEDs of over half their NB-selected sample of LAEs at $z\sim4.5$. Vargas et al. (2014) also found that their sample of 20 NB-selected LAEs at $z=2.1$ favored clumpy ISM models.

Lastly, we found a moderate correlation between W_λ(Lyα) and galaxy mass-weighted age. The strength of Lyα emission depends on both its production through recombination in HII regions and its escape through channels in the ISM with low neutral gas covering fractions, so the interplay between these processes determines W_λ(Lyα). As noted by Marchi et al. (2019), who obtained a similar result, the trend with age could arise from older systems having experienced intense star formation in their past, where stellar winds and radiation cleared out neutral gas and dust, leaving channels for Lyα escape. Through ongoing star formation or recent bursts, these objects can still produce Lyα photons, and the ISM conditions favor their escape. For the youngest galaxies, even though the most massive, ionizing photon-producing stars are present, it is possible that a significant amount of dust and neutral gas has yet to be swept away, hindering the escape of Lyα.

6. Predicting Lyα Emission in the Epoch of Reionization

Using our knowledge of Lyα emission from HETDEX galaxies situated in an ionized IGM, we can attempt to predict the intrinsic emission strength of LAEs at $z>7$, an era where starlight from galaxies was still actively reionizing the universe.

6.1. An LAE Sample in the Epoch of Reionization

Our sample at $1.9<z<3.5$ provides a view of Lyα emission unobscured by a significant IGM neutral fraction. By creating a predictive model that connects global galaxy properties to their intrinsic W_λ(Lyα) in this pristine era, we can apply it to LAEs in the EoR to derive their expected intrinsic W_λ(Lyα), then attributing any deficiency of Lyα emission from objects in the EoR to an increasing neutral fraction. This does require the assumption that the production and escape of Lyα photons do not evolve with redshift for fixed galaxy properties, which will require further testing. As a pilot attempt here, we took advantage of the sample of $z>7$ LAEs that Jung et al. (2020) found in GOODS-N to test our ability to predict Lyα emission from EoR galaxies.

Using a deep, spectroscopic survey conducted with Keck/MOSFIRE, Jung et al. (2020) found 10 $>4\sigma$ Lyα detections at $z>7$ among 72 high-z candidate galaxies. Such objects likely reside in ionized bubbles of the IGM, allowing the Lyα photons to redshift away from the resonant frequency, therefore lowering the absorption cross section with neutral hydrogen. These emitters thus serve as direct tests of our understanding of the galaxy properties that modulate Lyα emission strength from the ISM/CGM.

Because the photometric catalog for the GOODS-N field contains the LAEs discovered by Jung et al. (2020), we performed the same SED analysis detailed in Section 3.1 for those objects. We again masked all photometric bands including and blueward of Lyα given the object's spectroscopic redshift. For most of the $z>7$ LAEs, this left three HST filters and both Spitzer/IRAC channels. We again used BAGPIPES to estimate the galaxy properties, adopting our fiducial model (delayed-τ SFH, Calzetti et al. 1994 dust law). Figure 15 shows an example fit for an object at $z=7.51$.

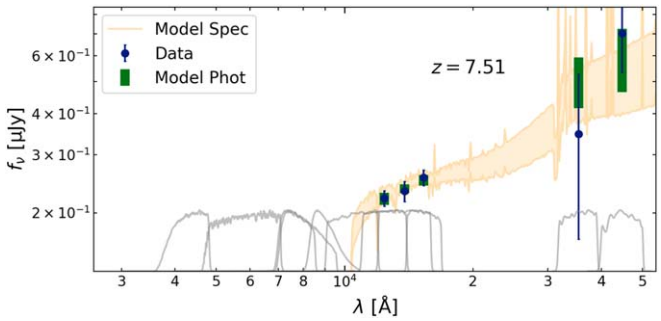


Figure 15. An example BAGPIPES SED fit for LAE ID z7_GND_42912 at $z = 7.51$ detected by Jung et al. (2020). For scaling purposes, we do not show the upper limits for nondetections in the HST bands blueward of the Lyα break. From our photometric data, we constrained the stellar population properties of 10 LAEs in the EoR, allowing us to predict their intrinsic Lyα emission using our HETDEX sample.

6.2. A Predictive Model for W_λ(Lyα)

To predict the Lyα equivalent widths of the $z > 7$ sample, we chose several properties that strongly impact the emergent Lyα emission from galaxies: stellar mass, dust extinction, and SFR. As discussed above, stellar mass may determine the amount of neutral hydrogen gas (and thus dust) in the galaxy and the total path length needed to escape. In the presence of dust, Lyα photons may terminate their resonant scattering process through absorption by a dust grain following reemission at longer wavelengths, limiting likelihood of escape. Finally, the global SFR impacts the production of UV photons that can create Lyα through recombination, and feedback from star formation may impact the structure of the ISM itself, creating ionized channels for escape.

Using the posterior distributions sampled by BAGPIPES, we matched each $z > 7$ emitter to LAEs in the HETDEX sample based on SED-derived properties. To do this, we calculated the "separation" in the log mass–SFR–dust attenuation parameter space from the EoR LAEs to each LAE in the HETDEX sample. For the separation calculation, we divided each parameter value by the full range of values in the sample to normalize the parameter space. For example, for log stellar mass, an object in the HETDEX sample with log mass halfway between the sample minimum and maximum would have a value of 0.5, so the difference between 0.5 and the EoR LAE log stellar mass scaled the same way would become input to the Euclidean distance formula. We then ranked the HETDEX LAEs by separation in parameter space and constructed the prediction using the $N = 3$, 5, and 7 closest neighbors. We computed the posterior W_λ(Lyα) distribution by co-adding Gaussian distributions with mean and standard deviation set by the W_λ(Lyα) measurements and error bars in our sample. To give more importance to those LAEs that closely resembled the EoR galaxy, we weighted each Gaussian distribution by the inverse of its squared distance in parameter space from the EoR galaxy when co-adding to obtain the final prediction. The predicted W_λ(Lyα) distributions are normalized such that the integral over all equivalent widths equals unity.

Figure 16 shows our predicted W_λ(Lyα) distributions for LAEs in the Jung et al. (2020) sample with Lyα S/N > 4. We show predictions using three different values of N, the number of nearest neighbors in parameter space, to reveal any stochasticity in the prediction. The measured Lyα equivalent widths from Jung et al. (2020) are indicated by vertical dashed lines, with 1σ error intervals shaded gray. Importantly, we only expect our predictions to match the observed equivalent widths of EoR LAEs if they exist in ionized bubbles. If the EoR LAEs instead exist in regions of the IGM with significant neutral fractions, we expect to overpredict the Lyα emission. On the other hand, an underprediction of the Lyα emission from an EoR object would imply that our sample size is too small to account for the diversity in physical properties of the LAE population.

