The Astrophysical Journal, 943:177 (30pp), 2023 February 1 https://doi.org/10.3847/1538-4357/aca962

HETDEX Public Source Catalog 1: 220 K Sources Including Over 50 K Lyα Emitters from an Untargeted Wide-area Spectroscopic Survey*

Erin Mentuch Cooper[1,2], Karl Gebhardt[1], Dustin Davis[1], Daniel J. Farrow[3,4], Chenxu Liu[1], Gregory Zeimann[5], Robin Ciardullo[6,7], John J. Feldmeier[8], Niv Drory[2], Donghui Jeong[6,7], Barbara Benda[9], William P. Bowman[6,7], Michael Boylan-Kolchin[1], Óscar A. Chávez Ortiz[1], Maya H. Debski[6], Mona Dentler[10], Maximilian Fabricius[3,4], Rameen Farooq[11], Steven L. Finkelstein[1], Eric Gawiser[11], Caryl Gronwall[6,7], Gary J. Hill[1,2], Ulrich Hopp[3,4], Lindsay R. House[1], Steven Janowiecki[12], Hasti Khoraminezhad[13], Wolfram Kollatschny[10], Eiichiro Komatsu[14,15], Martin Landriau[16], Maja Lujan Niemeyer[14], Hanshin Lee[2], Phillip MacQueen[2], Ken Mawatari[17,18], Brianna McKay[19], Masami Ouchi[15,18,20], Jennifer Poppe[1], Shun Saito[13,15], Donald P. Schneider[6,7], Jan Snigula[4], Benjamin P. Thomas[1], Sarah Tuttle[19], Tanya Urrutia[21], Laurel Weiss[1], Lutz Wisotzki[21], and Yechi Zhang[18,22]

The HETDEX collaboration

[1] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA; erin@astro.as.utexas.edu
[2] McDonald Observatory, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA
[3] University Observatory, Fakultät für Physik, Ludwig-Maximilians University Munich, Scheinerstrasse 1, D-81679 Munich, Germany
[4] Max-Planck Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany
[5] Hobby-Eberly Telescope, University of Texas, Austin, Austin, TX 78712, USA
[6] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[7] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[8] Department of Physics, Astronomy, Geology, and Environmental Sciences, Youngstown State University Youngstown, OH 44555, USA
[9] Physics and Astronomy Department, Rutgers, The State University, Piscataway, NJ 08854-8019, USA
[10] Institut für Astrophysik, Universität Göttingen, Friedrich-Hund-Platz 1, D-37077 Göttingen, Germany
[11] Physics and Astronomy Department, Rutgers, The State University, Piscataway, NJ 08854, USA
[12] University of Texas, Hobby-Eberly Telescope, McDonald Observatory, TX 79734, USA
[13] Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N Pine Street, Rolla, MO 65409, USA
[14] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany
[15] Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, the University of Tokyo, Kashiwanoha, Kashiwa, Chiba 277-8583, Japan
[16] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[17] National Astronomical Observatory of Japan, Osawa 2-21-1, Mitaka, Tokyo 181-8588, Japan
[18] Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan
[19] Department of Astronomy, University of Washington, Seattle, 3910 15th Avenue NE, Room C319, Seattle, WA 98195-0002, USA
[20] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[21] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany
[22] Department of Astronomy, Graduate School of Science, the University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
Received 2022 August 18; revised 2022 December 2; accepted 2022 December 4; published 2023 February 7

Abstract

We present the first publicly released catalog of sources obtained from the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX). HETDEX is an integral field spectroscopic survey designed to measure the Hubble expansion parameter and angular diameter distance at $1.88 < z < 3.52$ by using the spatial distribution of more than a million Lyα-emitting galaxies over a total target area of 540 deg^2. The catalog comes from contiguous fiber spectra coverage of 25 deg^2 of sky from 2017 January through 2020 June, where object detection is performed through two complementary detection methods: one designed to search for line emission and the other a search for continuum emission. The HETDEX public release catalog is dominated by emission-line galaxies and includes 51,863 Lyα-emitting galaxy (LAE) identifications and 123,891 [O II]-emitting galaxies at $z < 0.5$. Also included in the catalog are 37,916 stars, 5274 low-redshift ($z < 0.5$) galaxies without emission lines, and 4976 active galactic nuclei. The catalog provides sky coordinates, redshifts, line identifications, classification information, line fluxes, [O II] and Lyα line luminosities where applicable, and spectra for all identified sources processed by the HETDEX detection pipeline. Extensive testing demonstrates

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

that HETDEX redshifts agree to within $\Delta z < 0.02$, 96.1% of the time to those in external spectroscopic catalogs. We measure the photometric counterpart fraction in deep ancillary Hyper Suprime-Cam imaging and find that only 55.5% of the LAE sample has an r-band continuum counterpart down to a limiting magnitude of $r \sim 26.2$ mag (AB) indicating that an LAE search of similar sensitivity to HETDEX with photometric preselection would miss nearly half of the HETDEX LAE catalog sample. Data access and details about the catalog can be found online at http://hetdex.org/. A copy of the catalogs presented in this work (Version 3.2) is available to download at Zenodo doi:10.5281/zenodo.7448504.

Unified Astronomy Thesaurus concepts: Redshift surveys (1378); Catalogs (205); Emission line galaxies (459); Lyman-alpha galaxies (978)

1. Introduction

Systematic wide-area spectroscopic surveys undertaken in the past two decades, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000), BOSS (Dawson et al. 2013), eBOSS (Dawson et al. 2016), and DESI (Abareshi et al. 2022) have resulted in orders of magnitude increase in the number of moderate-resolution spectra available for study. These investigations, thus far, select their spectroscopic targets based upon multiwavelength photometric imaging. Targets are chosen based on continuum brightness, color, morphology, determined stellar mass, and determined star formation rates. These surveys, with their well-defined observing limits and well-characterized systematic uncertainties, have greatly advanced our understanding of the universe.

The above surveys have compiled extensive galaxy samples out to $z \sim 1$. At higher redshifts, spectroscopic surveys of galaxies are limited to relatively small solid angle regions, where deep imaging aids in the construction of magnitude-limited samples that are sufficiently bright to yield spectroscopic redshifts. Examples of these surveys include the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007), and the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004), which both provide unprecedented views of our universe with Hubble Space Telescope (HST) and complementary ground-based imaging data. Spectroscopic redshifts in both of these fields number in the tens of thousands (e.g., Wirth et al. 2004, 2015; Reddy et al. 2006; Barger et al. 2008; Ferreras et al. 2009; Cooper et al. 2011; Kriek et al. 2015; Momcheva et al. 2016; Hasinger et al. 2018) and provide important benchmarks for photometric redshifts, as well as numerous targeted investigations in these legacy fields.

At redshifts larger than two, galaxy samples are often targeted based upon color and magnitude, depending on the science goals. In most cases, these data sets will be biased toward bright, high stellar-mass objects (e.g., Kriek et al. 2008; Marsan et al. 2017) and come from a variety of observatories and heterogeneous sensitivity limits. However, at high redshift, the higher spatial densities of low-mass galaxies provide a stronger tracer of the galaxy distribution (Muzzin et al. 2013; Finkelstein et al. 2015; Song et al. 2016). For these faint galaxies, spectroscopic redshifts are difficult to obtain from absorption features, and it is most efficient to rely on emission lines for redshifts. The strong line emission from Lyα-emitting galaxies (LAEs) allows detection over a wide range of stellar mass (e.g., Shapley et al. 2003; Hu & Cowie 2006) and redshifts for objects generally too faint for detection in broadband images (Hagen et al. 2016; Oyarzún et al. 2017; Santos et al. 2020). See Ouchi et al. (2020) and references therein for a thorough review.

LAE surveys are traditionally conducted by comparing an object's flux through a narrowband filter with that seen in broadband imaging (e.g., Cowie & Hu 1998; Rhoads et al. 2000; Gronwall et al. 2007; Ouchi et al. 2008; Konno et al. 2016; Sobral et al. 2018; Spinoso et al. 2020; Ono et al. 2021). Such searches can be quite successful, but cover relatively small slices in redshift space, as only those objects that have Lyα redshifted into the bandpass of the narrowband filter are detected. Recent searches (Benitez et al. 2014; Eriksen et al. 2019; Bonoli et al. 2021) are optimizing the technique by utilizing a high number of narrowband filters, providing for higher redshift coverage, improved source identification, and efficient, homogeneous sky coverage.

Alternatively, an efficient method to survey large volumes of high-redshift (high-z) space is through Integral Field Unit (IFU) observations (van Breukelen et al. 2005; Adams et al. 2011; Bacon et al. 2015; Urrutia et al. 2019). IFU observations provide simultaneous redshift coverage along with spatial information in the field of view (FOV), limiting the need for follow-up spectroscopy and providing spectral information for neighboring sources, which can aid in identifying contaminants due to spatially extended line emission from low-z galaxies. Though IFU surveys can still be subject to occasional contamination by lower-redshift galaxies and active galactic nuclei (AGNs), especially when the wavelength coverage of the spectrograph is limited, such instruments are more efficient at detecting high-z LAEs than narrowband imaging.

One such instrument is the Visible IFU Replicable Unit Spectrograph (VIRUS; Hill et al. 2021), on the 10 m Hobby-Eberly Telescope (HET; Ramsey et al. 1998; Hill et al. 2021), which can obtain ≈35,000 spectra simultaneously, each covering the wavelength range $3500\,\mathring{A} \lesssim \lambda \lesssim 5500\,\mathring{A}$ with spectral resolving power $750 < R < 950$. VIRUS is the primary instrument of the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX; Gebhardt et al. 2021), whose goal is to measure the Hubble parameter, $H(z)$, and the angular diameter distance, $D_A(z)$, to better than 1% accuracy in the redshift range $1.9 < z < 3.5$. HETDEX uses LAEs as a (biased) tracer of dark matter density; by measuring their clustering, HETDEX characterizes the evolution in the universe's dark energy density and tests for potential evolution (Shoji et al. 2009). To achieve the desired accuracy, HETDEX needs to measure at least one million LAEs over 540 deg^2 of sky, or 10.9 Gpc^3 in the targeted redshift range. The project does not need complete coverage within this sky area to accomplish its scientific goals; as discussed by Chiang et al. (2013), a fill factor of 0.22 (1/4.6), which optimizes the number of IFUs given the area of the focal plane of the HET, is sufficient. For the target number of LAEs, HETDEX needs an exposure time sufficient to reach about 2.5 LAEs per IFU. The typical total exposure time is 20 minutes per field. With 468,000 IFU observations, at 2.5 LAEs per IFU, we reach the goal of one million LAEs. If the experiment falls short of this goal, the sky

area can be adjusted if needed to reach the target number of objects.

The first observations of HETDEX were obtained in 2017 January, with VIRUS in commissioning mode at a fraction of its current capability. In 2017 the project started with 11 working IFUs; by 2021 August, the maximum number of 78 IFUs were operational in the focal plane. This paper presents the first general public catalog of HETDEX sources acquired over the first 3 yr of the survey. These sources come from HETDEX's dual object detection method, described in Gebhardt et al. (2021), that searches for line emission sources and continuum emission sources independently within the HETDEX IFU data set. Although designed to find LAEs, the untargeted IFU data also observes a wide range of astronomical sources. This catalog provides coordinates, redshifts, spectra, and measured properties of 223,641 objects, which we organize into five source types that are referenced throughout this paper: Lyα-emitter as `lae`, [O II]-emitting galaxy as `oii`, active galactic nuclei as `agn`, low-z galaxy (with no measured [O II] line emission) as `lzg`, and $z = 0$ sources as `star`. Transient objects such as meteors and satellites are not included, nor are large nearby galaxies: these objects will be published at a later time.

The outline of this paper is as follows. Section 2 describes the observations obtained for the HETDEX survey and details concerning the quality assessment of the observations. Section 3 describes the process of going from raw object detection to an astronomical source. Section 4 describes source classification and redshift assignment. In Section 5, we provide the data format of the catalog, and Section 6 presents properties of the catalog samples.

Accompanying this paper are two separate catalogs. The first is the *Source Observation Table* (columns described in Table 3), which is a summary of information for each HETDEX observation of a single astronomical source. Its position, classification, redshift, as well as line flux and luminosity measurements are provided for each observation. Here, the group of detections that comprise the source are reduced to one representative detection per source observation, and we provide the spectrum for that detection in a separate FITS file. In addition to this aggregate table, we provide the *Detection Info Table* (columns described in Table 6), which provides information for every HETDEX detection that has passed a series of quality checks and object detection criteria. Line emission detection information, such as observed wavelength and line fluxes are provided for every HETDEX detection in this table and can include a variety of line species, unlike Table 3, which is limited specifically to Lyα and [O II] line flux and luminosity measurements for simplicity if they are relevant for a source (e.g., a star or low-z galaxy will not have an accompanying Lyα or [O II] measurement).

All positions reported in this paper are in the International Celestial Reference System (ICRS). We adopt the flat Λ cold dark matter cosmology with $H_0 = 67.7\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $\Omega_{m,0} = 0.31$ measured by Planck Collaboration et al. (2020). All magnitudes are expressed in the AB system (Oke & Gunn 1983). We assume a rest-frame vacuum wavelength of $\lambda = 1215.67$ Å for Lyα and rest-frame air wavelength of $\lambda = 3727.8$ Å for the [O II] doublet, integrated to our instrumental resolution. Observed wavelengths expressed in this paper and associated data products are as measured in air. All redshifts are appropriately calculated for any differences between air and vacuum wavelengths using the standard in Morton (1991).

2. Observations

The data on which these catalogs are based were all obtained in HETDEX survey observations (Gebhardt et al. 2021) using the IFUs of VIRUS, the fiber-fed, multispectrograph instrument of the HET (Hill et al. 2021). Each IFU feeds a pair of VIRUS spectrographs with 448 1.″5 diameter fibers positioned on a rectangular array with the fiber center separations of 2.″5. At a given pointing, three exposures, each typically 6–7 minutes in duration (the exposure times range from 3.6–12 minutes, depending upon observing conditions), are obtained; the telescope is dithered in a triangular pattern to obtain a complete fill factor for each of the $51'' \times 51''$ IFU fields (see Gebhardt et al. 2021; Hill et al. 2021). Figure 2 provides an example IFU fiber layout for this three-dither pattern. In a single IFU observation, 1344 fiber spectra are obtained providing full sky coverage of the IFU. Also shown in varying shades of color are the four amplifiers that compose the IFU. Because each amplifier channel is fed to its own detector channel, we consider these components individually in the quality assessment of HETDEX observations. At full completion of the VIRUS instrument, its 78 IFUs cover approximately 21.7% (a factor of 4.6) of the HET's 22′ diameter FOV.

Survey data for this catalog release come from the internal HETDEX Data Release 2 (HDR2). This release consists of 2797 observations obtained starting in 2017 January, when the VIRUS IFU assembly contained just 16 operational IFUs, and ending on 2020 June 26 when 71 IFUs were installed within the VIRUS array. The full complement of 78 IFU units became operational in 2021 August. This catalog is generated from 134,831 IFU observations of which 124,472 (92.3%) pass our quality control pipeline described in detail in the following sections. A tally of IFU observations in each field is given in Table 1. The total sky coverage of the catalog is 25.0 deg^2.

The HETDEX footprint consists of two primary fields that allow for full-year surveying as shown in Figure 1. The spring field, labeled as the `dex-spring` field throughout this paper and in the associated catalog, covers 390 deg^2 of high decl. ($\delta \sim 51°$) sky while the fall field, labeled as `dex-fall` in the catalog, covers 150 deg^2 along the celestial equator (see Gebhardt et al. 2021 for full details on field selection). To reach the survey science requirements, 468,000 IFU observations are needed at the current technical specifications, resulting in 94 deg^2 of complete sky coverage. In addition to the two primary fields, HETDEX obtained a number of science verification observations of COSMOS ($\sim$2.0 deg^2) and GOODS-N ($\sim$0.09 deg^2). While most of these data were acquired using the exposure times and dithering pattern described above, several of these fields were taken with longer exposures or were visited multiple times during the survey.

HETDEX observations are expected to be completed in 2024, and eventually cover 540 deg^2 with partial fill. The final effective sky coverage is expected to be about 94 deg^2 with noncontiguous tiling over the two main HETDEX fields in combination with the 21.7% fill factor of the VIRUS IFU array. Figure 1 shows the survey boundaries in red and the source positions in blue, which effectively map out the HETDEX IFU field boundaries in this release. Examples of IFU boundaries are overlaid in blue over DESI Legacy Imaging Data (Dey et al. 2019) in the left panel of Figure 3 for a 1 deg^2 region in the HETDEX fall field. The right panel zooms into a much smaller $6' \times 6'$ region, indicating the IFU boundaries in white and source positions of HETDEX sources as described in the

Table 1
Catalog Release Survey Statistics

Field	Field ID	Center (R.A., Decl.)	*N*(IFU) Observed	*N*(IFU) Included	Area (deg^2)	*N*(source)
HETDEX-Spring	dex-spring	$13^h + 51°$	96,955	89,603	17.97	172,831
HETDEX-Fall	dex-fall	1.5^h 0°	35,741	33,001	6.62	56,251
COSMOS	cosmos	$10^h + 2°$	1506	1340	0.27	2447
GOODS-N	goods-n	$12.5^h + 62°.2$	638	528	0.11	1121

Note. Listed is the count of IFU observations observed and the count included after observation quality inclusion criteria described in Section 2.1. The `field` ID is the string match to find each field column in the catalog.

legend. This cropped region covers only a quarter of the HET FOV. Overlapping observations are seen from two independent observations (taken at different HET track angles). Overlapping IFUs such as this provide valuable repeated observations for validation tests discussed later in Sections 6.3.2 and 6.6.

The data processing of HETDEX frames is detailed in Gebhardt et al. (2021). Briefly, bias frames, pixel flats, twilight sky flats, and the background on the science frames themselves are used to 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 from fiber counts between 4400 and 5200 Å and comparing their IFU positions to the stars' equatorial coordinates in SDSS (York et al. 2000) and Gaia (Gaia Collaboration et al. 2018) catalogs. This process typically results in global solutions that are accurate to $\sim 0''.2$. The absolute flux calibrations are produced by using $g < 24$ SDSS field stars as in situ standards and using their *ugriz* colors (Padmanabhan et al. 2008), Gaia parallaxes (Gaia Collaboration et al. 2018), and foreground reddenings (Schlafly & Finkbeiner 2011) to determine their most likely spectral energy distribution in a grid of model spectra (Cenarro et al. 2007; Falcon-Barroso et al. 2011). The final system throughput curve is derived from the most likely flux distribution of ~20 stars, and is generally accurate to ~5% (Gebhardt et al. 2021).