In Figure 17, we plot the predicted versus observed equivalent widths with a one-to-one line drawn to facilitate comparison. Each object's predicted value and error were calculated as the first moment and square root of the second moment of the $N = 5$ curves in Figure 16, respectively. In 5 out of 10 cases (ID z7_GND_18626, z7_GND_44088, z7_GND_42912, z7_GND_22233, and z7_GND_39781) the 1σ interval of our W_λ(Lyα) predictions overlapped with the 1σ interval of the observational measurement, indicating moderate agreement. For strong emitters (observed W_λ(Lyα) > 20 Å), our prediction overlapped with observation five out of eight times. Furthermore, two strong emitters (z7_GND_42912 and z7_GND_16863), postulated by Jung et al. (2020) to inhabit ionized bubbles, had observed equivalent widths greater than or equal to the majority of our predicted W_λ(Lyα) distributions, as one might expect for sources with little IGM attenuation.

It is not surprising that our model failed to predict weak Lyα emission accurately. First, our model predicts Lyα EWs in the absence of IGM absorption; thus, an underprediction could imply significant absorption of Lyα photons by neutral hydrogen in the IGM. Second, as our sample by construction contains far more strong emitters than weak ones (see Figure 12), this could presently bias us toward an overprediction of Lyα emission strength drastically. We note that we underpredicted the emission from ID z7_GND_34204 (indicated by an arrow in Figure 17), which could be attributed to the dearth of objects in our sample with very high equivalent widths to match with that object's value, $\sim$280 Å.

ID z7_GND_42912 offers a good example of how challenging predicting Lyα emission can be. As N increases, the peak of the predicted distribution shifts from agreeing well with the observation to underpredicting it. It is clear that our sample is currently too small to fully span the parameter space in both W_λ(Lyα) and physical properties. Future analyses with much larger samples made possible by HETDEX should be able to better capture the mean trends and variance in galaxy parameters that determine Lyα emission strength.

Some of the predictions in Figure 16 bode well for constraining the expected W_λ(Lyα) given a suite of galaxy properties measured from broadband SED fitting. With larger samples that suffer less from the inherent idiosyncratic behavior of Lyα emission (e.g., its dependence on the observer's line of sight), a rigorous, statistical understanding of the properties that drive that emission will arise, unlocking the potential of LAEs to probe cosmic reionization. We further note that, with larger samples, machine learning (ML) may prove an invaluable tool in making the nuanced connection between global galaxy properties and Lyα emission strength, as the problem requires a regression analysis well suited for ML techniques.

7. Summary

We used SED fitting to study the properties of a sample of LAEs from the HETDEX survey in GOODS-N to better understand the phenomenology behind Lyα emission and ultimately leverage these beacons of light in the distant universe as probes of cosmic reionization.

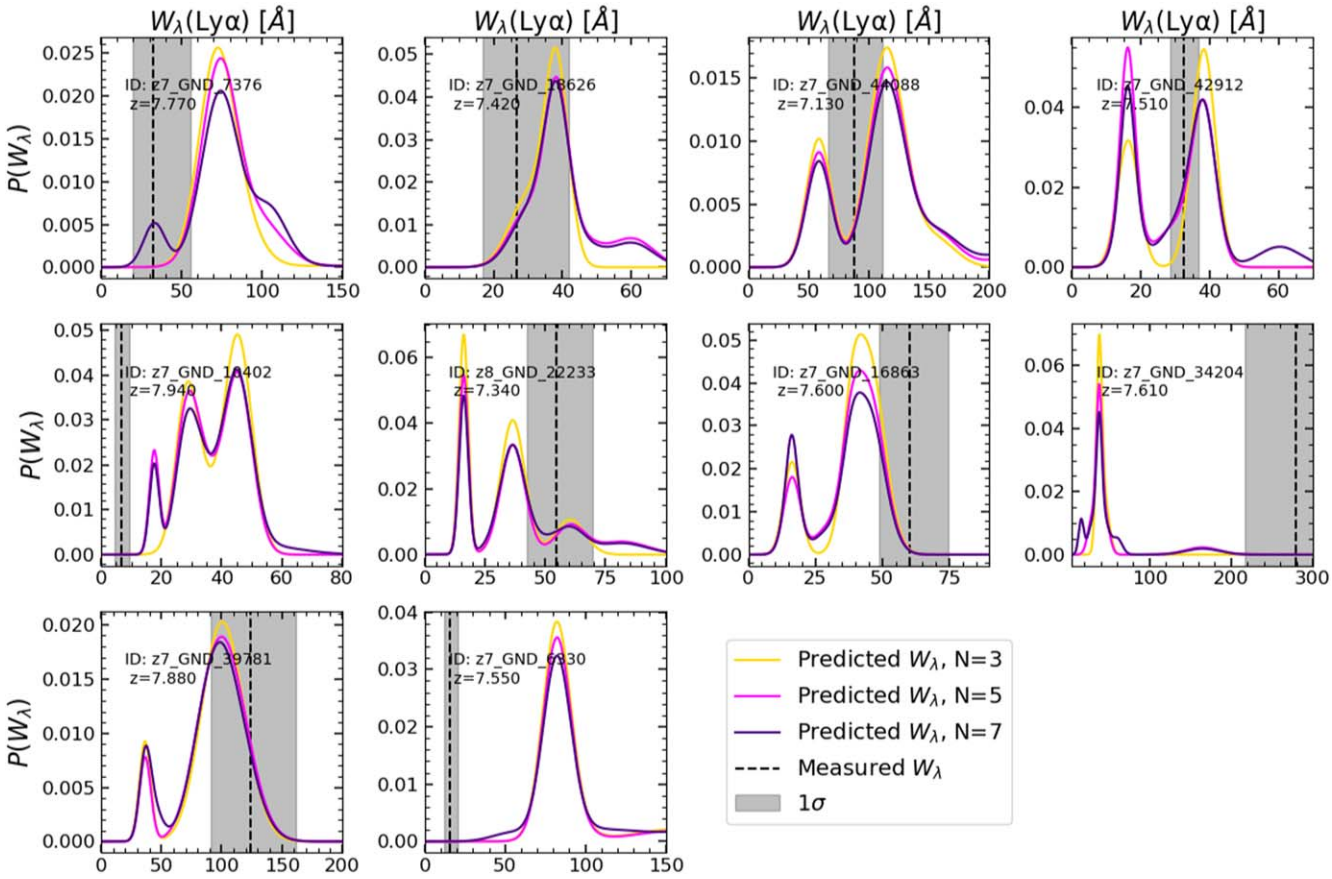


Figure 16. Probabilistic predictions of W_λ(Lyα) for 10 LAEs at $z > 7$ having emission-line S/N > 4. The distributions were normalized by setting their integrals to unity. We chose stellar mass, SFR, and dust extinction as predictive properties for this calculation. Gold, magenta, and indigo lines show the probability distribution of our predictions using $N = 3$, 5, and 7 nearest neighbors, and the gray shaded region shows the 68% confidence interval for the equivalent width measurements from Jung et al. (2020). Object IDs and redshifts are indicated with text for each plot. We find good agreement between prediction and observation for the majority of strong emitters.

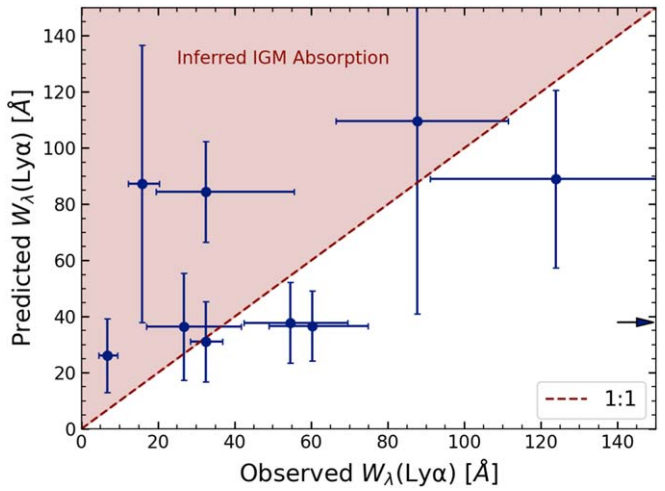


Figure 17. Predicted vs. observed W_λ(Lyα) for the LAE sample in Figure 16. We computed the first moment and square root of the second moment of each $N = 5$ distribution in Figure 16 for the predicted values and their error bars. A right-pointing arrow indicates the predicted value of object z7_GND_34204. A one-to-one dashed line is drawn to guide the eye, and points above this line (the region shaded red) could be the result of IGM absorption.

To build the sample, we inspected 842 emission-line detections to determine whether the line was Lyα or a feature from a low-redshift galaxy, such as [O II]. We then created a procedure to synthesize information about angular separation from the emission-line detection position, extracted emission-line flux, and χ^2 of SED fit assuming $z_{\mathrm{Ly}\alpha}$ to identify the continuum counterpart in our deep, multiband HST imaging in GOODS-N. After removing detections with no counterparts, AGN contaminants, and sources with insufficient photometric data, we analyzed a sample of 72 LAEs using SED fitting performed by BAGPIPES.

Our sample's properties were consistent with studies of LAEs from NB imaging surveys at similar redshifts. Our median sample mass was $0.8^{+2.9}_{-0.5} \times 10^9\ M_\odot$, and the galaxies' SFRs appeared to put them approximately on the star-forming main sequence, except for at $M_\star < 10^9\ M_\odot$. Using Lyα emission-line flux measurements from HETDEX, we also studied correlations between W_λ(Lyα) and galaxy properties. We found strong correlations between W_λ(Lyα) and stellar mass and SFR. We additionally found a moderate correlation where galaxies with older stellar populations had larger Lyα equivalent widths. Interestingly, we did not find a significant impact of dust extinction on W_λ(Lyα), whereas many other studies have. Overall, this paints a picture of LAEs as low-mass systems with moderate star formation activity wherein Lyα photons can escape even in the presence of dust. In addition, the LAEs detected by HETDEX do not stand out significantly in terms of their stellar population properties from LAEs found using NB imaging with comparable flux limits.

Finally, we used our LAE sample to try to predict the value of $W_\lambda(\mathrm{Ly}\alpha)$ for 10 LAEs at $z > 7$ by matching the distinct samples in the parameter space of mass, SFR, and dust extinction. Our prediction matched the data at the 1σ level 5 out of 10 times ($5/8$ for strong emitters); the three over-predictions could indicate significant absorption by a neutral hydrogen in the IGM. With large sample sizes in the near future and tools such as ML, we are optimistic about the ability of HETDEX LAEs to unlock the potential of Lyα as a reliable reionization probe.

A.P.M. and S.L.F. acknowledge support from the National Science Foundation, through grants AST-1908817 and AST-1614798.

HETDEX is led by the University of Texas at Austin McDonald Observatory and Department of Astronomy with participation from the Ludwig-Maximilians-Universität 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, 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, 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 München, Pennsylvania State University, Institut für Astrophysik Göttingen, University of Oxford, and the Max-Planck-Institut für Astrophysik (MPA).

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.

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 Pennsylvania State University.

I.J. acknowledges support from NASA under award number 80GSFC21M0002.

Software: get_spectrum.py (https://github.com/HETDEX/hetdex_api), emcee (Foreman-Mackey et al. 2013), Bagpipes (Carnall et al. 2018), Numpy (Harris et al. 2020).

Appendix A
Model Dependence of Measured Galaxy Properties

Bayesian approaches to SED fitting, like the one implemented in BAGPIPES, provide robust constraints on the parameter uncertainties and their interdependence, but the model chosen for comparison to the data (as well as the chosen priors) determines the accuracy of those estimates. In other words, an inaccurate model yields inaccurate measurements of galaxy properties. Many galaxy SED-fitting studies have shown that model choices, such as the SFH, systematically impact the measured galaxy properties (see Conroy 2013 for review).