2.1. Data Quality Control

An accurate description of sky sensitivity and coverage is essential for HETDEX. Each IFU consists of 448 fibers that are divided into two spectrograph channels. Each channel has a 2064×2064 detector. We use two amplifiers per spectrograph channel and bin $2\times$ in the spectral direction during readout. Thus, each IFU consists of four amplifier channels, labeled "RU," "RL," "LL," and "LU" as demonstrated in the left panel of Figure 2. Each amplifier generates an FITS image that is 1032×1032, each with 112 fiber spectra. With a full 78 IFU installation, each single exposure consists of data from 312 CCD amplifiers, which corresponds to about 35,000 fiber spectra. Our standard three-dithered observation set generates 936 FITS files, each an image of a single amplifier, and 104,000 fiber spectra. Although the IFU spectrographs are designed to be identical, in practice, there are important variations from amplifier to amplifier that we track (see, for example, Figure 6 from Gebhardt et al. 2021). Over its lifetime, including calibration frames, HETDEX will generate about 20 million FITS files. Each one of the amplifier images consisting of 112 fiber images is considered individually for quality assessment.

Over the first 3 yr of the HETDEX survey, we have seen a variety of detector and calibration issues. These include dead amplifiers, variable electronic noise, low count rates, and scrambled pixels. Calibration issues include vignetting of some IFUs, saturation problems from bright objects, astrometric uncertainties from fields with low number of stars, large variation in throughput over a dither set, large variations in the wavelength solution for some spectrographs, among others. We robustly track these issues and find that for a given exposure set, about 92% of the FITS files are useful and make it into the catalog. This percentage has increased over the years as we have fixed various detector issues, and we expect an even higher rate of return in the future.

2.1.1. Detector Issues

Instrument deficiencies can result in a number of failures. Specific detectors may exhibit low response or spatially distorted features that cannot be removed by flat-field corrections. These failures can vary with time and significantly impact our detection methods. Certain failures result in the creation of many false detections and can dramatically affect our detected sample. Building from an initial sample that was visually flagged, we have developed a set of criteria from statistics generated by our image calibration that automatically identifies substandard amplifier readouts. These criteria are summarized in Table 2. We provide the quality inclusion criterion limits for each statistic and a short description. We also indicate the fraction of amplifiers that pass each criteria. Ultimately, an amplifier must satisfy all of the criteria to be included. Unfortunately not all issues can be caught automatically, and an extensive list of detector issues for each detector are maintained[23] so that both the catalog and the survey selection function are consistently masked. For this catalog release, about 92% of the FITS files pass our quality control, although we note that the first year of observations were particularly poor with the fraction of usable data below 90%. The quality fraction generally averages about 94% in recent years. An example where a specific amplifier is removed from the survey is shown in the right panel of Figure 3. The IFU at roughly $20° 11'$, $00° 02'$ has a single amplifier masked out of the catalog.

2.1.2. Calibration Failures

Science frames that cannot be calibrated to the HETDEX specification (Gebhardt et al. 2021) are also removed from our catalog. These data are usually produced by the presence of bright stars or large galaxies on or near an IFU. While the HETDEX tiling attempts to avoid the very brightest stars, the

[23] https://github.com/HETDEX/hetdex_api/tree/master/known_issues

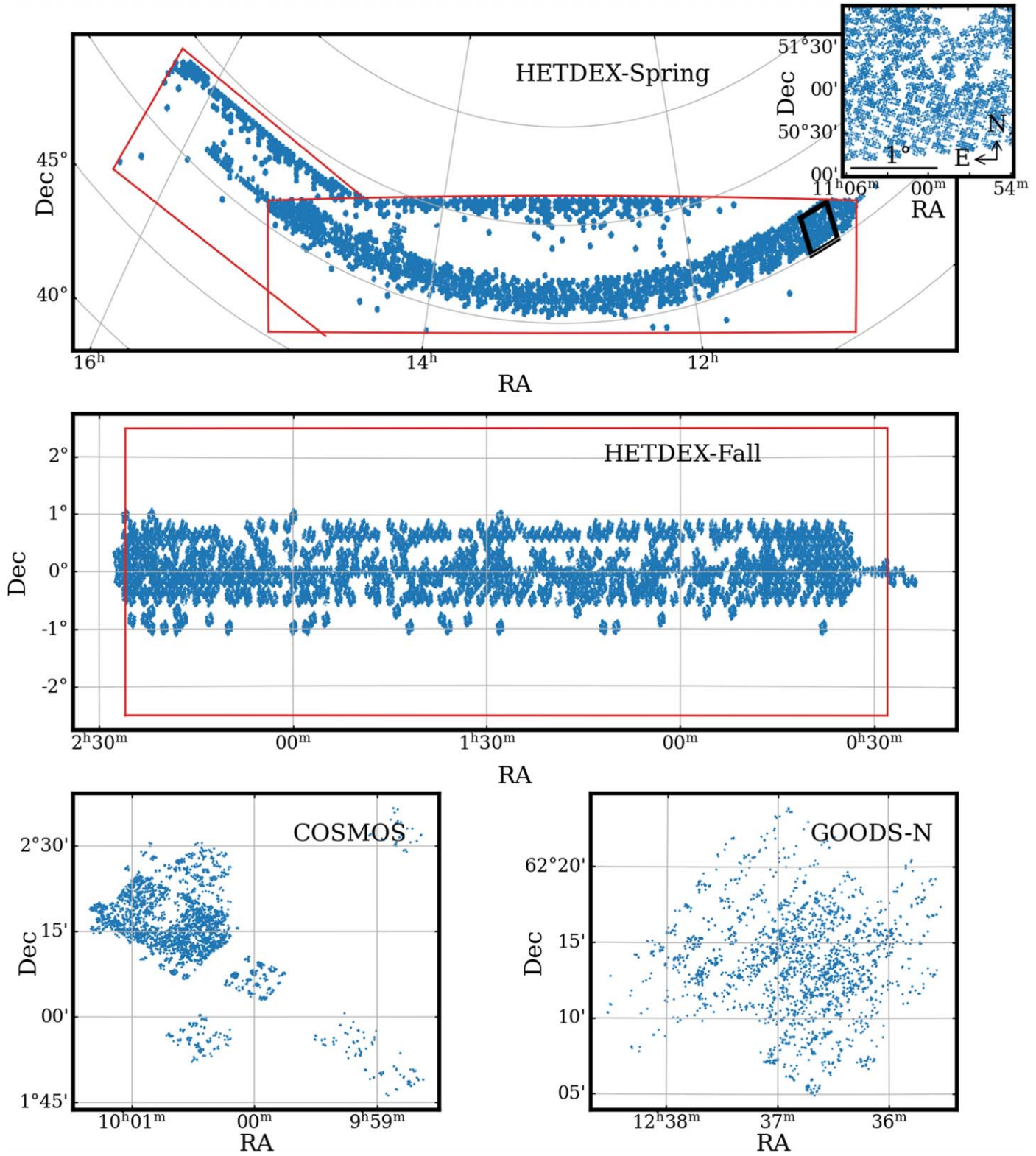

Figure 1. Sky coverage of the planned HETDEX science fields (in red) and the footprint of this catalog release (in blue): (1) the high decl. spring field (top), which is centered at $13^{\rm h}.5$, +51° and covers ∼390 deg^2 of the sky, and (2) the equatorial fall field (middle), which is centered at $1^{\rm h}.5$, 0°, covers ∼150 deg^2 of the sky. Also highlighted are the two legacy fields, COSMOS (bottom-left) and GOODS-N (bottom-right), where some coverage is included in this catalog. The blue points represent catalog sources, which effectively trace the IFU array footprint of this release. Each IFU has an FOV of $51'' \times 51''$, which means that the full VIRUS IFU array covers a 0.22 fill factor in the HET's 22′ diameter FOV. The expanded inset on the top-right presents a typical 2 deg^2 region in the spring field (representing the rectangular region in the top panel).

spectra of objects brighter than $g \sim 14$ will typically saturate a detector, and flood nearby fibers (on the detector) with excess signal. The counts in these fibers cannot be flux calibrated. The criteria set out in the previous section and summarized in Table 2 also help to automatically remove any frames that have calibration issues.

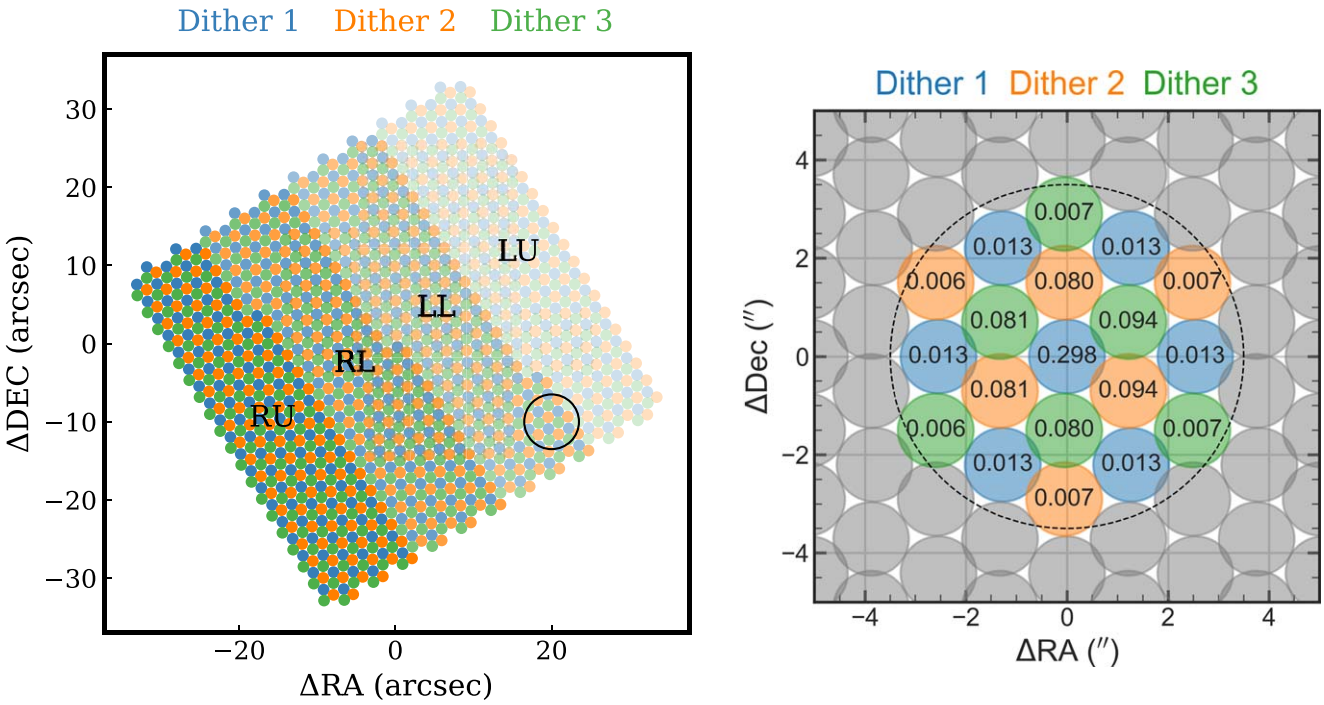

Figure 2. Fiber layout for a typical three-dither HETDEX/VIRUS observation for a single 51″ × 51″ IFU is shown on the left panel. The filled circles represent the fiber footprint on the sky illustrating the 1.″5 diameter fiber locations for three separate dithered exposures as colored. The three dithers provide complete coverage of the IFU. The different shades indicate the four amplifier channels of the IFU. The black outlined circle indicates an example aperture used for an HETDEX point-spread function (PSF)-weighted spectral extraction ($r = 3.''5$). The right panel provides a close-up example of a spectral extraction with the axis centered relative to aperture center. The values in each circle represent the fractional contribution of the specific fiber to the extracted summed spectrum at 4500 Å for a point-source model centered at 0,0 with 1.″8 FWHM. These fractions change with wavelength due to differential atmospheric refraction resulting in asymmetrical fiber weights, depending on the zenith direction. The dashed line is a 3.″5 aperture and is the radius at which we collect fibers for extraction.

2.1.3. Observation Quality Criteria

Each dither in a HETDEX observation is individually flux calibrated, as there may be small differences in their relative throughputs due to variations in the observing conditions. For the HETDEX catalog, we require that a nominal throughput, assuming a 360 s exposure time, must be greater than 0.08, and that the relative throughputs of each dithered exposure cannot differ by more than a factor of 3. The most common reason for rejection by this criterion is a significant drop in transparency during the third dithered exposure when clouds drifted into the FOV.

2.1.4. Pixel Masks

As described in Gebhardt et al. (2021), several detectors have significant features, including large dust spots, many charge traps, and a "pox" contamination where the quantum efficiency of individual pixels can be suppressed by 10%–40%. While the flat-field calibrations can identify many of the worst features automatically, many low-count defects remain in the data and can produce false-positive line detections.

For each pixel, we track the sky-subtracted residuals divided by the sky at that location. We then average over all observed fields (a few thousand in this case), and generate the scatter of the residuals for each pixel. We use these "residual maps" to highlight regions that have poor or variable sky subtraction. In this way, flat-field defects, charge traps, and pixel defects show up more easily. Pixel masks are then created from the visual inspection in these residual science frames. Additionally, charge traps related to a deficiency of counts are identified as vertical features in the detectors with a width of one pixel. (They can start at any y-position on the detector and either continue to the top or bottom of the frame depending on the readout direction.) A mask three pixels wide, centered on the affected line and covering the length of the defect, is then applied to the 2D fiber data frame, and propagated in the 1D flux-calibrated fiber spectra.

2.2. Large Galaxy Masks

Galaxies larger than roughly 1′ are excluded from our catalog using the positions and optically defined elliptical apertures provided by the Third Reference Catalog of Bright Galaxies (RC3; de Vaucouleurs et al. 1991) and the Uppsala General Catalogue of Galaxies (UGC; Nilson 1973). The spring field contains 644 such galaxies; the fall field, 447. For each galaxy, we use the catalogs' basic parameters for position, position angle, ellipticity, and D25 semimajor axis (i.e., the size of the galaxy defined by its B-band isophote at 25.0 mag arcsec^{-2}). Each galaxy is visually inspected through photometric imaging to confirm that these default values are reasonable. Where the parameters are uncertain, they are corrected to values listed in the NASA/IPAC Extragalactic Database[24] or through visual inspection of the galaxy in SDSS g-band images. All detections that fall within 1.5× the D25 scale of a bright galaxy's elliptical aperture are removed. This factor was determined by examining the HETDEX spectra at different scalings and ensuring that all detections related to the bright galaxy were encompassed in the aperture mask. These galaxy masks are consistently applied to the HETDEX catalog and accompanying survey area mask through the HETDEX

[24] http://ned.ipac.caltech.edu

Table 2
Statistics in Image Processing Used for Amplifier Quality Assessment

Quality Criteria	Quality Fraction	Description
im_median > 0.05	98.7%	Median counts in unprocessed amplifier image frame
$-10 <$ background < 100	98.0%	Median counts value in sky-subtracted image
$0.2 <$ sky_sub_rms < 1.5	99.3%	rms in sky-subtracted image counts
sky_sub_rms_rel < 1.5	97.7%	Ratio of sky rms in individual amplifier relative to all other amplifiers in all IFUs within the same exposure
n_cont < 35	98.7%	Number of fibers above a certain count threshold. A good indication of an amplifier saturated by a bright star or nearby galaxy.
norm < 0.5	99.3%	Relative normalization for a dithered exposure.
maskfraction < 0.2	98.3%	Rejected if more than 20% of the frame is masked

Note. Inclusion quality criteria for each image statistic, the fraction of amplifiers that pass the criteria and a short description are listed. Each HETDEX IFU is fed to four amplifiers, each containing spectral information for 112 fibers. For HETDEX survey data, we consider each amplifier to be an independent observation with its own quality criteria.

python-based, software repository `hetdex-api`[25] to provide proper survey volume accounting.

2.3. Satellites and Meteors

Both satellites and meteors generate signals that produce detections in both our emission-line and continuum emission catalogs. Meteors largely appear as line emission at multiple wavelengths and therefore contaminate our LAE samples because of their lack of strong continuum underlying the emission. Fiber spectra from HDR2 contain at least 31 meteors resulting in thousands of spatially extended emission-line detections, as the meteor travels across the HET focal plane. We use a systematic search method for these objects as part of our Emission Line Explorer software tool (`ELiXer`; Davis et al. 2023). Strong line emission appearing in just a single dithered HETDEX observation is flagged as a meteor candidate. For any observation with over 10 associated meteor candidate emission-line detections, we visually inspect the detections to confirm the presence of the meteor. We create a simple linear mask by fitting to the positions of the flagged meteor detections. This mask extends 12″ above and below the linear fit to the meteor positions; in many cases, a smaller mask could be used, but this width is needed for the brightest events. We therefore chose to be conservative with this mask. This linear mask is consistently applied to both the line emission and continuum emission raw catalogs as well as to our survey area mask.

Satellites are identified when the continuum flux density measured in the HETDEX spectral data differs from that estimated within photometric imaging data (see Figure 3 in Gebhardt et al. 2021 for an example of a satellite trail across the HETDEX FOV). Each HETDEX detection is processed with the `ELiXer` software tool, in which forced aperture photometry is performed on all available imaging. If any reported photometric measurement is more than two magnitudes fainter than the measured HETDEX value, the source is rejected as it indicates the signal is from a transient source, such as a satellite, only briefly observed at that point in the sky. Visual inspection confirmed that the majority of these detections are indeed satellites, scattered light from bright star or artifacts caused by improper flux calibration. For more discussion about finding transient sources in HETDEX, see Vinko et al. (2022).

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

3. Catalog Generation

With the HETDEX data frames reduced and verified, the data is organized into a database of flux-calibrated, 1D fiber spectra each with their own corresponding sky coordinate. In this section, we describe the steps taken to create a catalog of astronomical sources from initial object detection to final source identification. This process involves assessing data quality, as outlined in the previous sections, performing a grid search for potential object detection, then reducing the initial raw databases of potential line and continuum emission detections into high-quality detections. These two independent catalogs of high-quality detections are combined to create a single list of astronomical sources through detection grouping.

The flowchart in Figure 4 illustrates the process of producing a source classification from the raw line and continuum emission detection pipeline. This section describes the steps from detection to source object, including spectral extractions from the IFU data, detection search methods, and line parameter measurements. We also detail our method of deriving spatially resolved line fluxes for resolved sources that are applied to the low-z galaxy sample exclusively. Following this section, we describe source type identification and redshift assignment in detail in Section 4.

3.1. Object Detection

Two independent, but complementary, object detection search techniques are performed as part of the main HETDEX reduction pipeline: one to identify emission lines, the other to detect continuum sources. In the second internal data release for HETDEX (HDR2), a search was performed across 210 million flux-calibrated fibers as described in detail in Section 7 of Gebhardt et al. (2021). We briefly summarize the procedures here. During this process, no imaging preselection is used and the HETDEX data itself provides object detection. Both the emission line and continuum detection algorithms are designed to identify point-sources and account for the variable image quality, or point-spread function (PSF) of each independent HETDEX three-dither observation. To move from object detection to source classification, the outputs from both object detection methods are combined as described in Section 3.3 below.