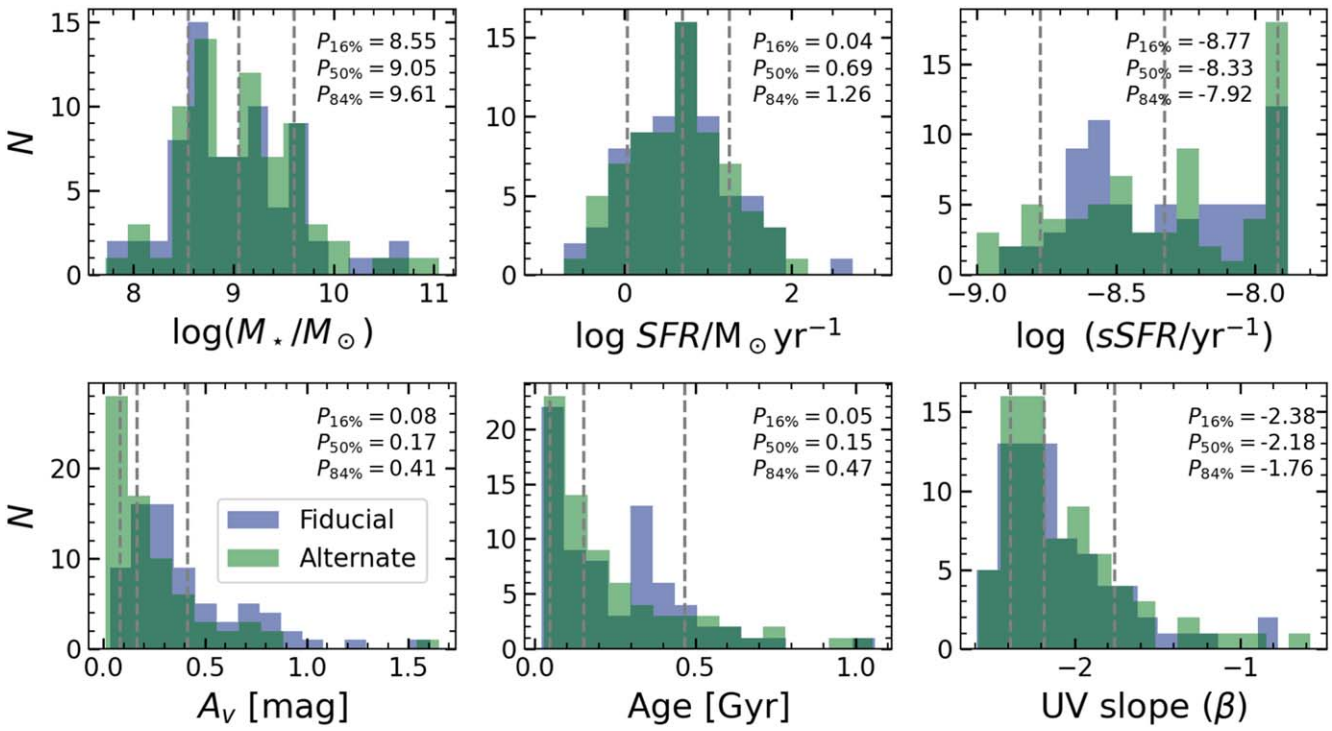


Figure 18. Comparison of galaxy properties as measured using our "fiducial" model (light blue) vs. our "alternate" model (sea green). The 16th, 50th, and 84th percentiles calculated using the alternate model are indicated by vertical dashed gray lines, and their values are indicated with text. The distributions are consistent, save for dust extinction (A_V), which has lower values by a factor of ∼2 for the alternate model.

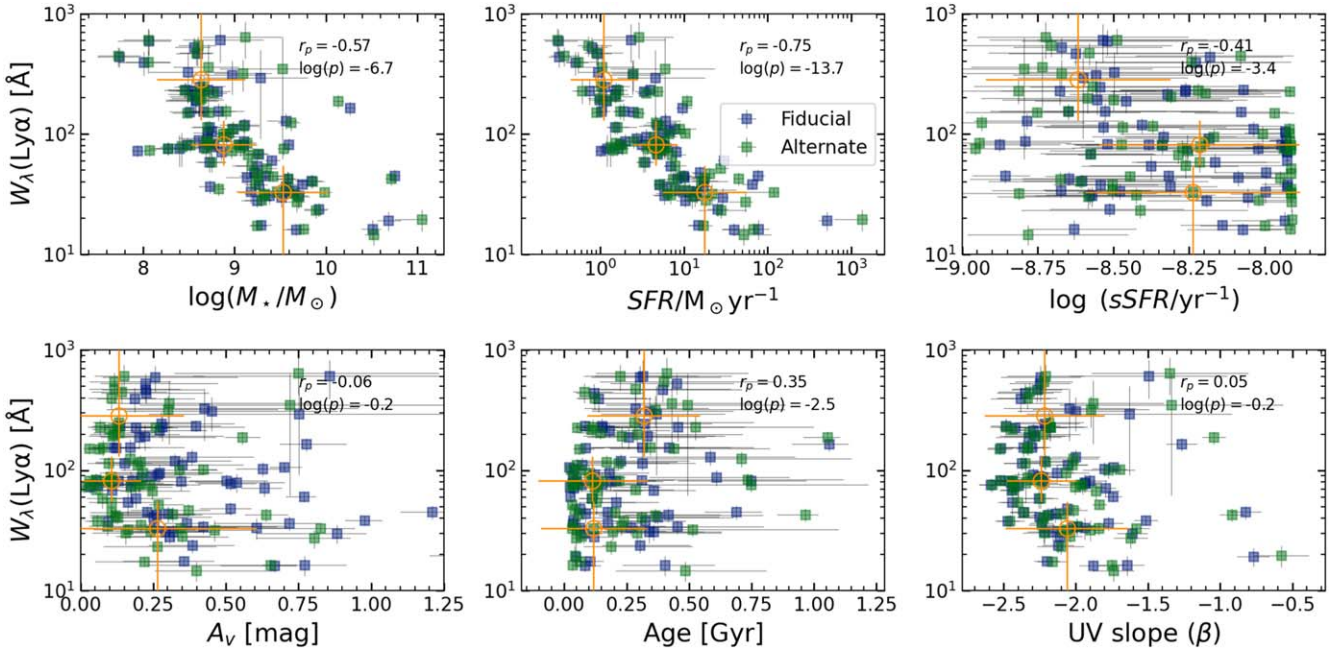


Figure 19. Comparison of correlations between W_λ(Lyα) and galaxy properties as measured using our "fiducial" model (light blue) vs. our "alternate" model (sea green). Binned values from the alternate model are indicated as open gold circles. The correlations presented in Section 4.3 appear robust when different models are adopted.

To test the robustness of our results to different modeling choices, we performed an additional analysis of our entire sample using an alternate model. We did not seek to find a more (or less) accurate model; we simply wanted a different model to determine whether the median properties or correlations between Lyα emission and galaxy properties changed. To this end, we adopted a constant SFH parameterization and the dust absorption model of Charlot & Fall (2000). The constant SFH required two parameters: the time when star formation began and the constant SFR. For dust attenuation, we adopted the recipe given in Charlot & Fall (2000) by using an absorption curve proportional to $\lambda^{-0.7}$ and a factor of three reduction in the dust extinction normalization for stellar populations older than 10^7 yr to account for the dispersal of stellar birth clouds. The authors found this recipe to match the absorption of stellar continuum and nebular emission for nearby starburst galaxies very well, and the differential extinction toward young stars differs markedly from the treatment by Calzetti et al. (1994) used in our "fiducial" model presented above.