3.1.1. Spectral Extraction

Each 51″ × 51″ IFU observation consists of 448 fibers × 1036 spectral elements × three dithers. A demonstrative example of the

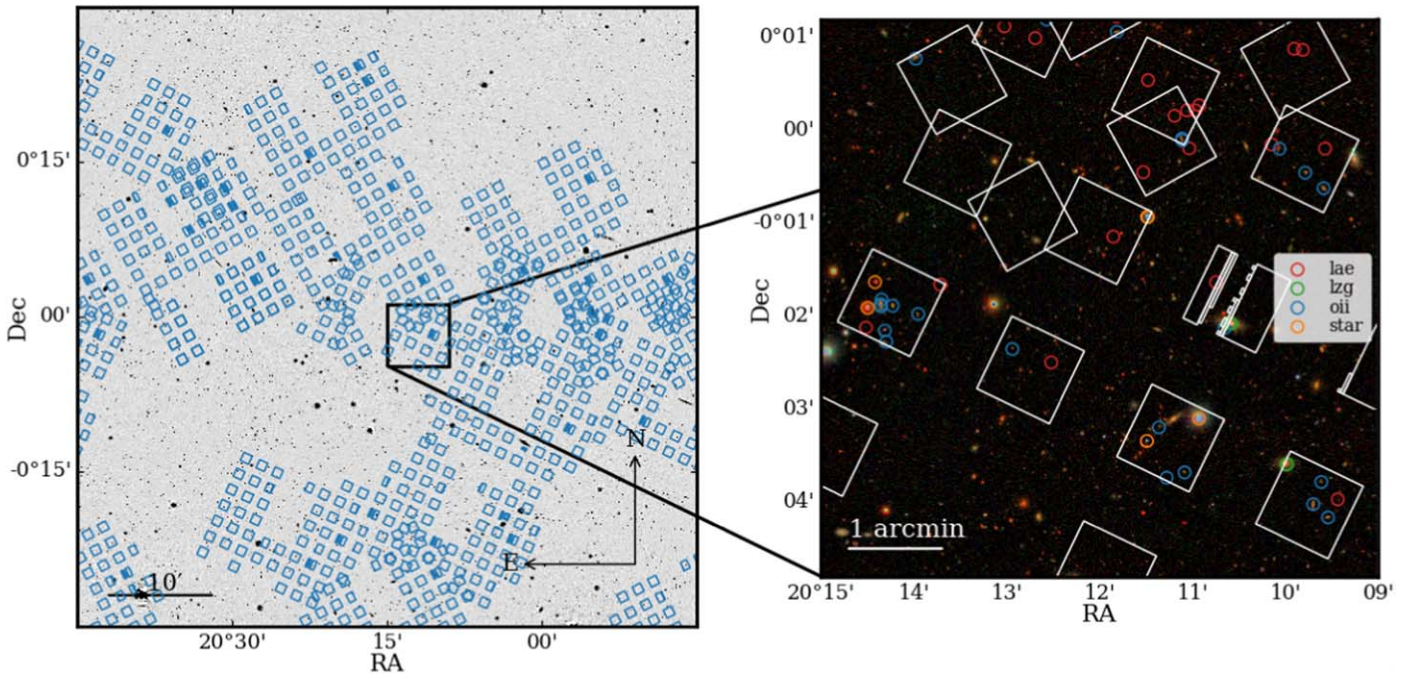

Figure 3. Left panel: observational footprint of a 1 deg^2 area of the HETDEX fall field. The blue outlines show 51″ × 51″ IFU footprints. The IFU array design results in gaps in observations with a 0.22 fill factor. Since the IFUs are oriented in the direction of the parallactic angle at the time of observation, they are oriented in differing directions. Right panel: an expanded view of a 6′ × 6′ region of the field, superposed on a color image from the DESI Legacy Imaging Surveys (Dey et al. 2019). White squares indicate the boundaries of HETDEX IFU observations. Catalog sources, identified by our classification scheme (Lyα emitter, low-z galaxy, [O II] emitter, star), are coded according to the legend. Note that one of the IFU observations was affected by a bad amplifier so that some coverage is missing in the IFU located at 20° 11′, 00° 02′.

fiber layout is found in the left panel of Figure 2. Each IFU is searched individually for line emission in a grid of 1D, PSF-weighted spectral extractions. A single fiber alone does not provide evidence for line emission; instead, we assume that the signal-to-noise ratio (S/N) of an object can best be measured in a collection of nearby fibers rather than individual fibers. We therefore use the collection of all fibers within a 3.″5 radius aperture about a candidate line. The image quality of the observation, assumed to be described by a symmetric 2D Moffat function ($\beta = 3.0$, Moffat 1969), assigns the weights to each fiber in an aperture according to the optimal extraction algorithm of Horne (1986).

An example 3.″5 radius aperture is displayed in both panels in Figure 2. The text in each circle in the right panel displays the fractional flux contribution from each individual fiber for the case of average HETDEX image quality (1.″8) with a detection centered on the central fiber. The fraction that each fiber contributes depends on the location of each fiber with respect to the aperture center, the wavelength (due to the effects of atmospheric diffraction), and the measured image quality PSF. In this extraction example, the central seven fibers contribute 80% of the extracted flux at 4500 Å.

The fiber weights are dependent on image quality. For the best HETDEX image quality (∼1.″2), a fiber centered on the source contains more than 50% of the flux; the poorest-quality HETDEX observation (∼2.″5) has 10% of the flux in the central fiber. As the HET does not have an atmospheric dispersion corrector, the weighted signal contribution from each fiber varies as a function of wavelength due to differential atmospheric refraction. Since the HET is a fixed-altitude telescope, the magnitude of the differential refraction is essentially constant for all observations. As described in Gebhardt et al. (2021), our data demonstrate that from 3500–5500 Å, a source position moves by 0.″95.

3.1.2. Line Emission Search

The initial grid search for an emission line is performed in steps of 0.″5 in the spatial direction and steps of 8 Å in the spectral dimension, guided by the simulations described in Gebhardt et al. (2021). At each grid step, a Gaussian line profile with the instrumental line width (σ = 1.7 Å) is used for the initial fit. Continuum emission is subtracted by fitting a constant intensity value to the spectrum ±50 Å around the Gaussian's central wavelength. The signal-to-noise of the line fit is measured by integrating the flux in the Gaussian model out to ±3.5 Å around the central wavelength, then dividing by the noise, which is the quadratic sum of the uncertainties in the same wavelength range. All line fits with S/N > 4.0 and $\chi^2 < 3.0$ are submitted to the next stage of line fitting to better constrain the line parameter measurements.

For the emission-line candidates identified in this first search, an optimized grid of spectral extractions is performed at a higher (0.″15) rastering resolution, with a Gaussian line width, σ, that is now allowed to be a free parameter. The location within the raster that provides the highest S/N of the emission line is assumed to be the true source position. The amplitude of the Gaussian fit then yields the measured continuum-subtracted line flux. In the case of duplicate detections (defined as emission lines lying within 3″ and 3 Å of each other), only the line fit with the highest S/N detection is accepted. The resulting line-fit parameters, including the measured observed line flux (`flux_obs`), the continuum measurement (`continuum_obs`) line width, (Gaussian σ—listed as `sigma`),

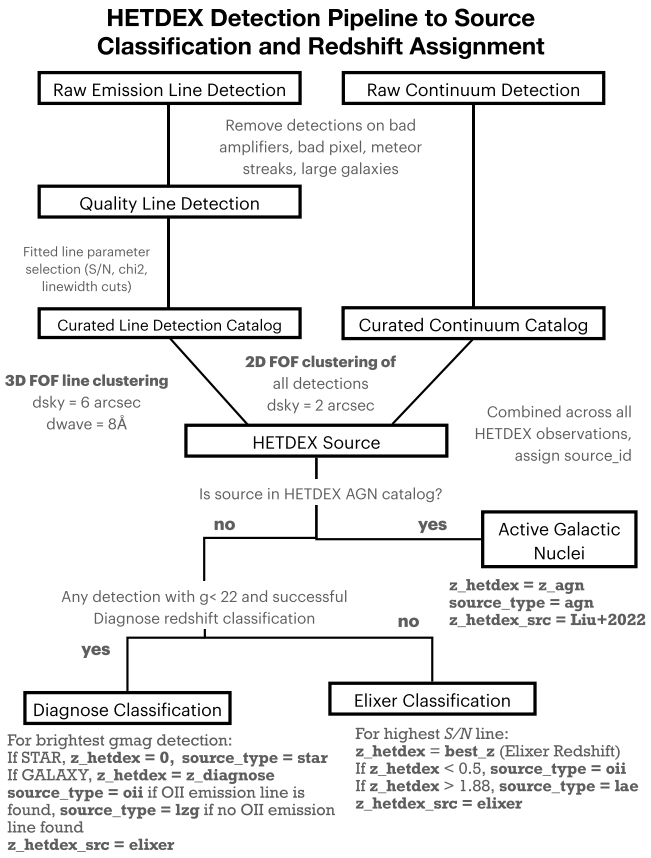

Figure 4. The steps and decisions made to generate the HETDEX source catalog from the raw detection pipelines. A HETDEX source is created from detection grouping in both 3D and 2D space through friends-of-friends (FOF) clustering to create unique sources on the sky. Multiple observations of the same source are assigned just one identifier (`source_id`). Described in detail in Section 4.2, each source is assigned a redshift and classification. If any detection has been identified as an AGN from Liu et al. (2022), it is assigned the redshift from the AGN catalog and is classified as an AGN. If a source has $g_{\rm HETDEX}$ brighter than 22 mag, the `Diagnose` classification and redshift are assigned to the source; otherwise, the `ELiXer` redshift and classification is used.

and quality of fit (χ^2—listed as `chi2`), are included in Table 6 described in the Appendix.

3.1.3. Emission-Line Fits and Criteria

The raw emission-line database is produced from all available HETDEX observations and consists of both real sources and artifacts that can arise from the data quality issues described in Section 2.1. Each raw line detection is subjected to a series of tests, which check whether the candidate line is close to a large galaxy, a meteor trail, a known detector feature, or subject to a poorly performing amplifier. A criterion is then applied, based on the S/N measured in the continuum-subtracted emission line, the Gaussian width of the fitted line in Angstroms, σ, and the quality of line fit, χ^2, measured in a $\pm 2 \times \sigma$ wide spectral window. Specifically, to be classified as a detection, a line must satisfy either

$$5.5 < \mathrm{S/N} < 6.5 \;\&\; 1.7\,\text{Å} < \sigma < 6\,\text{Å} \;\&\; \chi^2 < 1.2 \quad (1)$$

or

$$\mathrm{S/N} > 6.5 \;\&\; 1.7\,\text{Å} < \sigma < 14\,\text{Å} \;\&\; \chi^2 < 2.5 \;\&\; g > 19 \quad (2)$$

where g, hereafter labeled $g_{\rm HETDEX}$, is an equivalent broadband photometric measurement obtained by summing up the flux densities in the HETDEX spectrum, weighted by the SDSS g-band filter curve.

This combination of constraints means that high line width sources can have poorer fits (i.e., a higher χ^2) if they also have a relatively high S/N and faint $g_{\rm HETDEX}$. Sources in the high line width regime tend to suffer from a higher contamination rate, due to calibration issues or the existence of broad continuum emission from nearby galaxies and late-type stars. However, the broad-line identifications do contain interesting high-z sources, including AGNs (see the HETDEX AGN Catalog; Liu et al. 2022) and extended Lyα emitters (Mentuch Cooper et al. 2023, in preparation). We therefore allow a more liberal χ^2 quality of fit for these objects, as their lines are not typically well described by a single Gaussian line model (especially in the case of AGNs). Additional to the above criteria, narrow emission-line detections are cataloged in the wavelength range between 3510 and 5490 Å while broad-line features are only cataloged between 3550 and 5460 Å, as many spurious high line width sources were identified by the detection software on the spectral edges.

Detector artifacts are a major issue with some HETDEX spectra. In some cases, the fiber spectrum is poorly calibrated, resulting in measured continuum flux that is negative, leading to a false-positive detection. To mitigate this issue, we apply a lower cut of -3×10^{-17} erg s^{-1} cm^{-2} Å^{-1} to the local continuum measured in the Gaussian line fit. Additionally, the fiber profile quality of fit can also help identify detector artifacts. If the quality fit of the fiber profile solution, $\chi^2_{\rm fiber}$, is high, we remove the detection. In practice this value is measured for each of the five highest weight fibers in an aperture extraction, 5 Å above and below the central wavelength of the detected emission line. If any of the fibers in this spectral window have a $\chi^2_{\rm fiber} > 4.5$ or $\chi^2_{\rm fiber} > 3$ and continuum $< 0.5 \times 10^{-17}$ erg s^{-1} cm^{-2} Å^{-1}, the detection is removed from further consideration. We opted for this dual criterion because fibers with a significant continuum signal can produce higher reduced $\chi^2_{\rm fiber}$ values; we are particularly concerned with finding artifacts in the low continuum regime, where their presence can lead to false-positive LAE candidates.

Our final curated line emission detection catalog consists of 236,354 line emission detections. They can be identified in the *Detection Info Table* (see Table 6) in the column `det_type==line`. The line flux sensitivity limit for a HETDEX line emission source depends on observed wavelength, image quality, exposure time, other observing conditions, and instrument component inconsistencies, but, on average, 50% completeness is reached at roughly $\sim 7 \times 10^{-17}$ erg s^{-1} cm^{-2}. The reader is referred to Section 8.2 of Gebhardt et al. (2021) for a detailed discussion on HETDEX emission-line sensitivity and completeness with an updated discussion on these topics and the HETDEX selection function to be presented in D. Farrow et al. (2023, in preparation).

3.1.4. Continuum Emission Search

For each of the 448 fibers in an IFU, the detector counts are measured in two 200 Å regions: one in a blue region of the spectrum (from 3700–3900 Å) and one in a red region of the spectrum (from 5100–5300 Å). If either region contains more than 50 counts per 2 Å pixel on average (corresponding to $g \sim 22.5$), it is collected as a possible continuum source. The 50-count limit is arbitrary and designed to be conservative (future HETDEX catalogs reach significantly lower fluxes, as

objects can be detected more than two magnitudes fainter than this limit). Once we detect a possible source, we search about the fiber position, using a 15×15 element raster with $0\farcs1$ spatial bins. The spatial location that achieves the lowest χ^2 fit to our PSF model defines the center of the source (as opposed to the line emission search, which peaked up on the S/N of the Gaussian fit), and a point-source extraction at that position provides the detection spectrum.

Each continuum source undergoes a series of quality checks to ensure it lies on a high-quality detector and is not flagged as an artifact or satellite. The final curated continuum emission detection catalog consists of 60,907 detections. These detections can be identified in the *Detection Info Table* (see Table 6) in the column `det_type==cont`. The sensitivity of the continuum catalog is based on a photon count threshold and depends on observing conditions and exposure time. On average, the HDR2 continuum detection sensitivity is equal to $g_{\rm HETDEX} \approx 22.5$ mag. But we note that the counts threshold can be adjusted to reach fainter sensitivities down to $g_{\rm HETDEX} \approx 24$–$25$.

3.2. AGN Catalog

The HETDEX AGN catalog from the same internal data release, HDR2, is presented in Liu et al. (2022), and contains the same base sample that is included in the catalog presented in this paper. There are, however, several selection differences between Liu et al. (2022) and the current work; some of these add candidates that are not in the Liu et al. (2022) HETDEX AGN catalog, while others reject Liu et al. (2022) objects.

The catalog presented here includes additional data quality criteria that limit the sample relative to Liu et al. (2022). For example, some frames with poor observational conditions or sources on amplifiers that failed our quality assessment remain in the HETDEX AGN catalog, which mitigated these issues through visual inspection. Our sample is also limited to the HETDEX fall and spring fields, as well as the COSMOS and GOODS-N legacy fields, whereas the AGN catalog includes additional data from a North Ecliptic Pole survey (Chavez Ortiz et al. 2023, submitted).

Roughly a quarter of the AGN sample overlap with the curated line emission catalog and the continuum detection catalog; the main divergence between the catalogs arises from AGNs that exhibit broad-line emission that is not well fit by the Gaussian model implemented in the HETDEX line detection algorithm. These detections occupy both high line width and high χ^2 parameter space and do not meet the line parameter criteria for line emission in our curated line detection catalog. As described in Liu et al. (2022), visual inspection of this broad-line emission sample is essential for classifying a source as an AGN rather than a calibration artifact.

We include the AGN catalog in our combined source catalog and allow individual detections to be grouped according to the process described in Section 3.3. This approach allows for additional line and continuum emission detections to be associated with the AGN source and assigns an AGN classification and its associated redshift.

3.3. Detection Grouping

Both the line emission and continuum emission pipelines are designed to identify point-source emission. For LAEs, the primary target of interest for HETDEX, and many [O II] emitting galaxies, the point-source approximation is valid for HET image quality. However, ∼40% of the S/N > 5.5 detections identified in the emission-line and continuum source pipeline have multiple identifications. This situation can arise in extended objects, where emission is found at more than one spatial location, or with sources where more than a single emission-line has been detected. Similarly, bright astronomical sources can have both line emission and continuum emission, leading to entries in both the continuum and line curated catalogs.

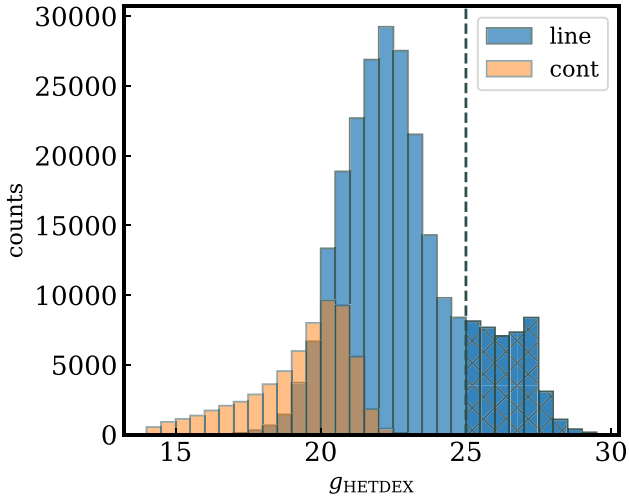

Figure 5. Distribution of emission-line detections and continuum detections as a function of $g_{\rm HETDEX}$, a pseudo magnitude calculated by integrating the extracted spectrum at the detection location, weighted by the SDSS g' filter transmission curve. At magnitudes brighter than 22.5 mag, continuum and line emission detections overlap. The bimodality in the line emission detections is due to [O II]-emitting galaxies and LAEs being the dominant samples HETDEX can detect. The vertical dashed line and hatched region at $g_{\rm HETDEX} = 25$ represents the average HETDEX continuum sensitivity limit.