Figure 18 shows the distribution of LAE properties measured using the alternate model compared with the fiducial model. The sample median stellar mass increased by 0.1 dex, as did the median SFR. These two changes do not affect our results or interpretation significantly. The median dust dropped from $A_V = 0.30$ to 0.17, a fairly substantial change, but not unusual given the common factors of roughly a few discrepancies between different models and SED-fitting codes (see Leja et al. 2017). Nonetheless, the correlations between galaxy properties and W_λ(Lyα) remained unaffected by the model modifications, as shown in Figure 19. Stellar mass and SFR correlated strongly and negatively with Lyα emission strength, while other parameters, like dust extinction, continued to show no significant correlations.

Appendix B
Imaging, Emission Lines, and SED Fits for LAEs in This Study

In this Appendix, for all 72 LAEs in our sample, we present HST imaging cutouts in Figure 20 showing the sources and any neighbors, the HETDEX Lyα emission-line detections in Figure 21, and the SED fits with BAGPIPES (Carnall et al. 2018) used to measure physical properties in Figure 22.

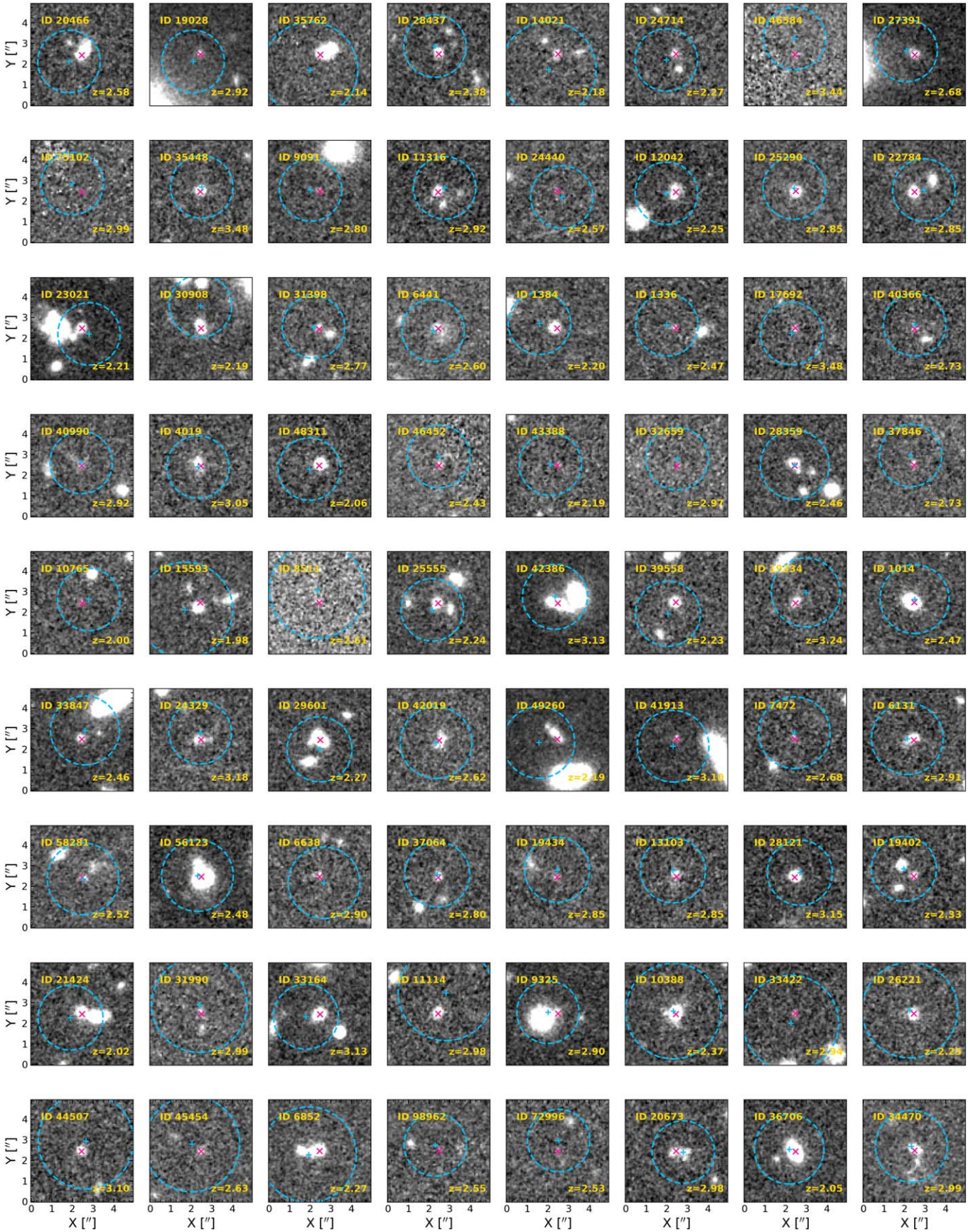


Figure 20. All HST F160W (*H*-band) images of LAEs in the sample. Each cutout shows a $5'' \times 5''$ image centered on each galaxy in our sample. The galaxy centroid is indicated with a pink diamond, and the HETDEX detection position and PSF FWHM are indicated by a light-blue cross and dashed circle, respectively. We also include object IDs and redshifts with text.

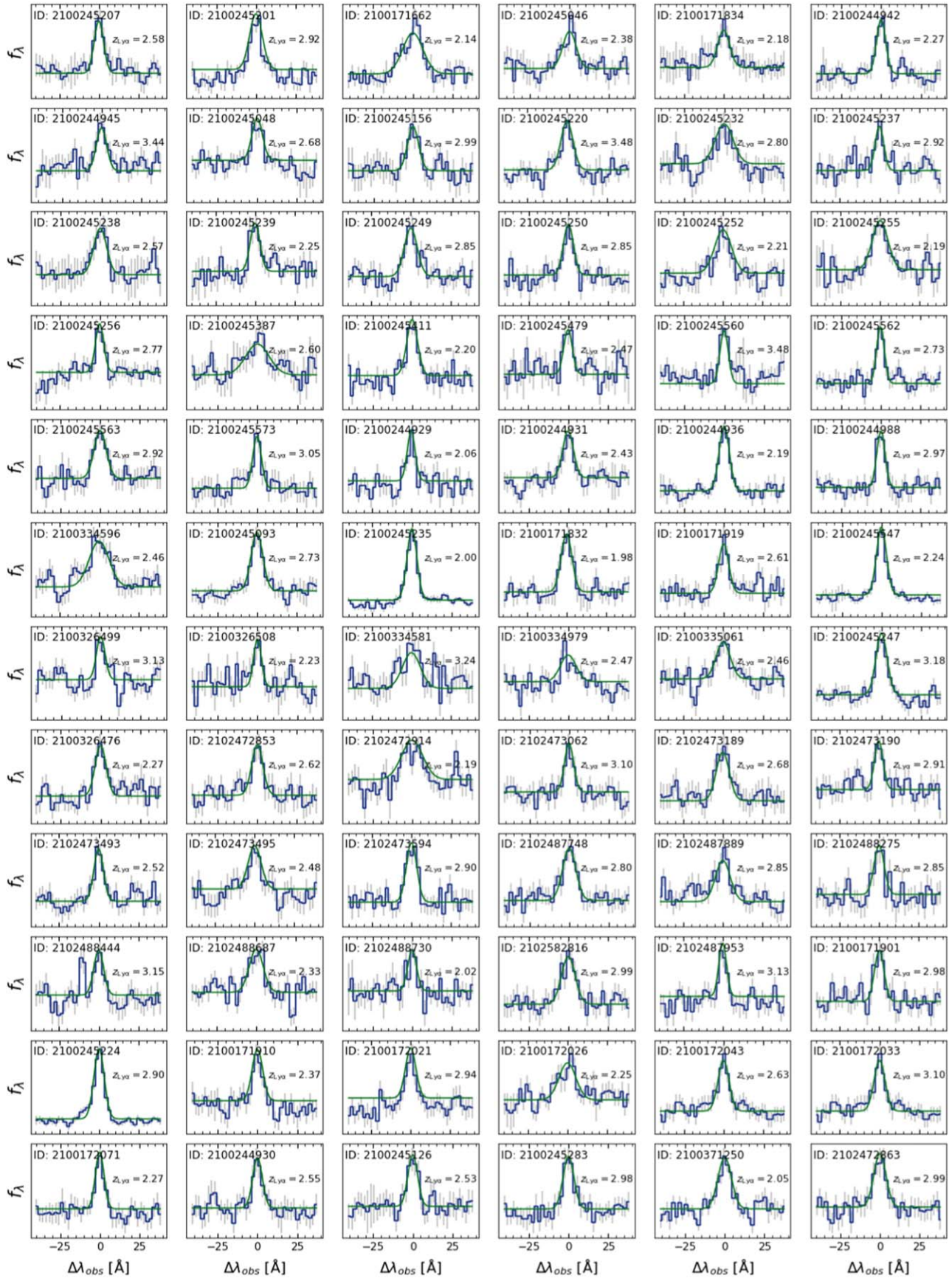


Figure 21. All emission-line detections from the HETDEX survey for LAEs in the sample. The observed data are indicated by the blue lines with gray error bars. A Gaussian model fit to the data is shown in green. The x-axis is scaled in angstroms relative to the line center. Detection IDs and Lyα line redshifts are indicated with text.

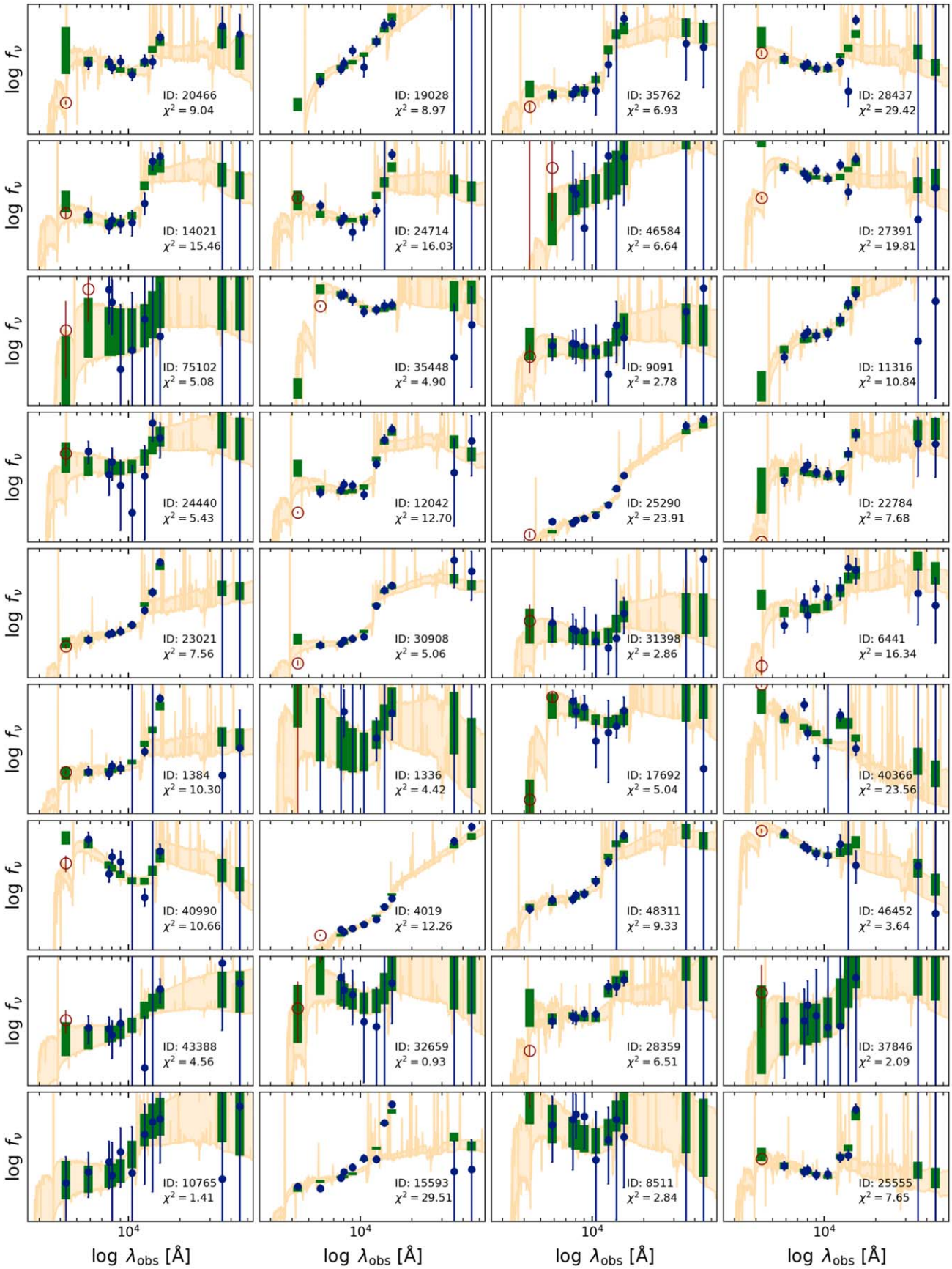


Figure 22. All SED fits for LAEs in the sample (see Figure 9 for a description of the plots). The χ^2 value for each fit is also given with text.