The overlap in point-source brightness between the detection samples is shown in Figure 5. As expected, the continuum sources have much brighter $g_{\rm HETDEX}$ values than the objects found by the line detection algorithm. Starting near $g_{\rm HETDEX} \approx 20$, there is considerable overlap in the catalogs, while at magnitudes fainter than $g_{\rm HETDEX} \approx 22$, the line emission sources dominate. A particularly important challenge to creating a robust LAE sample is extended [O II] line emission surrounding low-z galaxies; these features can often be confused for Lyα emission due to the lack of detectable continuum emission at large galactocentric radii. To mitigate the impact of these contaminants and to properly associate extended line emission to a single emission-line source, we apply a 3D friend-of-friends (FOF) clustering algorithm to our list of line detections. Our code[26] uses `cKDtree` from `SciPy` (Virtanen et al. 2020) but is modified to use normalized coordinates in a pseudo-spherical space consisting of projected separation on sky and normalized wavelength difference. Specifically, we adopt a spatial linking length of $6''$ and a linking length in the spectral direction of 8 Å. Information for the 3D clustering of the emission-line detections is provided in the columns `wave_group_XX` of the *Detection Info Table* (see Table 6). These columns contain the identification for the wave group, its mean equatorial coordinate, and mean central wavelength. Also included is the group's semimajor `wave_`

[26] https://github.com/HETDEX/hetdex_api/blob/hdr2.1.3/hetdex_tools/fof_kdtree.py

group_a and semiminor wave_group_b axes, as determined from the line flux-weighted first-order moment of the line detection group. These values can be considered as a rough approximation to the extended group's emission-line size, but we caution that many sources that consist of just two matched line detections will be elongated in shape, while sources that have incomplete IFU coverage (as in the cases with extended [O II]-emitting galaxies) will be limited by the IFU edge of fiber coverage.

In addition to 3D clustering in wavelength and position, we also link all detections on sky together with a spatial linking length of 2″. This will ensure that sources with emission lines at multiple wavelengths will be grouped as one source. If those lines are themselves extended, then this step will group extended emission at multiple wavelengths together, as is the case for nearby galaxies that might, for example, exhibit extended [O II] Hβ and [O III] emission at multiple wavelengths. For blended objects or cases where a background object lies behind an extended foreground group, we accept that this linking may cause background detections to be lost and ultimately merged into the foreground object. Some of these sources may be quite interesting, particularly those with the potential to be gravitationally lensed, as demonstrated for the sample in Laseter et al. (2022). For fainter source groups that are ultimately classified as LAEs after detection grouping, we separate these sources spatially and assign a redshift according to each detection's observed wavelength assuming it to be Lyα. Although we note a possible exception is if the line emission wavelengths contain a pair of emission lines that can be associated with a common redshift, such as Lyα, He II, and C IV. Redshift assignment and source classification are described further in Section 4.

For each source, we select a single representative detection for each source observation. This is listed as selected_det==True in the *Detection Info Table* (columns described in Table 6) and detectid in the *Source Observation Table* (column reference is found in Table 3). This detectid corresponds to the detection member with the brightest (i.e., smallest) $g_{\rm HETDEX}$ value for all sources that are not LAEs. For LAEs, we use the highest S/N Lyα line detection as the selected representative detectid. To remove detections that are identified by the HETDEX detection pipeline due to sharp discontinuities in the detection spectrum that do not correspond to true line emission, we opt to remove all emission lines with a line width greater than 6 Å that are not identified as a representative source (with selected_det==True) or are not included in the AGN catalog.

Examples of the data (broadband images, reconstructed IFU images, and spectra) for a range of objects in the catalog are presented in Figures 6 ($z < 0.5$) and 7 ($1.9 < z < 3.3$). Individual detections are overplotted on the imaging data in the left panel in each row. Line emission detections are marked by orange crosses, and continuum detections are marked by green crosses. The spectrum for a single representative detection for the source group (identified by selected_det==True) is shown at the right. Yellow bars indicate line emission detections. In the first example in Figure 6, multiple emission lines are found, but notably a line near 5250 Å is missing from the curated catalog. This is because the other emission-lines causes the single Gaussian fit of that detection to have a poor quality of fit and does not make it to the curated catalog based on line parameter quality criteria. Continuum-subtracted emission, line flux maps, shown in the middle column, at the observed wavelength indicated in the text at the top, demonstrate differences between the line emission distribution and continuum emission morphology as shown in the imaging data on the left.

Table 3
Source Observation Table Column Descriptions

Name	Description
source_name	HETDEX IAU designation (i.e., HETDEX J123449.19 +511733.7)
source_id	HETDEX Source Identifier
shotid	integer representing observation ID: int(date+obsid)
RA	source_id R.A. (ICRS deg)
DEC	source_id decl. (ICRS deg)
gmag	($g_{\rm HETDEX}$) SDSS g-magnitude measured in HETDEX spectrum
Av	applied dust correction in the V band
z_hetdex	HETDEX spectroscopic redshift
z_hetdex_src	HETDEX spectroscopic redshift source
z_hetdex_conf	0 to 1 confidence HETDEX spectroscopic redshift source
source_type	options are star, lae, agn, lzg, oii, and none
n_members	number of detections in the source group
detectid	detection ID of representative detection for the source (selected_det==True in Table 6)
field	field ID: cosmos, goods-n, dex-fall, dex-spring
flux_aper	Dust-corrected, O II line flux measured in elliptical galaxy aperture in 10^{-17} erg s^{-1} cm^{-2}
flux_aper_err	error in flux_aper
flag_aper	1 = aperture line flux used for lum_oii, 0 = PSF-line flux used from "flux" column
major	major axis in arcseconds of aperture ellipse of resolved O II galaxy defined by imaging
minor	minor axis in arcseconds of aperture ellipse of resolved O II galaxy defined by imaging
theta	angle in aperture ellipse
lum_lya	Lyα line luminosity and error calculated from "flux" column (i.e., dust-corrected Lya line flux) in rg s^{-1}
lum_lya_err	error in lum_lya
lum_oii	O II line luminosity calculated from "flux" column if flag_aper = 0 or 'flux_aper' column if flag_aper = 1 in erg s^{-1}
lum_oii_err	error in flux_oii
flux_lya	Lyα line flux calculated from "flux" column for (i.e., dust-corrected Lya line flux) in erg s^{-1}
flux_lya_err	error in flux_lya
flux_oii	[O II] flux in erg s^{-1} calculated from "flux" if flag_aper = 0 or "flux_aper" if flag_aper = 1 (i.e., dust-corrected)'
flux_oii_err	error in flux_oii
sn	signal-to-noise for line emission
apcor	aperture correction applied to spectrum at 4500 Å

3.4. Spatially Resolved Line Fluxes

For every galaxy with z_hetdex < 0.5, we measure spatially resolved [O II] line fluxes at the galaxy's redshift in addition to those provided by the HETDEX line detection algorithm that are point-source, PSF-weighted line flux values. We note that while a resolved line emitter can appear as multiple detections in the line database, some flux will inevitably be missed even if each detection is summed. In addition, the line detection pipeline used for this catalog release contained an upper limit on the continuum value, so some very bright line emitting galaxies are completely missing from our

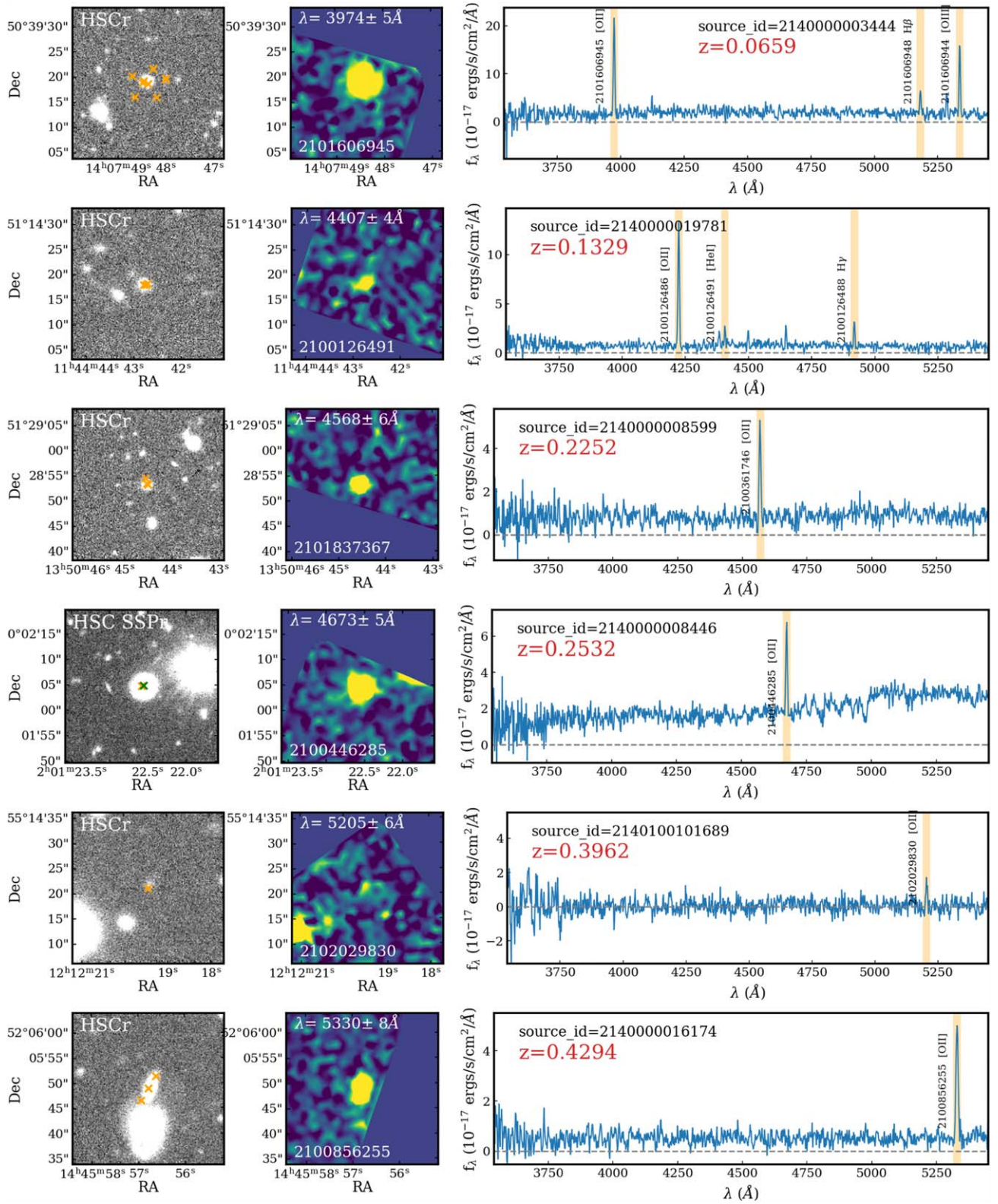

Figure 6. Example low-z sources in the catalog in order of increasing redshift. For each object, the left panel displays a 30″ × 30″ HSC-r band image; the middle panel is a wavelength-collapsed, continuum-subtracted, flux map of line emission (same area and orientation as left panel), at the HETDEX detected line indicated in the text. Solid blue in the line flux map indicates an area that is not covered by an IFU (i.e., no HETDEX data exists). The right panel presents the HETDEX spectrum for the best detection for the source (indicated by `selected_det==True` in Table 6 described in the Appendix). Text on this panel indicates the `source_id`, the individual emission-line detections (as indicated by the vertical `detectid` text), and the HETDEX redshift, `z_hetdex`. Multiple detections can often arise from a single source if they possess both continuum and line emission. Extended emitters can appear multiple times in the detection catalog, as can be seen in the top two examples. Note that some emission lines are missed from the catalog in the second spectrum from the top where a single Gaussian model cannot sufficiently identify Hδ. The position of each detection is indicated in the left-hand column images as orange and green crosses, for line and continuum detections, respectively.

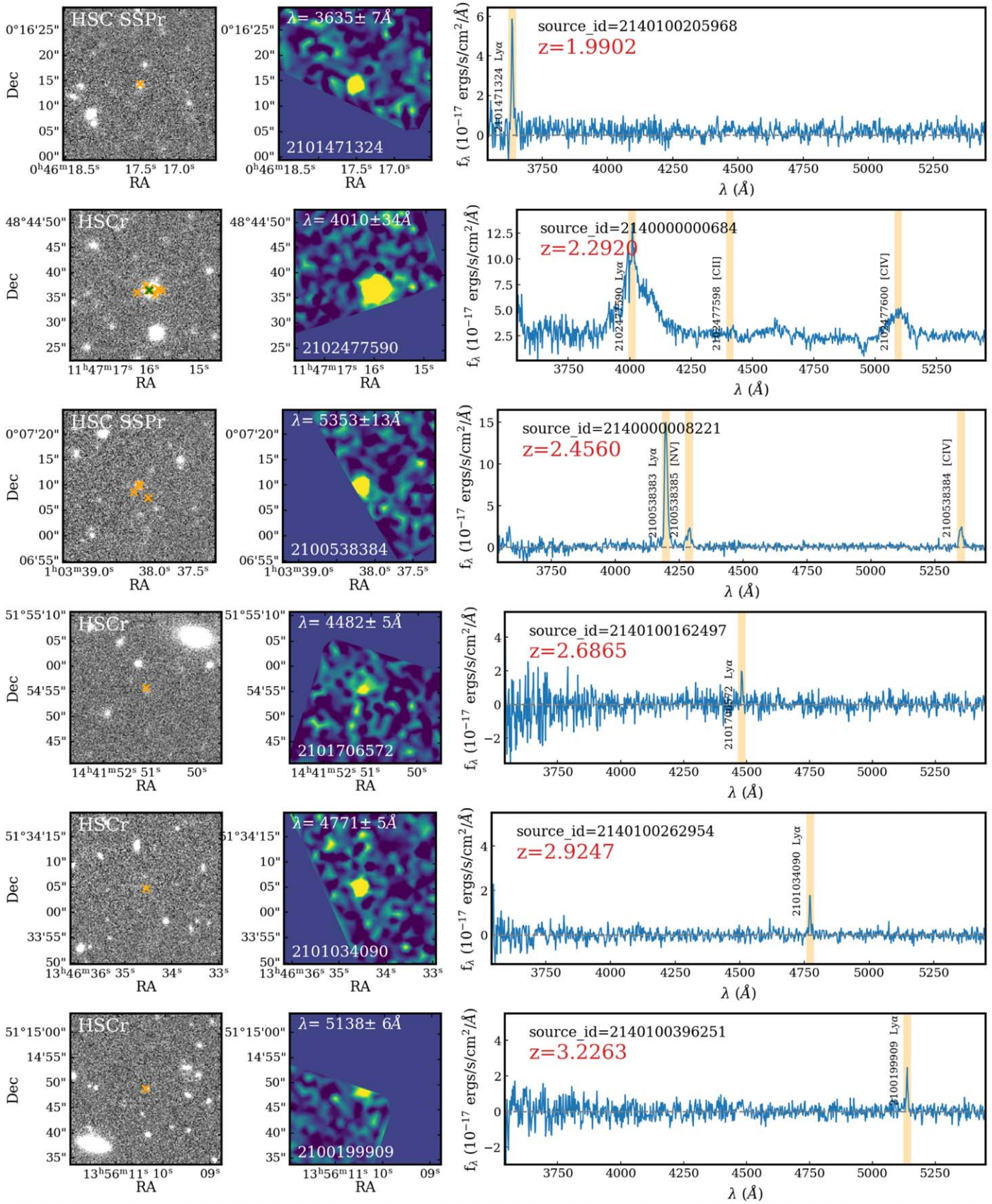

Figure 7. Examples of high-z HETDEX sources in increasing redshift order. The panels are described as in Figure 6. The second row is an AGN exhibiting broad-line emission; the third row is an AGN with narrow-line emission. Both of these objects exhibit multiple line emission detections as shown by the orange crosses in the left image. The broad AGN is also a continuum emission detection source (indicated as a green cross in the left image). The remaining objects in the figure are typical LAEs in the catalog, i.e., pure line emission sources, and are frequently found without an imaging counterpart.

line emission database, even though they are found in the HETDEX continuum catalog.

A major strength of the wide-IFU (dithered) coverage with HETDEX is that the observations automatically produce an emission-line map of resolved galaxies. However, due to the IFU layout in the HET's focal plane, many of these systems have incomplete coverage, as their light extends off the edge of their IFU. The angular resolution of our imaging observations is substantially better than the IFU fibers. As a result, object shapes and sizes are better measured from direct imaging.

Object shapes are automatically included in HETDEX's `ELiXer` classification tool (Davis et al. 2023), as it applies `Source Extraction and Photometry` (`SEP`; Barbary 2016) to all available broadband imaging at the location of each HETDEX detection. This step provides the major and minor axes of an ellipse fit to the second-order moment of each object's surface brightness distribution (Bertin & Arnouts 1996). We use the `ELiXer` catalog `selected=True` option and preferentially choose *r*-band over *g*-band measurements to define each galaxy's elliptical aperture. In general, the *r*-band imaging we have obtained has a fainter limiting magnitude, and significantly better image quality. The image selection can be found in the columns `catalog_name_aper` and `filter_name_aper` in Table 6. Elliptical parameters for each low-*z* galaxy can be found in both Tables 3 and 6 under the columns `major`, `minor`, and `theta`. Additionally the aperture center and the measured continuum aperture magnitude are in Table 6 in the columns `ra_aper`, `dec_aper`, `mag_aper`, and `mag_aper_err`.

Figure 8 presents the major axis distribution of the low-*z* sample. More than three-quarters of the sample has a major axis larger than 3″ and is thus spatially resolved by VIRUS. We create continuum-subtracted line flux and flux uncertainty maps for each source's [O II] emission by summing the fiber data in a ±15 Å window around the wavelength of observed [O II], redshifted from λ3727.8 Å according to `z_hetdex`. We then subtract local spectral continuum by making two narrowband-like images, each 50 Å wide, shifted by an additional 10 Å blue and red of the line emission. We subtract the average of these two images from the line flux map to produce a continuum-subtracted line-flux map.

The flux and associated error in the galaxy's elliptical aperture is summed using the `photutils` software package (Bradley et al. 2021). The resulting aperture and dust-corrected aperture fluxes are found in columns `flux_aper` and `flux_aper_err`. The photometric information has an aperture correction applied, `im_apcor`, for sources that lack full IFU coverage. For each source, we opt to use the `flux_aper` value for `flux_oii` if the value is positive and the major axis of the low-*z* galaxy is greater than 2″.

3.4.1. Comparison to SDSS [O II] Line Fluxes

Since the HETDEX survey fields lie completely within the SDSS footprint, we can compare spectra that are in common between the two surveys. Figure 9 presents HETDEX measurements of the continuum-subtracted emission-line fluxes for [O II] emitting galaxies found in the MPA-JHU value-added catalogs from SDSS DR8[27] (based on the methods described in Brinchmann et al. 2004; Tremonti et al. 2004). In the top-left, for optimal comparison, the HETDEX line fluxes are measured

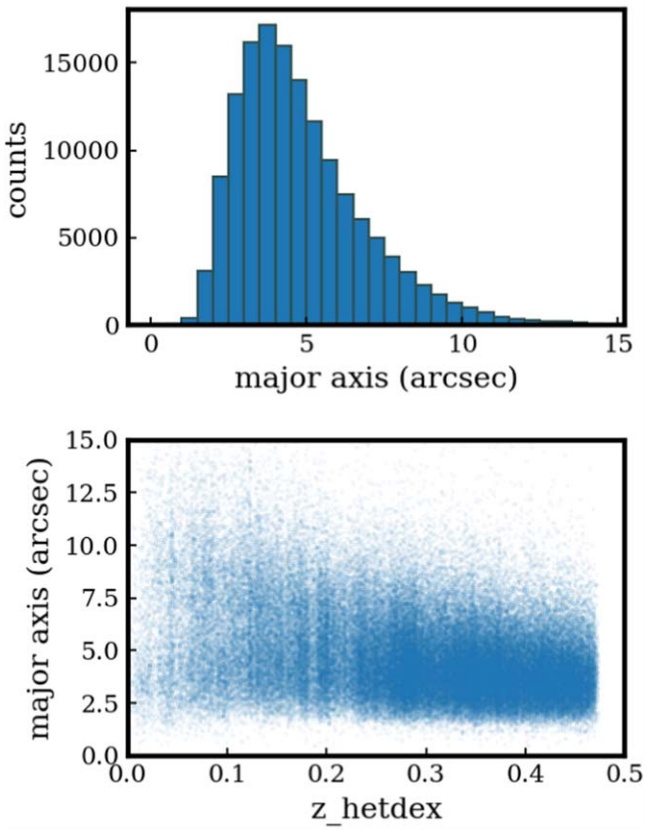

Figure 8. Distribution of sizes for the $z < 0.5$ galaxy sample as derived from object detection using Source Extraction and Photometry (SEP) on ancillary *r*-band or *g*-band photometric imaging. For the low-*z* galaxy sample, 78.5% of the galaxies have sizes greater than 3″, stressing the need for aperture flux values that encompass a galaxy's full extended emission.

in a circular 3″ diameter aperture to match the 3″ diameter fibers of SDSS. The figure demonstrates that for forced aperture line fluxes, the HETDEX measurements are well matched to SDSS 3″ values to an rms of 26% for objects with line fluxes above of $\sim 3 \times 10^{-16}$ erg s^{-1} cm^{-2}. Differences in positioning are mitigated by placing apertures exactly at the location quoted by SDSS; however, the derived line fluxes also depend on how the continuum is measured, and our IFU data have a different spatial profile than that of the single SDSS fiber.

Comparisons between the pipeline point-source fluxes show much greater scatter in the top-right panel in Figure 9. As with the forced line fluxes, differences in the measurement method can create scatter, but here the bigger culprit is positioning. The HETDEX detection pipeline is designed to peak on the highest S/N line detection in a spatial grid of 1D extracted spectral data. This peak can vary from the location of the SDSS fiber by up to 1″. The bottom-right panel shows the distribution of sky separations between the SDSS fiber and the position of best HETDEX peak detection. In cases where the SDSS fiber is on the edge of the IFU, some flux is lost and underestimated in the pipeline point-source fluxes. In addition, fewer data points are shown here because the continuum detection search was performed with an upper limit threshold on counts, thus excluding the brightest [O II]-emitting galaxies.

The middle-left panel compares HETDEX aperture fluxes, `flux_aper`, with SDSS measurements. Here, the HETDEX values lie significantly above those of SDSS due to the larger aperture area. The middle-right shows the optimal HETDEX

[27] https://www.sdss.org/dr16/spectro/galaxy_mpajhu/

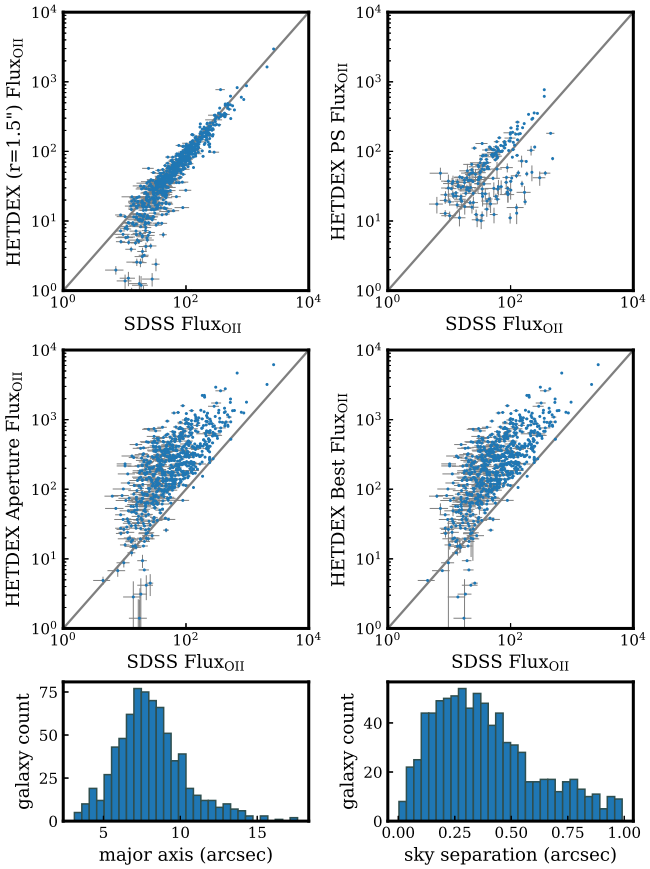

Figure 9. HETDEX [O II] line fluxes compared to SDSS [O II] line fluxes. All units in 10^{-17} erg s^{-1} cm^{-2}. The top-left panel is a comparison of 3″ diameter aperture [O II] line flux measurements from HETDEX in which the continuum-subtracted [O II] line fluxes within a 3″ diameter aperture are measured at the exact location of the SDSS fibers. The top-right panel shows the HETDEX pipeline continuum-subtracted [O II] line fluxes, which are measured using PSF-weighted extracted spectra at the location of the object's peak [O II]. The middle-left panel compares HETDEX aperture line fluxes with the SDSS. Here, the HETDEX values lie significantly above those of SDSS due to the larger aperture area. The middle-right shows the optimal HETDEX [O II] line flux, which can be either from the HETDEX pipeline or the aperture flux measurement, depending on conditions. The range of major axis sizes for the HETDEX measurements is shown in the bottom-left panel and separations between HETDEX catalog detections and SDSS fiber positions is shown in bottom-right.

[O II] line flux, `flux_oii` in the catalog, which can be either from the HETDEX pipeline or the aperture flux measurement. It is assigned `flux_aper` if it is a positive value and the major axis of the galaxy based on broadband imaging is greater than 2″. Otherwise the line flux measured comes from the `flux` measured from the HETDEX pipeline. The flag listed as `flag_aper` is 1 if `flux_aper` is used, 0 if `flux` is used, and -1 if it is not relevant, as is the case for LAEs, AGN, stars, and low-z galaxies (LZGs). As shown in the bottom-left of Figure 9, the galaxies in the comparison have a wide range of sizes, so it is not surprising that the HETDEX spatially resolved fluxes are much larger than the fiber fluxes from SDSS.

3.5. Dust Correction

Reported fluxes are provided both as measured at the top of the atmosphere, and corrected for Galactic extinction. The python software package `dustmaps` (Green 2018) is employed to access the local Milky Way dust reddening values for each source's coordinates as measured by Schlegel et al. (1998). The software returns the locally measured color excess value, $E(B-V)$, based on a source coordinate. We assume the ratio of V-band extinction, A_V, to color excess, $E(B-V)$, to be $R_V = 3.1$ and apply a factor of 2.742 to measure the local V-band extinction as $A_V = 2.742 \times E(B-V)$ according to the re-calibration using SDSS stars of the Schlegel et al. (1998) maps by Schlafly & Finkbeiner (2011). A_V values range from 0.01–1.44 with a median value of 0.04. Dust correction of line fluxes is applied at the central wavelength of the line emission according to the $R_V = 3.1$ extinction curve of Fitzpatrick (1999), implemented using the open-source python software `extinction`.[28] The measured values are designated with the notation `XXX_obs`, where `XXX_obs` can be `flux_obs`, `flux_obs_err`, `continuum_obs`, or `continuum_obs_err` without any dust correction, while those without the `obs` suffix have the local dust correction applied. Included source spectra are also offered with and without an applied dust correction. Their format is described further in Section 5.

4. Source Classification and Redshift Determination

Once emission-line and continuum detections are placed into common source groups, we assign a source classification and redshift to each group. As the two HETDEX detection methods probe different astronomical sources, we must take a multi-pronged approach to classify our sample. In this section, we first outline the three methods of source classification and redshift assignment (Section 4.1) and then present the decision logic to assign a classification to a source (Section 4.2). In Section 4.3 we compare our measured redshifts to those available in the literature and quantify the accuracy of our redshift measurements.

4.1. Methods

The continuum sample comprises objects in the magnitude range of $14 \lesssim g \lesssim 21.5$. In contrast, the line-emission sample probes a broad range of continuum levels, with a quarter of the curated, high-quality sample having $g_{\rm HETDEX} > 25$, which we consider to be the approximate sensitivity limit of HETDEX 1D spectra.

The line emission algorithms probe a wide range of sources, which include galaxies with no detectable continua and bright objects with multiple emission-line entries in our detection catalog. Depending on the choice of line-fit parameters, the detection algorithms also include high line width detections that are actually sharp discontinuities in the spectra, caused by absorption features in late-type stars, while others are due to the broad, complex emission lines of AGNs.

For the brighter continuum spectra, we employ the software package `Diagnose` to determine a source's classification and redshift (see Section 4.1.1). For fainter objects, we rely on the properties of the line emission and assumptions about the expected luminosity function and equivalent width distribution of line emitting sources to assign a redshift; this process is described in detail in Davis et al. (2023) and briefly summarized in Section 4.1.2. If any source contains a detection that is found within the HETDEX AGN Catalog, then the

[28] https://github.com/kbarbary/extinction

redshift and classification from Liu et al. (2022) is applied, as detailed in Section 4.1.3.

4.1.1. Diagnose

Diagnose, a software package developed for the HET VIRUS Parallel Survey (HETVIPS; G. Zeimann et al. 2023, in preparation), uses a principal component analysis algorithm to classify sources as stars, galaxies, quasars, or unknown. The redrock[29] spectral templates used by Diagnose are the same as those employed by SDSS-IV (Ross et al. 2020) for their classification/redshift measurements. The templates include spectra from 10 galaxies, four quasars, and three cataclysmic variables. Stars are classified by spectral type, and are assigned to the subclasses B, A, F, G, K, M, and white dwarf. G. Zeimann et al. (2023, in preparation) report that for objects with both SDSS and HETVIPS classifications, the Diagnose values match those of SDSS for 96.9%, 94.7%, and 92.3% of stars, galaxies, and quasars, respectively.

Unsurprisingly, the fraction of sources that achieve a successful Diagnose classification and redshift assignment decreases as a function of VIRUS spectral signal-to-noise, and is correlated with a source's g-magnitude. G. Zeimann et al. (2023, in preparation) demonstrated that they reached ∼90% recovery of classifications at a spectral $\langle S/N \rangle = 8$, where $\langle S/N \rangle$ is the mean S/N measured per 2 Å spectral resolution element, and a value of 8 corresponds roughly to $g = 20$. For our sample of sources with $g_{\rm HETDEX} < 22$, Diagnose reports a confident classification for 98.5% of the detections. At brightnesses in the range of $22 < g < 23$, classifications are reported for 86.4% of the detections. However, we do not rely on Diagnose at these fainter magnitudes because of possible confusion between [O II] and Lyα line emission. Often these sources have little detected continuum signal, causing Diagnose to automatically default to a low-z star-forming template with a significant amount of Lyα-emission being falsely identified as [O II], Hβ, or [O III] emission. Faint line emission classification is better assessed by ELiXer.

4.1.2. Emission Line eXplorer (ELiXer)

The majority (≳60%, although this number is much higher when considering lower signal-to-noise detections) of HETDEX line emission detections consist of just a single emission line, and line identification cannot be trivially deduced from the spectrum itself. For LAEs, the largest contaminant is $z < 0.5$ [O II] emitting galaxies. Historically, a 20 Å equivalent width cut (in the rest frame of Lyα) has been used to segregate [O II] from Lyα (Gronwall et al. 2007; Adams et al. 2011) where the continuum is measured from either the spectrum itself (if sensitive enough) or in accompanying deep photometric imaging. In practice, this criterion typically results in more than 4% contamination, and excludes all lower equivalent width Lyα lines (Acquaviva et al. 2014). For HETDEX, this can be a problem as the $H(z)$ and $D_A(z)$ measures are sensitive to interloper clustering (Leung et al. 2017; Grasshorn Gebhardt et al. 2019; Farrow et al. 2021), and HETDEX requires contamination in the LAE sample to be ≲2% (Gebhardt et al. 2021). Leung et al. (2017) improved upon the 20 Å cut by adopting a Bayesian approach and including additional information about the equivalent width distributions of Lyα and [O II] using g- and r-band photometric info, and the systems' emission-line luminosity functions. From their modeled data, Leung et al. (2017) reported an expected contamination rate of Lyα by [O II] of between approximately 0.5% and 3.0% at a cost of ∼6.0%–2.4% lost LAEs. HETDEX implements this line discrimination approach and builds upon it in its line emission classifying software Emission Line eXplorer (ELiXer; Davis et al. 2023). It adds in a suite of additional information such as multiple line emission considerations, photometric imaging counterpart information (galaxy size and magnitude, for example) and additional data quality checks that assign a probability likelihood, P(Lyα), that a HETDEX emission-line detection is due to Lyα. This value, plya_classification, and a number of other measurements from ELiXer related to the detection's imaging counterpart are presented for each detection in Table 6 (described in the Appendix).

Davis et al. (2023) report a projected HETDEX LAE contamination rate from [O II] of 1.3% (±0.1%) and an additional 0.8% (±0.1%) from all other sources, along with an LAE recovery rate of 95.7% (±3.4%) with ELiXer version 1.16.5 and the current internal HETDEX catalog (based on its third internal data release). For the work in this paper, we use an earlier ELiXer version (1.9.1) and find an LAE contamination rate from [O II] of 2.4% with an LAE recovery rate of 95.2% for galaxies with $g_{\rm HETDEX} > 22$ mag. These rates do not include the bias in the spectroscopic sample used to measure the contamination rate and recovery rates; this sample tends to be brighter than the main HETDEX LAE sample. For details on projecting these values to unbiased rates, please see Davis et al. (2023).

4.1.3. AGN Catalog Redshifts

As discussed in Section 3.2, a systematic search for AGNs within HDR2 is performed to identify both broad-lined emitting AGNs and narrow-line AGNs with two confirmed emission lines. This AGN catalog (Liu et al. 2022) consists of 5322 AGNs, of which 3733 have spectroscopic redshifts secured by either (1) two emission-line confirmations and/or (2) a positional match to AGNs within the SDSS DR14 Quasar Catalog (Paris et al. 2018). These sources are identified with zflag=1 in Liu et al. (2022) and are identified in this catalog release by agn_flag=1. The remaining single broad-lined sources are assumed to be due to Lyα emission from AGNs and are identified in the catalog by agn_flag=0. Sources that are not AGNs are given agn_flag=-1.

4.2. Assignment

A sequence of logic is implemented to assign a redshift and classification to each source. We highlight the general logic in the bottom part of the flowchart presented in Figure 4 and describe the details here.

The method of the assigned redshift is found in the column z_src_redshift. If any detection in the group is found in the HETDEX AGN catalog, the redshift from Liu et al. (2022) is assigned to the source group, the source type is labeled agn, and the source's redshift confidence is taken from the z_flag column from the HETDEX AGN Catalog. This value is 1 if either the redshift is derived from multiple emission lines or the object has an SDSS counterpart with a measured redshift consistent with the HETDEX observations. A small (2.2%)

[29] https://github.com/desihub/redrock-templates

fraction of our catalog (4976/223,641) is assigned its redshift from the HETDEX AGN Catalog and can be found in the catalog under z_src_redshift==Liu+2022.

For source groups that contain one or more detections with $g_{\rm HETDEX} < 22$ mag and a Diagnose classification, we adopt the Diagnose redshift, z_diagnose. We assign a confidence to this redshift of z_conf=0.9 (an arbitrary high-confident number here, but we aim to provide better calibration of our redshift assignment in the future) as redshifts assigned from Diagnose are highly reliable, with a 97.1% accuracy for our sample (described further in Section 4.3). If the detection has a STELLAR classification, we assign source_type=star and z_hetdex=0. We note that additional classification information can be found from the Diagnose spectral fits in columns (z_diagnose, cls_diagnose, stellartype in Table 6 (described in the Appendix). If the Diagnose classification is GALAXY, we label the source type as oii if an [O II] emission line is present in the spectrum with a line flux value, flux_oii, reported, and lzg (low-z galaxy) if no emission line has been detected. If the Diagnose classification is QSO, we assign a source type of agn. Less than half (41.1%, 91,885/223,641) of the catalog redshifts are assigned using Diagnose and can be isolated in the catalog under z_src_redshift==Diagnose.

For all other source groups, we rely on ELiXer to assign source redshifts. For the public HDR2 catalog that is limited to higher S/N > 5.5 detections, 60.7% (135,789/223,641) sources are classified by ELiXer. A few steps of logic are involved when making the final selection, which we briefly outline here. For a single emission-line source group, we simply assign the ELiXer redshift, best_z, to z_hetdex. We also transfer the ELiXer redshift confidence, best_pz (described in detail in Davis et al. 2023) to z_conf. The redshift confidence should not be used for selection criteria, however, as we have not calibrated it. This will be applied in later HETDEX catalogs.

If multiple line detections are found, we first check to see if any of the detections are part of a common wavelength linked group. We then use the redshift for the detection closest to the center of the wave group (listed as the minimum value of src_separation for the detection group). Next we check to see if any of the detections are confidently at low-z, with plya_classification < 0.4. This value is empirically chosen to maximize the LAE recovery fraction (96%), while minimizing the [O II] contamination fraction (at 3%). It also differs from the built-in threshold (plya_classification = 0.5) that ELiXer users for its redshift assignment, as this earlier software version was found to put low-quality LAE candidates at 0.4 < plya_classification < 0.5. If this is the case, we assign the ELiXer best_z to the detection closest to the source group center. This can result in background line emitters getting blended with the foreground source and ultimately not classified as an LAE or more distant galaxy. If neither of these cases are found, we go through the source group assigned redshifts and make a choice of which is the best redshift to use.