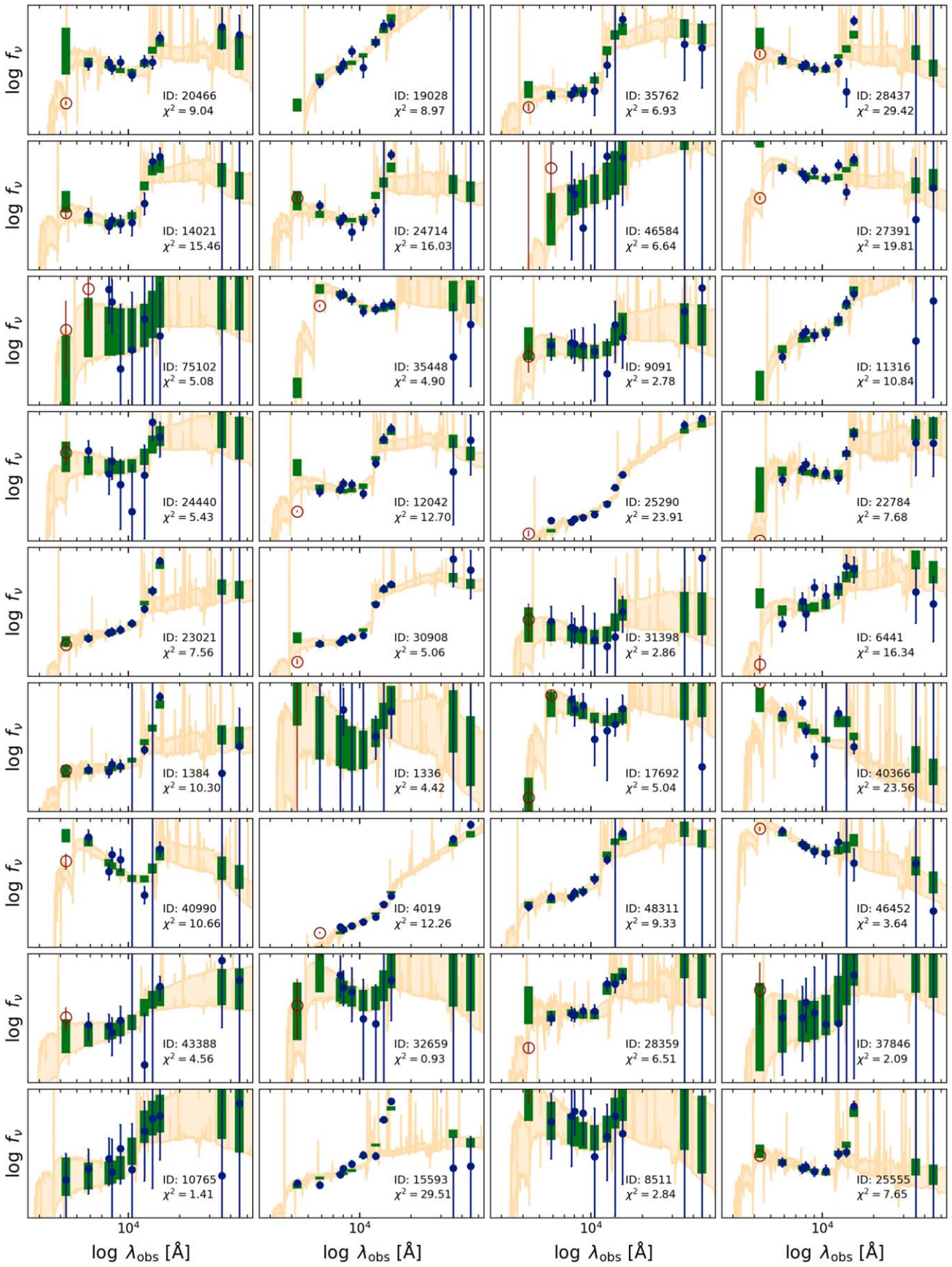


Figure 22. (Continued.)

ORCID iDs

Adam P. McCarron https://orcid.org/0000-0002-3912-9368
Steven L. Finkelstein https://orcid.org/0000-0001-8519-1130
Oscar A. Chavez Ortiz https://orcid.org/0000-0003-2332-5505
Dustin Davis https://orcid.org/0000-0002-8925-9769
Erin Mentuch Cooper https://orcid.org/0000-0002-2307-0146
Intae Jung https://orcid.org/0000-0003-1187-4240
Delaney R. White https://orcid.org/0000-0002-7707-9437
Gene C. K. Leung https://orcid.org/0000-0002-9393-6507
Karl Gebhardt https://orcid.org/0000-0002-8433-8185
Viviana Acquaviva https://orcid.org/0000-0002-6788-6315
William P. Bowman https://orcid.org/0000-0003-4381-5245
Robin Ciardullo https://orcid.org/0000-0002-1328-0211
Eric Gawiser https://orcid.org/0000-0003-1530-8713
Caryl Gronwall https://orcid.org/0000-0001-6842-2371
Gary J. Hill https://orcid.org/0000-0001-6717-7685
Wolfram Kollatschny https://orcid.org/0000-0002-0417-1494
Martin Landriau https://orcid.org/0000-0003-1838-8528
Chenxu Liu https://orcid.org/0000-0001-5561-2010
Daniel N. Mock https://orcid.org/0000-0003-4237-2470
Ariel G. Sánchez https://orcid.org/0000-0003-1198-831X