If the collection of redshifts has a standard deviation less than 0.02, then we can simply assign the redshift of the detection closest to the source center. This will happen if the source is an extended Lyα-emitter or if observed emission lines are a pair match such as Lyα and C IV. Extended Lyα-emitters will be analyzed in a future HETDEX paper but can be found in this catalog by searching for sources with a defined wave_group_id at z_hetdex >1.88. Both AGN and LAE source types (e.g., via a logical search of source_type = =lae or agn) can exhibit extended emission. If the standard deviation of z_hetdex in the detection grouping is larger than 0.02 and all detections are classified as high-z according to ELiXer's best_z, then we assume the detections to be independent from each other and be line-of-sight interlopers. We disassociate the group of detections and assign each detection a unique source_id, and classify each detection as an LAE at redshift corresponding to Lyα the observed line wavelength. Sources that have been assigned their redshift from ELiXer are found in Table 6 under z_src_redshift=='elixer'.

Examples of the source classification and redshift assignment for a range of objects in the catalog are presented in Figures 6 ($z < 0.5$) and 7 $1.9 < z < 3.3$).

4.3. Accuracy

To assess the accuracy of our redshift assignments, we can compare our cataloged values, z_hetdex, to spectroscopically determined redshifts from other surveys. These literature redshifts are generally quite reliable, as they tend to be derived from spectra with higher spectral resolution, and/or broader wavelength coverage than that from HETDEX. However, because these targets were often pre-selected from broadband imaging data, they also tend to have continuum magnitudes significantly brighter than the bulk of the HETDEX detections.

In the COSMOS legacy field, we use redshifts from zCOSMOS DR3 (Lilly et al. 2009) and the DEIMOS 10k Sample (Hasinger et al. 2018); in the GOODS-N legacy field, the redshifts come from Reddy et al. (2006), Barger et al. (2008), Wirth et al. (2004, 2015), PEARS (Ferreras et al. 2009), and DEEP3 (Cooper et al. 2011). Redshifts are also used from MOSDEF (Kriek et al. 2015) and 3D-HST (Momcheva et al. 2016), which cover multiple deep legacy fields. Finally, we also use measurements from the SDSS DR16 Redshift Catalog (Ahumada et al. 2020), which covers brighter objects over all of our fields. For all of these data, we apply quality criteria to select the surveys' most confident redshifts as described in their corresponding papers.

Shown in Figure 10, 4675 of our catalog sources have a spectroscopic redshift match within 1.″2. The source type breakdown of the catalog-matched sources is 134 LAEs, 1592 AGNs, 2560 [O II]-emitting low-z galaxies, 311 other LZGs without [O II] line emission, and 78 stars. The accuracy of our combined redshift assignment method is 96.1% (88.4%), where accuracy is defined as agreement with an external spectroscopic redshift value to within $\Delta z < 0.02$ ($\Delta z < 0.005$). Restricting the matches to those sources with redshifts assigned from the HETDEX AGN Catalog, 98.2% (1564/1592) are in agreement. The redshift assignment for AGNs, however, included cross-matches to AGN spectroscopic redshifts from SDSS DR14 (Paris et al. 2018), which were adopted if they agreed with line emission measured in the HETDEX data. Removing the AGN population from the catalog and considering just Diagnose and ELiXer redshift assignments, the net redshift accuracy to within $\Delta z < 0.02$ is 95.0% (2929/3083). The sample assigned redshifts by Diagnose are 95.7% (1994/2084) accurate; this result is slightly higher than that reported by G. Zeimann et al. (2023, in preparation) because we have excluded the poorer performing AGN/Quasar assignments, which are affected by narrower wavelength coverage of VIRUS. The remaining

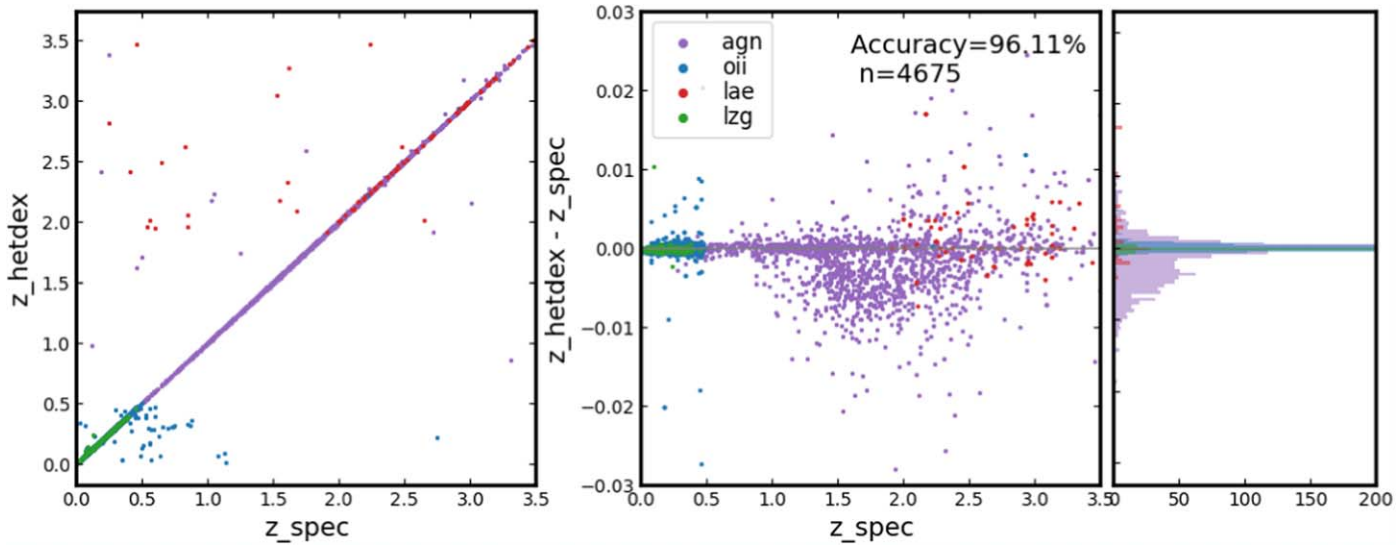

Figure 10. In the left panel, a comparison plot of HETDEX spectroscopic redshifts, z_hetdex, and spectroscopic redshifts, z_spec, from multiple publicly available redshift catalogs is shown. In the right panel, a zoom-in on redshift differences for well-matched sources is shown. The histogram in the far-right panel shows the distribution of differences. Our net accuracy of spectroscopic redshift agreement to within $\Delta z < 0.02$ is 96.11%.

redshifts assigned by ELiXer have a relatively poorer redshift accuracy of 93.6% (935/999) primarily because ELiXer is the catchall for all of the remaining sources that are not bright enough for Diagnose nor are they visually classified as an AGNs. False-positive line-emission detections, which generally do not have associated continuum emission, fall into this category, as do transient detections (such as missed meteor and satellite features), and local line emitters with little continuum (e.g., young stellar objects, active late-type stars, and planetary nebulae). Finally, the overall density of faint line emitters is considerably larger than that for bright continuum sources, so some line-of-sight mismatches with external redshift catalogs are likely.

One approach to measure the true success rate of identifying [O II] and Lyα sources and mitigate possible contamination from false positives and other interlopers is to only consider line detections at observed wavelengths that match the spectroscopic redshift from the external catalogs at either the observed frame of [O II] or Lyα. This requirement increases the accuracy of ELiXer to 95.24% and the accuracy of HETDEX redshift assignments to 96.8%. For greater in-depth discussion on assigning classifications with ELiXer and its success rates, see Davis et al. (2023).

5. Catalog Format

The information in this release is presented in two separate catalog formats: the *Detection Info Table* (with columns described in Table 6), which contains information about every HETDEX line emission and continuum detection that has passed the quality checks and line parameter criteria described in Sections 2.1 and 3.1.3, respectively; and the *Source Observation Table* (with columns described in Table 3), which contains aggregate information from the more detailed Table 6 for each source observation. It contains fundamental information on a source (position, redshift, physical size if relevant, [O II] or Lyα flux, and luminosity where appropriate) and is repeated for each separate HETDEX observation of the source. For most users, Table 3 will be sufficient, and it is a limited, easier-to-parse summary of Table 6 (which is provided in the Appendix).

A HETDEX source, identified by source_id, is a collection of all detections at the same on-sky position combined through the detection grouping method described in Section 3.3. If the source is observed more than once, its source_id and source_name remain the same, but the observation ID (shotid) will be different as will the reported catalog measurements. We report a single representative detection identifier, detectid, which may be matched to Table 6 for each source observation in the detectid column; this column corresponds to the detection member with the brightest (i.e., smallest) $g_{\rm HETDEX}$ value for all sources that are not LAEs. For LAEs, we use the highest S/N Lyα line detection as the selected representative detectid. A user may search Table 6 for this representative detect ID by selecting the column selected_det==True.

For the [O II] and Lyα line fluxes, we provide the columns flux_oii and flux_lya and corresponding error columns for sources identified as lae and oii. As discussed in Section 3.4, for each low-z galaxy, an aperture [O II] line flux is measured: flux_aper at z_hetdex. This flux is assigned as the source's flux_oii if it is a positive value and the major axis of the galaxy based on broadband imaging is greater than 2″. Otherwise, the line flux measured comes from the flux measured from the line fit to the extracted spectrum of the brightest detectid in the source group. Line fluxes and associated errors are converted to intrinsic [O II] and Lyα line luminosities using our best measured redshift, z_hetdex, and the cosmology defined by Planck Collaboration et al. (2020).

The following files are included in this release:

1. The *Source Observation Table* (column descriptions in Table 3):
   hetdex_sc1_vX.X.dat/.ecsv.
   This table consists of one row per source observation. For each source observation, it provides the source's J2000 equatorial coordinates, and redshift (z_hetdex). Every source is classified into one of the following source_type options: lae, oii, agn, lzg, or star as

Table 4
Format of the *Source Observation Spectra* (Table 3 + Source Spectra) FITS File

HDU No. and Name	Type	Dimensions	Format	Description
0:PRIMARY	PrimaryHDU			Empty
1:INFO	BinTableHDU	7367R × 27C		Source information for each catalog source, one row per source observation
2:SPEC	ImageHDU	(1036, 232,650)	float32	Dust- and aperture-corrected, PSF-weighted 1D spectrum at `detectid` in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}.
3:SPEC_ERR	ImageHDU	(1036, 232,650)	float32	Uncertainty in SPEC 10^{-17} erg s^{-1} cm^{-2} Å^{-1}.
4:SPEC_OBS	ImageHDU	(1036, 232,650)	float32	Aperture-corrected, 1D PSF-weighted spectrum at `detectid` in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}.
5:SPEC_OBS_ERR	ImageHDU	(1036, 232,650)	float32	Uncertainty in SPEC_OBS in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}.
6:APCOR	ImageHDU	(1036, 232,650)	float32	Aperture correction applied to spectrum and catalog flux values
7:WAVELENGTH	ImageHDU	(1036)	float32	Wavelength array from 3470–5540 Å in 2 Å bins

Note. These contain Table 3 (also available in a simple .dat/.ecsv ASCII format) in HDU 1. The aperture correction value is applied to the spectra. The aperture correction is the fractional fiber coverage of the $r = 3\farcs5$ aperture centered on the detection identified in the column `detectid`.

described in Section 4.2. For sources with either Lyα or [O II] line emission, the table provides the optimal measurement for the dust-corrected, aperture-corrected flux and luminosity in `flux_lya`, `flux_oii`, `lum_lya` and `lum_oii`.

2. The *Source Observation Spectra* (this is Table 3 plus accompanying Source Spectra) in FITS format:
   `hetdex_sc1_spec_vX.X.fits`.
   For each row in Table 3, we provide a corresponding 1D extracted spectra in an FITS file format consisting of seven Header Data Units (HDUs). Multiple HDUs are included as listed in Table 4. The primary HDU is empty. HDU1:INFO contains a copy of Table 3. At the same row index for each source in this table, HDU2:SPEC and HDU3:SPEC_ERR contain the aperture-corrected, 1D PSF-weighted dust-corrected spectra and their associated uncertainties in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}, computed according to the procedure outlined in Section 3.5. HDU4:SPEC_OBS and HDU5:SPEC_OBS_ERR are the aperture-corrected, observed spectrum and associated uncertainty in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}. HDU6:APCOR contains the applied aperture correction for each spectral element in the spectrum. The correction varies by wavelength due to the atmospheric diffraction correction. The final HDU7:WAVELENGTH is a 1036 array corresponding to the spectral dimension in Å. All spectra have the same spectral range from 3470–5540 Å in steps of 2 Å.
3. The *Detection Info Table* (column descriptions provided in Table 6):
   `hetdex_sc1_detinfo_vX.X.dat/.ecsv/.fits`.
   This table contains specific information for every curated line emission and continuum detection. Every emission-line detection row contains all parameter information, including the emission-line's observed wavelength, its fitted parameters, and measured flux. If the observed wavelength corresponds to a commonly found spectral species[30] at redshift `z_hetdex`, the species is indicated in the column `line_id`. There are also several columns related to imaging counterpart matches, redshift assignments, and emission-line classification as found by `ELiXer` (Davis et al. 2023). Also included are a number of columns containing details about the specific observation, the instrument, and the detection grouping parameters. A full column description is provided in Table 6. Detailed information concerning this catalog is provided in the Appendix.

[30] http://classic.sdss.org/dr6/algorithms/linestable.html

6. Sample Properties

In this section we discuss some basic properties of the HETDEX Catalog sample. More in-depth studies will be reported in future HETDEX papers. A basic discussion of the source count and magnitude distribution for each source type is provided in Section 6.1. Section 6.2 provides an overview of line fit parameters. We outline some comparisons to imaging counterparts where available in Section 6.3 and measure the imaging counterpart fraction for both the LAE and OII samples. Sections 6.4 and 6.5 cover the line luminosity and redshift distributions for the high-z and low-z samples. Using an internally confirmed sample of emission-line sources, we provide an upper limit on the LAE false-positive rate in Section 6.6.

6.1. Source Distribution

The unique count of astronomical sources is 51,863 LAEs, 123,891 [O II] emitting galaxies, 4976 AGN, 5274 LZGs, and 37,916 stars. Since the HETDEX observations contain several fields with repeated observations, a number of the sources have more than one entry. Thus the total number of source detections is greater than the number of unique sources, and consists of 40,088 stars, 129,511 [O II] emitters, 4971 AGN, 5399 LZGs, and 52,681 LAEs. In the catalog, each source is allocated a common integer `source_id` and `source_name` where the latter descriptor follows the IAU standard, i.e., of the form `HETDEX J123449.19+511733.7`. To search for repeated observations of the same object, the unique integer observation identifier `shotid` can be used. Since multiple detections can comprise a single source, we have provided the additional column `selected_det==True` in Table 6 to indicate which detection best represents the source. If a source is observed multiple times, multiple `selected_det==True` entries are in the catalog.

By field, the breakdown of sources for dex-spring: 172,831, dex-fall: 56,251, cosmos: 2447, and goods-n: 1121, which is presented relative to field size and IFU count in Table 1. In Figure 11, the $g_{\rm HETDEX}$ magnitude distribution is shown. The vertical dashed line shows the approximate limiting magnitude

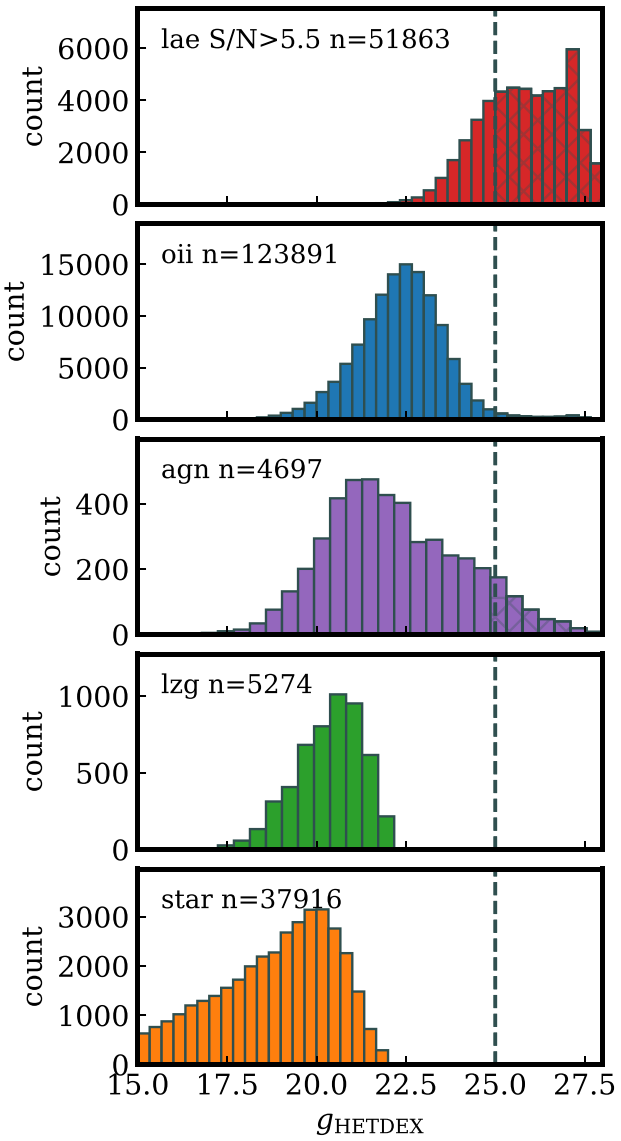

Figure 11. Histogram of $g_{\rm HETDEX}$ magnitudes for each source type as measured by summing the 1D extracted spectra, weighted by the SDSS g-band response curve. If multiple detections comprise the source, the brightest detection is used. The vertical dashed line at $g_{\rm HETDEX} = 25$ represents the HETDEX average sensitivity limit

of observations; while the precise value varies with the observing conditions, below $g_{\rm HETDEX} \approx 25$, the "magnitude" is mostly the result of summed noise. Note that the median continuum magnitude for S/N > 5.5 LAEs is $g_{\rm HETDEX}$ ∼25.9, far below this threshold. In this sample, 84.7% of the LAEs have $g_{\rm HETDEX} > 24.5$.

The LAE sample used to produce the bright-end Lyα luminosity function presented by Zhang et al. (2021) overlaps with the sample in this catalog. Their sample is derived from an earlier version of the catalog presented here. It also consists of an augmented LAE sample found by force extracting HETDEX spectra on known imaging sources in Hyper Suprime-Cam r-band (HSC-r; Miyazaki et al. 2018) data and performing an independent line emission search. As they required HSC-r band coverage for their work, their coverage consists of ∼45% (11.4 deg^2) of the coverage presented in this catalog (25.0 deg^2). Cross-matching between both catalogs shows that 91% of LAE candidates in Zhang et al. (2021) are recovered in our final source catalog and 95% in the raw line emission-line database. The primary reason LAE candidates (about 1000 LAE candidates) are not in the catalog presented here is they fail from stricter observational quality criteria cuts, updated masking, and different line parameter criteria. Visual inspection suggests about three-quarters of the missing sample are false positives due to noise and artifacts, while the rest are confident LAE candidates that are culled due to quality cuts. We do note that a substantial fraction (about 10%) of their LAE candidates are misclassified. Many of these LAEs are actually low-redshift emission-line sources (primarily due to [O II] emission, but some are due to Hβ, [O III], and other line emission) in the *HETDEX Source Catalog*.

Low-z, [O II] emitting galaxies are the most numerous in our catalog. These objects span a wide range of magnitudes ranging from $g_{\rm HETDEX} \sim 17.5$ to below the catalog's sensitivity limit (based on a threshold continuum requirement for object detection not the sensitivity limit of HETDEX observations), with a median value of $g_{\rm HETDEX} \sim 22.4$. The faint end of this distribution overlaps that of the LAEs and illustrates the need for additional observational criteria in determining whether a line emitter is from a low-z galaxy or from Lyα.

The third panel in Figure 11 shows the AGN $g_{\rm HETDEX}$ distribution. This distribution has a median magnitude that is slightly brighter than that for the [O II] galaxies, $g_{\rm HETDEX} \sim 22.0$, and contains both bright continuum sources with broad-line emission and a fainter population in which either multiple emission lines (i.e., Lyα, C IV, and He II) or a single broad emission feature is found. In some cases, broad extended line emission can have very little continuum associated with it, and it is often spatially extended, suggesting the possibility that the heating mechanism is perhaps not due to an AGN.

The LZG population is mostly detected in the continuum detection search, although a detection of line emission can also result if there is a continuum peak within two absorption troughs or if there is an abrupt change in an object's continuum level. Not surprisingly, LZG galaxies are generally brighter than [O II] emitters, with a median $g_{\rm HETDEX}$ of 20.4 mag, and they are much less numerous. But there is considerable overlap in the populations, and both object classes are present in our continuum catalog. Initially, many [O II] galaxies were missed by our line-emission search, and were only discovered by measuring spatially resolved [O II] aperture fluxes (see Section 3.4). If we consider just continuum detected sources, which can loosely be considered analogous to a magnitude-limited survey, then 16.3% of the low-z galaxies have no measured line emission, while 82.9% have measured continuum-subtracted [O II] emission. Alternatively, if we consider all HETDEX low-z detections (both emission-line and continuum objects) then just 2.7% of the catalog consists of LZGs, and 97.2% are [O II] systems.

Both the HETDEX line and continuum emission catalogs contain entries from stellar emission. As with LZGs, discontinuity jumps in the continuum can be mistaken for emission features; our spatial clustering ensures that duel detections are properly merged into a single source. The stars in the HETDEX catalog can be as bright as $g_{\rm HETDEX} \sim 12.3$ with a median value of $g_{\rm HETDEX} \approx 19.0$. The number of stars in this catalog is considerably smaller than that reported in the HETDEX-Gaia catalog (Hawkins et al. 2021), due to the very different selection criteria and methods of detection applied. In

Hawkins et al. (2021), spectral extractions were performed at the known locations of $10 < G < 22$ Gaia DR2 stars (Gaia Collaboration et al. 2018) using the same data release (HDR2) presented here. Their catalog consisted of 98,736 unique stellar candidates; our catalog contains 37,916 of these objects.

The major difference between the two catalogs lies in the heterogeneous sensitivity of the HETDEX continuum detection level. At magnitudes of $g_{\rm HETDEX} < 15$, bright stars and galaxies saturate the detectors, making it impossible to properly perform a flux calibration. When this happens, the detector may be useless for HETDEX science, but may still produce detectable and classifiable stellar spectra. Moreover, as discussed in Section 2.1, frames with bright stars or additional detector issues will fail our observation quality criteria and are removed from the survey; the forced extraction methods performed by Hawkins et al. (2021) did not apply these additional criteria, and so will include more objects. In addition, our continuum detection search was essentially count limited, rather than magnitude limited. Depending on observing conditions, this restricted our continuum-selected sample to objects brighter than $g \sim 21$–22.5. Finally, we note that it appears the HETDEX-Gaia catalog has some likely galaxy contaminants: out of the 37,916 overlapping sources, 87.3% were classified in the HETDEX Source Catalog as stars, 3.9% as AGNs, 6.6% as [O II] galaxies, and 2.0% as LZGs. Indeed when measuring radial velocities, Hawkins et al. (2021) report low-level (∼2%) contamination by galaxies.

6.2. Line Parameter Properties

As described in Section 3.1.3, the Gaussian amplitude (which provides the line flux), line width (σ), and continuum level are the free parameters of the emission-line fit. The central location of the line detection is determined by rastering the line fit at high spatial resolution to maximize on S/N of the fit. Table 6 reports theses values and the quality of fit, χ^2. We employ variable cuts on the line parameters when we create the curated line emission catalog (as discussed in Section 3.1.3 and described in Equations (1) and (2)). The internal HETDEX Source Catalog is limited to line emission detections with S/N > 4.8, while the public version excludes all LAEs with S/N < 5.5. At the lower value, confidence in the line detection goes down, and the visual line fits are of poor quality.

In Figure 12, we plot the S/N distribution for the LAEs and [O II] galaxies in the top and bottom panels, respectively. The lighter colors are for the full sample included in the catalog; the solid region is for an isolated "confirmed" sample with at least three independent observations. Many of these objects are from the science verification fields, which were subjected to multiple visits (see Section 6.3.2 for a complete discussion). The sample of sources confirmed by multiple observations has a similar S/N distribution as the objects in the full catalog, although the number counts go significantly down at higher S/N as sample variance plagues the "confirmed" sample. For both samples, as expected, the number of sources increases as S/N decreases. Unsurprisingly, the [O II] galaxies extend to much higher S/N than the LAEs.

The left panel of Figure 13 shows the distribution of Lyα and [O II] line fluxes for the two main line emission samples. Lyα line flux values range from 3.4–1030×10^{-17} erg s^{-1} cm^{-2} Å^{-1} where the 10th, 50th, and 90th percentiles are 8.3, 14.5, and 32.8×10^{-17} erg s^{-1} cm^{-2} Å^{-1}, respectively. The [O II] line fluxes range from 3.7–2960×10^{-17} erg s^{-1} cm^{-2} Å^{-1} where the 10th, 50th, and 90th percentiles are 10.2, 20.0, and 53.2×10^{-17} erg s^{-1} cm^{-2} Å^{-1}, respectively. Distributions for the AGN sample can be found in Liu et al. (2022). Some line emission sources were observed multiple times in this release, and have multiple measurements in the catalog. These can be isolated by searching for a common `source_id` value but a different observation ID (`shotid`) value. The rms scatter in repeated line measurements is 13.1% if comparison sources are required to be within 0.″5. This value increases to 13.8% without any sky separation requirement, suggesting source position plays a role in line flux accuracy, as discussed earlier in Section 3.4.1. Simulations described in Gebhardt et al. (2021) suggest an even higher statistical uncertainty of 25%–30% that is signal-to-noise dependent. A better flux accuracy is measured in the presented catalog likely due to a stricter S/N requirement of 5.5.

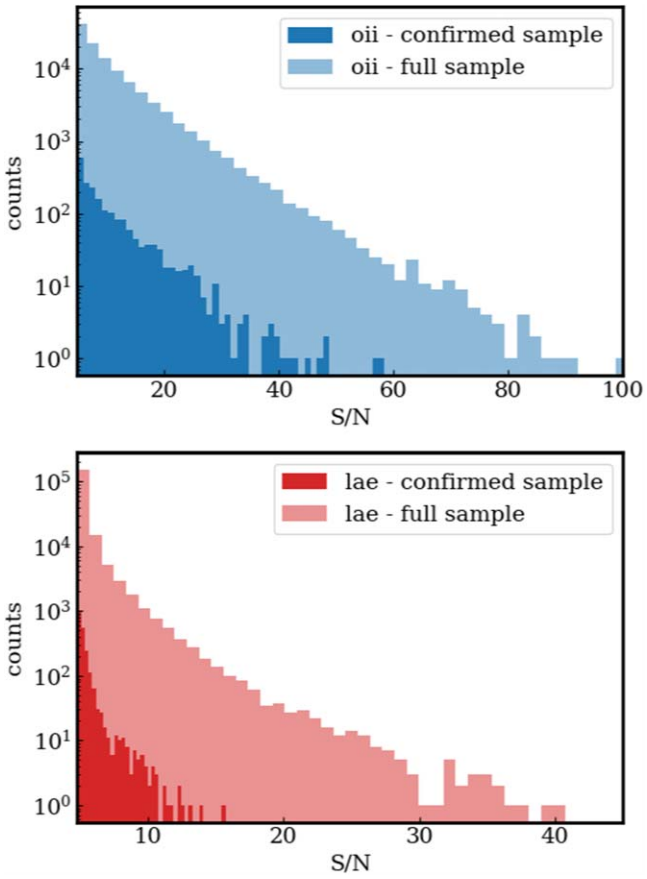

Figure 12. Signal-to-noise ratio (S/N) distributions. The dark colors show the confirmed sample of emission-line sources detected multiple times in independent HETDEX observations; the lighter colors display the full catalog distribution. LAEs are generally fainter than [O II] galaxies and have lower values of S/N. Note the difference in the *x*-axis ranges in the panels.

A comparison to two external samples of published line flux values is shown in Figure 14. Here we compare a strictly LAE sample from the SC4K survey (Sobral et al. 2018) shown in blue, as well as a mix of [O II] and Lyα-emitting galaxies from the HETDEX Pilot Survey (HPS; Adams et al. 2011). SC4K uses 16 different narrowband and medium-band filters over the COSMOS field to select a large sample of LAEs. For these emitters, we require a match to within 1″ spatially and within 300 Å of the HETDEX emission-line wavelength. We find that 50 LAEs and 17 AGNs overlap with our catalog. We find 30% of the HETDEX line fluxes are more than three times the

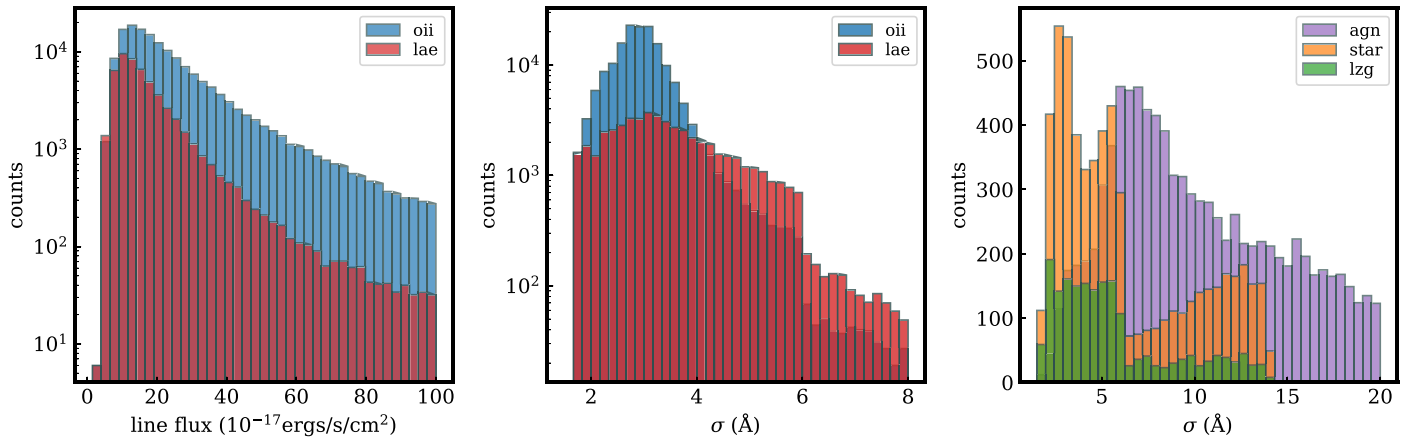

Figure 13. The left panel shows the observed line fluxes for the for LAEs (in red) and [O II] galaxies (in blue). The middle panel shows the fitted Gaussian line width distribution in Å for the LAEs (in red) and [O II] galaxies (in blue). The right panel gives the same information for AGNs, LZGs, and stars. These classes are less numerous and exhibit a broader line distribution than that for emission-line galaxies.

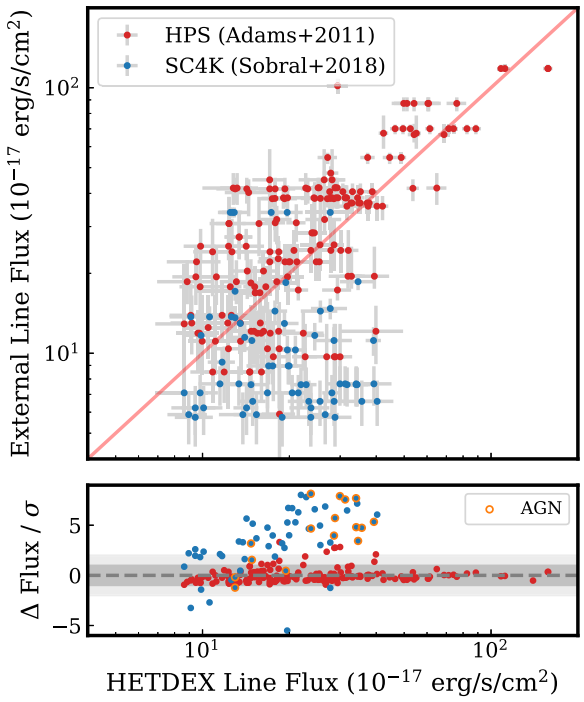

Figure 14. The top panel is a comparison of point-source line fluxes measured from the HETDEX line detection pipeline to two external surveys, SC4K (Sobral et al. 2018) in blue, and HPS (Adams et al. 2011) in red. The comparison comprises LAEs, AGNs, and O IIs, although the SC4K sample is limited to just LAEs and AGNs at high-z. The bottom panel compares the line flux difference (HETDEX minus the external values) in the samples relative to the combined uncertainty. The dark-gray shaded region indicates 1σ agreement, and the light-gray shaded region indicates 2σ agreement. For the HPS sample, they agree within 1σ 92% of the time. HETDEX line fluxes are overestimated relative to the SC4K sample. AGNs from the HETDEX-SC4K overlap sample are indicated in orange and represent one-third of the comparison sample whose fluxes disagree.

reported SC4K fluxes with 50% of these being AGNs in our catalog. SC4K measures their fluxes in 2″ apertures, and the continuum is measured in multiple narrowband filters potentially leading to inconsistencies in the measurement, which is most significant for the AGN sources. Designed as an HETDEX validation survey, HPS is similar in design to HETDEX. HPS used the VIRUS prototype IFU (VIRUS-P; Hill et al. 2008) on the 2.7 m Harlan J. Smith telescope at the McDonald. Three-dithered exposures provide similar coverage at similar resolving power, with the main difference being VIRUS-P's larger fiber size (4.″2 compared to VIRUS's 1.″5 diameter fibers). Data is reduced in a completely independent pipeline from the current HETDEX pipeline, and line fluxes are measured differently. In total, 179 detections overlap within 8 Å spectrally and 1″ spatially; 128 of these are O IIs, 37 are LAEs, and 12 are AGNs. The bottom panel shows that the line flux differences (HETDEX—HPS) are consistent to within 1σ of the combined lines flux uncertainties 92% of the time.

The fitted Gaussian line width distributions for the [O II] and LAE samples have roughly the same range. As shown in the middle panel of Figure 13, Lyα emitters show a broader distribution and have a higher frequency of broad-line emitters relative to the [O II] sample, as might be expected (Kulas et al. 2012; Chonis et al. 2013). The figure also shows a substantial decrease in the distribution at $\sigma = 6$: our fitting criteria removes all lower signal-to-noise ($S/N < 6.5$) lines with $\sigma > 6$ (see Equation (1)). The drop off suggests we are losing real sources with this criteria, especially in the LAE sample at a 0.001% level. Fortunately, the HETDEX AGN Catalog selection does explore this parameter space. In the higher line width regime, visual inspection confirms the existence of many artifacts due to issues involving both the detector and the calibration.

6.3. Imaging Counterpart

We run each line and continuum detection through our source classifying software `ELiXer` (Davis et al. 2023) to obtain photometric magnitudes at the direct location of the HETDEX detection. We use the multiwavelength coverage provided by the Hyper Suprime-Cam through the Subaru Strategic Program (HSC-SSP; Aihara et al. 2018, 2022) as well as internally obtained HSC-r-band imaging as described in Davis et al. (2023). Data reduction and source detections were performed with version 6.7 of the HSC pipeline, hscPipe (Bosch et al. 2018), and produced r-band images with a 10σ limit of $r = 25.1$ mag in a 2″ diameter circular aperture.

The magnitude distribution of imaging counterpart to HETDEX sources, as measured in the HSC r band, are shown in Figure 15 for the LAE and OII catalog sources. We show the full sample in shaded opacity while highlighting the confirmed samples in solid colors to show the overall consistencies in the two populations. [O II] galaxies are

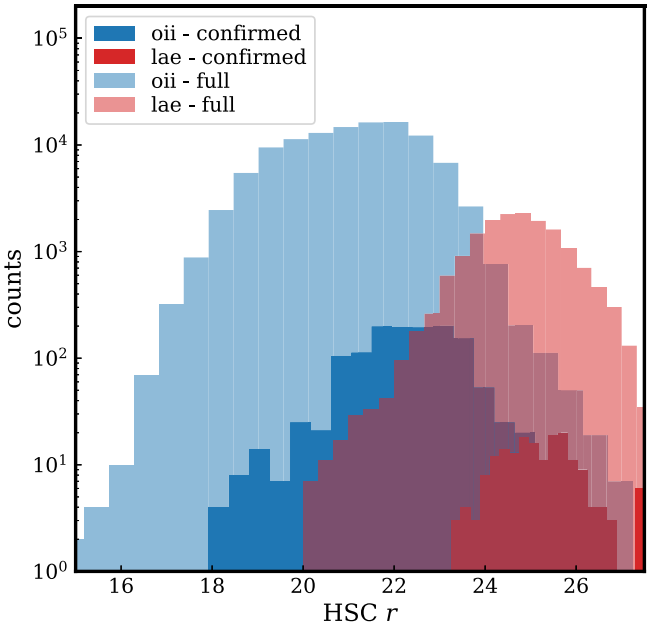

Figure 15. Distribution of HSC *r*-band image counterparts for the main catalog LAE and O II samples. Also shown are a "confirmed" sample of sources that are high-confidence HETDEX detections, which are detected on at least three independent nights. These values come from the ELiXer catalog SEP measurements performed on the image at the location of the HETDEX sources as described in Davis et al. (2023). The comparison between the samples is to show that the confirmed sample is a good representation of the full line emission.

brighter than the LAEs with a median value of $r = 21.6$. In contrast, the LAE sample has a median magnitude of $r = 24.7$, but we note that this distribution is biased due to the sensitivity limit of the *r*-band imaging data.