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Acquaviva, V., Gawiser, E., & Guaita, L. 2011, ApJ, 737, 47
Acquaviva, V., Vargas, C., Gawiser, E., & Guaita, L. 2012, ApJL, 751, L26
Adams, J. J., Blanc, G. A., Hill, G. J., et al. 2011, ApJS, 192, 5
Ando, M., Ohta, K., Iwata, I., et al. 2006, ApJL, 645, L9
Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2015, ApJS, 218, 33
Atek, H., Kunth, D., Schaerer, D., et al. 2014, A&A, 561, A89
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blanc, G. A., Adams, J. J., Gebhardt, K., et al. 2011, ApJ, 736, 31
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Chabrier, G. 2003, PASP, 115, 763
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319
Dijkstra, M. 2014, PASA, 31, e040
Du, X., Shapley, A. E., Reddy, N. A., et al. 2018, ApJ, 860, 75
Erb, D. K., Quider, A. M., Henry, A. L., & Martin, C. L. 2012, ApJ, 759, 26
Farrow, D. J., Sánchez, A. G., Ciardullo, R., et al. 2021, MNRAS, 507, 3187
Feltre, A., Maseda, M. V., Bacon, R., et al. 2020, A&A, 641, A118
Feroz, F., & Skilling, J. 2013, in AIP Conf. Proc. 1553, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 32nd Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. U. von Toussaint (Melville, NY: AIP), 106
Finkelstein, K. D., Finkelstein, S. L., Tilvi, V., et al. 2015, ApJ, 813, 78
Finkelstein, S. L., Bagley, M., Song, M., et al. 2022, ApJ, 928, 52
Finkelstein, S. L., D'Aloisio, A., Paardekooper, J.-P., et al. 2019, ApJ, 879, 36
Finkelstein, S. L., Papovich, C., Giavalisco, M., et al. 2010, ApJ, 719, 1250
Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, ApJ, 756, 164
Finkelstein, S. L., Rhoads, J. E., Malhotra, S., & Grogin, N. 2009, ApJ, 691, 465
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
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
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJL, 600, L93
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Gronke, M., Dijkstra, M., McCourt, M., & Oh, S. P. 2016, ApJL, 833, L26
Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79
Guaita, L., Acquaviva, V., Padilla, N., et al. 2011, ApJ, 733, 114
Guaita, L., Gawiser, E., Padilla, N., et al. 2010, ApJ, 714, 255
Hagen, A., Zeimann, G. R., Behrens, C., et al. 2016, ApJ, 817, 79
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hathi, N. P., Le Fèvre, O., Ilbert, O., et al. 2016, A&A, 588, A26
Hayes, M. 2015, PASA, 32, e027
Hayes, M., Östlin, G., Duval, F., et al. 2014, ApJ, 782, 6
Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, in ASP Conf. Ser. 399, The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results, ed. T. Kodama, T. Yamada, & K. Aoki (San Francisco, CA: ASP), 115
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
Horne, K. 1986, PASP, 98, 609
Huang, Y., Lee, K.-S., Shi, K., et al. 2021, ApJ, 921, 4
Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2020, ApJ, 904, 144
Jung, I., Finkelstein, S. L., Livermore, R. C., et al. 2018, ApJ, 864, 103
Kelz, A., Jahn, T., Haynes, D., et al. 2014, Proc. SPIE, 9147, 914775
Khostovan, A. A., Malhotra, S., Rhoads, J. E., et al. 2021, MNRAS, 503, 5115
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kornei, K. A., Shapley, A. E., Erb, D. K., et al. 2010, ApJ, 711, 693
Kusakabe, H., Shimasaku, K., Ouchi, M., et al. 2018, PASJ, 70, 4
Lang, D., Hogg, D. W., & Mykytyn, D. 2016, The Tractor: Probabilistic Astronomical Source Detection and Measurement, Astrophysics Source Code Library, ascl:1604.008
Lee, S.-K., Ferguson, H. C., Somerville, R. S., Wiklind, T., & Giavalisco, M. 2010, ApJ, 725, 1644
Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170
Leung, A. S., Acquaviva, V., Gawiser, E., et al. 2017, ApJ, 843, 130
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Malhotra, S., & Rhoads, J. E. 2004, ApJL, 617, L5
Marchi, F., Pentericci, L., Guaita, L., et al. 2019, A&A, 631, A19
Martin, C. L., Dijkstra, M., Henry, A., et al. 2015, ApJ, 803, 6
Matthee, J., Sobral, D., Hayes, M., et al. 2021, MNRAS, 505, 1382
Matthee, J., Sobral, D., Oteo, I., et al. 2016, MNRAS, 458, 449
Miralda-Escudé, J. 1998, ApJ, 501, 15
Neufeld, D. A. 1991, ApJL, 370, L85
Ono, Y., Itoh, R., Shibuya, T., et al. 2021, ApJ, 911, 78
Ouchi, M., Ono, Y., & Shibuya, T. 2020, ARA&A, 58, 617
Oyarzún, G. A., Blanc, G. A., González, V., Mateo, M., & Bailey, J. I. I. 2017, ApJ, 843, 133
Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620
Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, MNRAS, 412, 1123
Papovich, C., Shipley, H. V., Mehrtens, N., et al. 2016, ApJS, 224, 28
Partridge, R. B., & Peebles, P. J. E. 1967, ApJ, 147, 868
Pentericci, L., Grazian, A., Fontana, A., et al. 2009, A&A, 494, 553
Pentericci, L., Grazian, A., Scarlata, C., et al. 2010, A&A, 514, A64
Ramsey, L. W., Sebring, T. A., & Sneden, C. A. 1994, Proc. SPIE, 2199, 31
Reddy, N. A., Topping, M. W., Shapley, A. E., et al. 2022, ApJ, 926, 31
Rhoads, J. E., Malhotra, S., Dey, A., et al. 2000, ApJL, 545, L85
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2018, ApJ, 858, 99
Santos, S., Sobral, D., Matthee, J., et al. 2020, MNRAS, 493, 141
Scarlata, C., Colbert, J., Teplitz, H. I., et al. 2009, ApJL, 704, L98
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shimakawa, R., Kodama, T., Shibuya, T., et al. 2017, MNRAS, 468, 1123
Sobral, D., Santos, S., Matthee, J., et al. 2018, MNRAS, 476, 4725
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Trainor, R. F., Strom, A. L., Steidel, C. C., et al. 2019, ApJ, 887, 85
Vargas, C. J., Bish, H., Acquaviva, V., et al. 2014, ApJ, 783, 26
Weiss, L. H., Bowman, W. P., Ciardullo, R., et al. 2021, ApJ, 912, 100
Wold, I. G. B., Barger, A. J., & Cowie, L. L. 2014, ApJ, 783, 119
Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2016, ApJS, 224, 15
York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, AJ, 120, 1579