6.3.1. g-band Magnitude Comparison

In Section 6.1, we described the $g_{\rm HETDEX}$ magnitude distributions for the catalog. Continuum sensitivity from a single extracted HETDEX spectrum varies based on observing conditions, but is generally reliable to $g_{\rm HETDEX} \sim 25$, although the uncertainties can be large. We compare these magnitudes with *g*-band measurements calculated with our ELiXer software on *g*-band data from the HSC-SSP program, which reach a sensitivity of $g = 26.5$ mag.

In Figure 16, the difference in magnitude is plotted as a function of $g_{\rm HETDEX}$. The stellar sample shows an offset of 0.05 mag with a scatter of 0.51 mag. The [O II] sample has an offset of 0.14 mag with a scatter of 0.41 mag. The LZG has the largest offset of −0.34 mag with a scatter of 0.57 mag. The AGN sample has an offset of 0.13 mag with a scatter of 0.54 mag. For both the [O II] and LZG samples, the aperture effects are important, as HETDEX magnitudes are for point-source measurements and the HSC-SSP measurements are for varying apertures. We intentionally exclude any sources with a major axis greater than 3″ to mitigate aperture affects. Unsurprisingly, the faint LAE sample that pushes the sensitivity limits of HETDEX shows the largest scatter of 0.59 mag, but a modest offset of 0.17 mag. An overall trend line shows that the offset is largest at faint magnitudes where a lower S/N in the HETDEX spectra leads to a lower integrated flux.

6.3.2. Imaging Counterpart Fraction

In HSC *r*-band, our preferred photometric data set, many LAEs in our sample have no apparent counterpart. This is not surprising given that the LAE sample exhibits stellar masses ranging from $\sim 10^8$–$10^{10}\,M_\odot$ (Hagen et al. 2016; Oyarzún et al. 2017; McCarron et al. 2022). However, we also recognize that a lack of imaging counterparts can be an indication of potential false-positive contamination. This is specific to the high-*z* sample since the low-*z* identifications require that a continuum be detected in either the spectral data or the accompanying imaging. Thus, by definition, the [O II] sample has less contamination from noise and artifacts.

One way to investigate the amount of contamination present in our sample is to compare the fraction of confirmed LAEs with imaging counterparts to that of the full catalog sample. There are a number of ways to confirm the presence of line emission for an HETDEX detection. For example, sources confirmed by other instruments can confirm an object as real (as well as provide confirmation of the redshift and classification). Alternatively, we can use the HETDEX data themselves: if line emission is detected in three or more independent observations, we confirm the source to be real. Although HETDEX tiling is designed to visit the sky just once, there are a number of science verification fields (in legacy regions such as COSMOS and GOODS-N) that HETDEX has visited on numerous occasions. In a few cases, observations were redone, due to substandard observing conditions. Although the unacceptable observations are not in our final catalog, the emission-line detections remain in the raw line database and are used for object verification. Finally, in a few cases, the corners of IFUs overlap (see, for example, the overlap in the right panel of Figure 3), due to tiling changes associated with the increasing number of active IFUs over time. We call the set of objects identified in three or more independent observations our "confirmed" sample. We show the S/N and HSC-*r*-band imaging counterpart magnitudes for the confirmed sample relative to the full catalog sample in Figures 12 and 15. The similar distributions indicate the sample is representative of the full catalog.

We consider two different catalog samples, all of which require accompanying HSC-*r* imaging coverage, for this test: the publicly provided catalog LAE sample ($n = 39{,}083$) and OII sample ($n = 86{,}357$). We compare these to their counterparts in the confirmed data set: $n = 422$ for the LAE sample, $n = 1529$ for the OII sample. To be in these subsamples, we require that each source be contained on an HSC *r*-band image, which we assume to have a depth better than $r = 26.2$ mag (5σ limit). For each sample, we calculate the number of HETDEX detections with an *r*-band counterpart as measured in either one of two ways: (1) through on-demand source extraction applied to the HSC image using the ELiXer software tool, or (2) a forced extraction at the exact position of the HETDEX detection and measured within an $r = 1\farcs5$ circular aperture.

In Table 5, we summarize the net fraction of LAEs and OII galaxies with HSC *r*-band counterparts. The first thing to note is that the fractions for the confirmed sample and the main sample are consistent to within a few percentage points. This implies that the catalog is relatively free of contamination by false positives. In a catalog with many false detections, the fraction of LAEs with counterparts would be much greater in the confirmed sample than in the main sample.

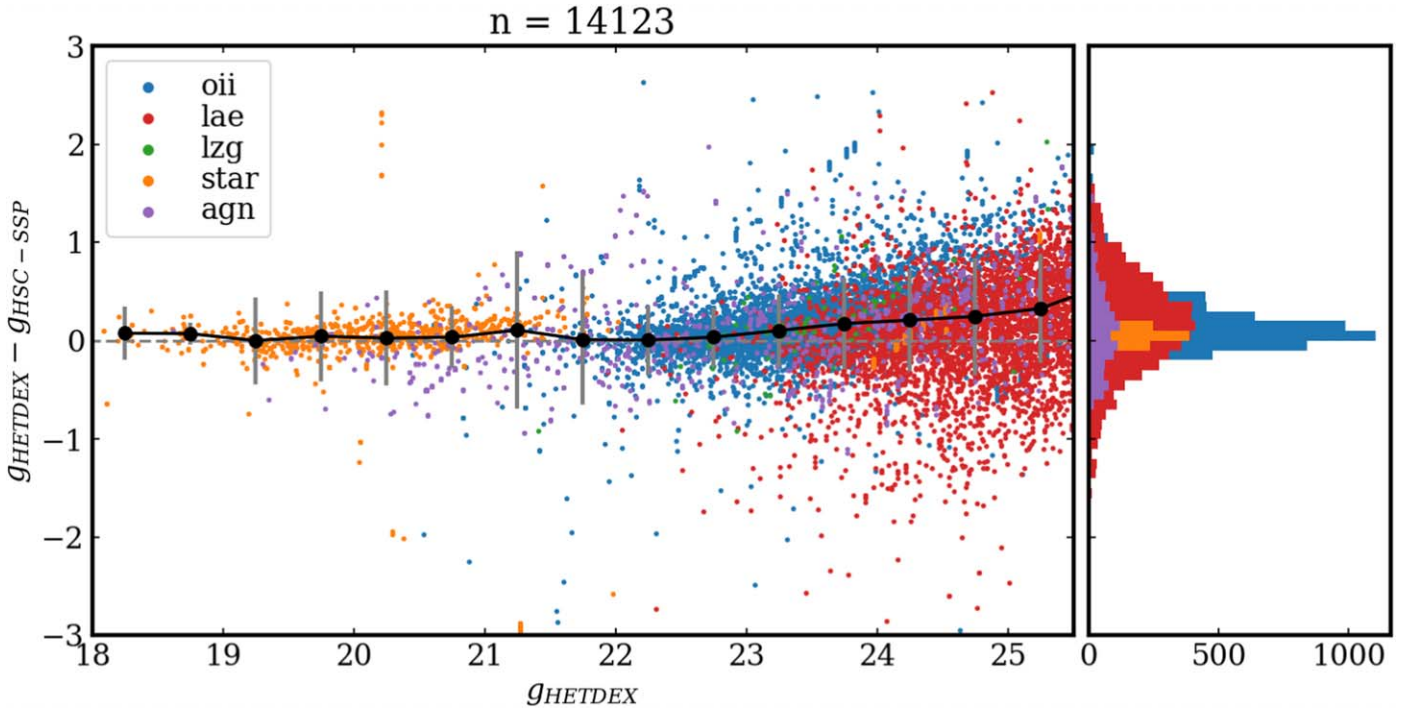

Figure 16. We compare the g-band magnitude to all imaging counterparts in deep g-band data from HSC-SSP. Depending on conditions, HETDEX can be sensitive to $g \sim 25$ mag.

Table 5
HSC r-band Imaging Counterpart Fraction

	Full Catalog	Confirmed
LAE	55.8% (21,824/39,083)	52.4% (221/422)
O II	99.3% (85,754/86,357)	99.2% (1517/1529)

Note. The "Confirmed" sample of line emitters consists of objects that have been independently detected in three different observations. The detected objects have r-band counterparts brighter than $r = 26.2$.

In Figure 17, we consider how the fraction of objects with r-band counterparts depends on the S/N of the line (top) and $g_{\rm HETDEX}$ (bottom). The counterpart fraction increases with the S/N on the LAE emission line. This implies the brighter emitters tend to also have higher line flux in our catalog as seen in previous LAE studies where equivalent width has a minimal dependence on line luminosity (e.g., Gronwall et al. 2007; Ciardullo et al. 2012). By design, nearly all OII sources are present in the r-band images, as a continuum detection (either through imaging or spectroscopy) is needed for a detection to be classified as OII. There are some OII detections that have faint $g_{\rm HETDEX}$ values and are consequently less likely to be observable in the r. These are sources that have been classified as OII by `ELiXer` even though they have a very weak continuum. Upon inspection, these often tend to be false positives, due either to calibration issues or satellites.

The measured imaging counterpart fraction demonstrates the strength of an IFU-based LAE survey. Just over half of the LAE sample have image counterparts brighter than $r = 26.2$. A survey based on imaging preselection at this sensitivity would miss half of the objects that HETDEX is using to trace the large-scale structure of the $1.9 < z < 3.5$ universe.

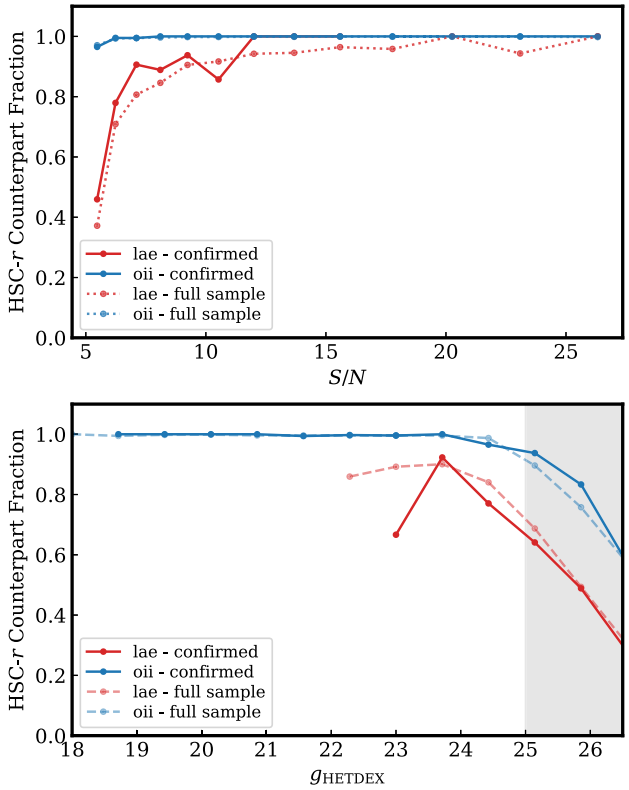

Figure 17. The fraction of O II and LAE sources with $r < 26.2$-band HSC counterparts, as a function of S/N (top panel) and $g_{\rm HETDEX}$ (bottom panel). The solid line is for a subset of sources that are confirmed by three independent HETDEX observations of the source.

6.3.3. Source Positioning

Table 6 contains selected output from the `ELiXer` catalog about the imaging counterparts of HETDEX detections.

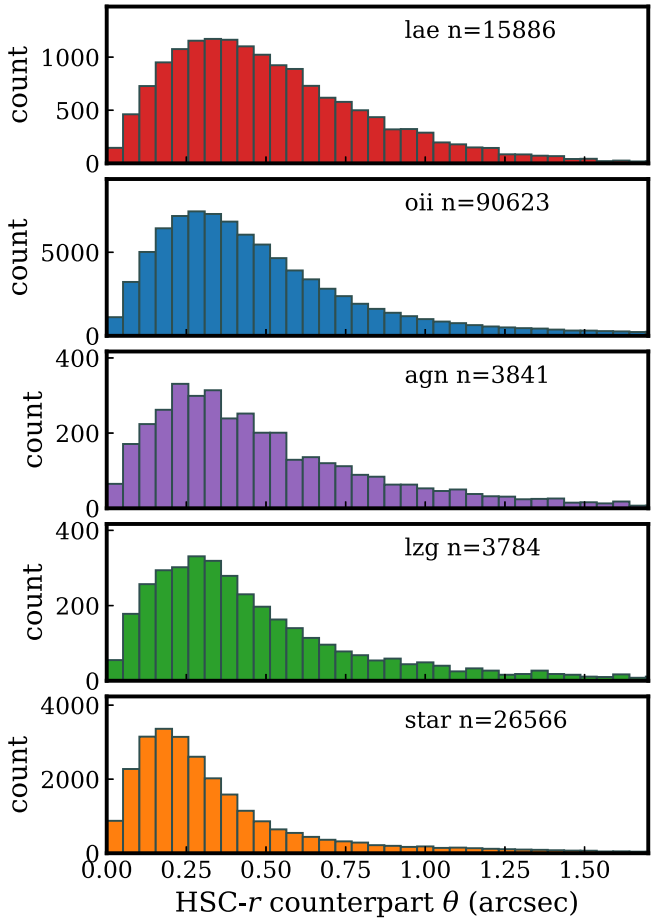

Figure 18. Distributions of sky separations between HETDEX source detections and their imaging counterparts. The LAE sample shows the widest spatial distribution is primarily due to poorer centroid positioning of low-S/N HETDEX detections. The star sample in orange at the bottom shows the tightest distribution with a median sky separation of $0\farcs27$.

Included in these data are the separations between the HETDEX sources and the position of the most likely imaging counterparts (labeled as `counterpart_dist`). Figure 18 shows distributions of these separations for each source type. The LAE sample shows the widest distribution. This is primarily attributed to a poorer ability to center low-S/N emission lines within the PSF-weighted VIRUS spectral extraction. McCarron et al. (2022) found in deep GOODS-N HST imaging that HETDEX imaging counterparts can be up to $1''$ in separation from the HETDEX detection center.

The star sample in orange at the bottom shows the tightest distribution with a median sky separation of $0\farcs27$. This is comparable to our astrometric uncertainties ($\sim 0\farcs2$). OII galaxies, AGNs, and LZGs, have a wider distribution as there can be differences in where the source center lies.

6.4. Luminosities

In Figure 19, we show the emission-line luminosities for LAEs and the OII galaxies. The Lyα line luminosities range from 1.84×10^{42} erg s^{-1} to 2.85×10^{44} erg s^{-1} with a median value of 8.31×10^{42} erg s^{-1}. For each LAE source, the corresponding Lyα flux and luminosity are found in the columns `flux_lya` and `lum_lya`. For the OII galaxies, 87.9% of the values (as indicated by `flag_aper==1`) are

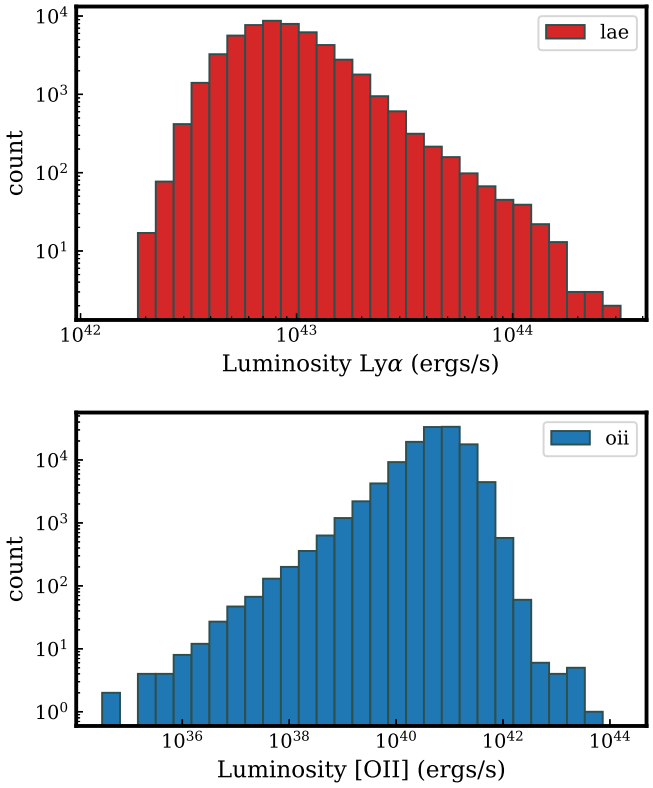

Figure 19. Top: Lyα luminosity distribution for 51,863 LAEs (solid red). The lower-S/N data set has a larger fraction of lower-luminosity sources, but both samples span a similar luminosity range. Bottom: [O II] luminosities for the low-z line emission galaxies. Here 88% of the values come from resolved aperture line fluxes (`flux_aper`). The rest are assumed to be point-source-like and are the HETDEX pipeline fluxes (`flux`).

from resolved aperture line fluxes (`flux_aper`); the rest are assumed to be pointlike sources and come from HETDEX pipeline (`flux`). The selected OII flux values can be found in the *Source Observation Table* (see Table 3) in the column `flux_oii`, and the luminosities are given in column `lum_oii`. The [O II] line luminosities of our sample range from 6.13×10^{32} erg s^{-1} to 1.76×10^{44} erg s^{-1} with a median value of 1.96×10^{40} erg s^{-1}.

6.5. Redshift Distribution

In Figure 20, we show the redshift distribution of the low-z and high-z galaxy samples. For the low-z sample, galaxies with (`source_type==oii`) and without (`source_type==lzg`) [O II] line emission are shown together; their counts increase with z, as the greater volume at higher z is more important than the accompanying decrease in survey depth. The dip at `z_hetdex` ~ 0.21 in the low-z sample (and at `z_hetdex` ~ 2.7 in the high-z data set) is due to a mask that is applied at the center of 50% of the detectors as well as an increase in night sky emission. Night sky emission, particularly in the blue, causes marked decreases in number counts in the lower-redshift regions of both data sets but is most notable at high-z. The brightest sky lines are marked by light yellow bars in Figure 20. At these epochs, the loss in depth due to increased distance outweighs the volume effect. Sample variance can play a small role in the variation in counts, but given the size of the survey, this should largely be mitigated. The remaining

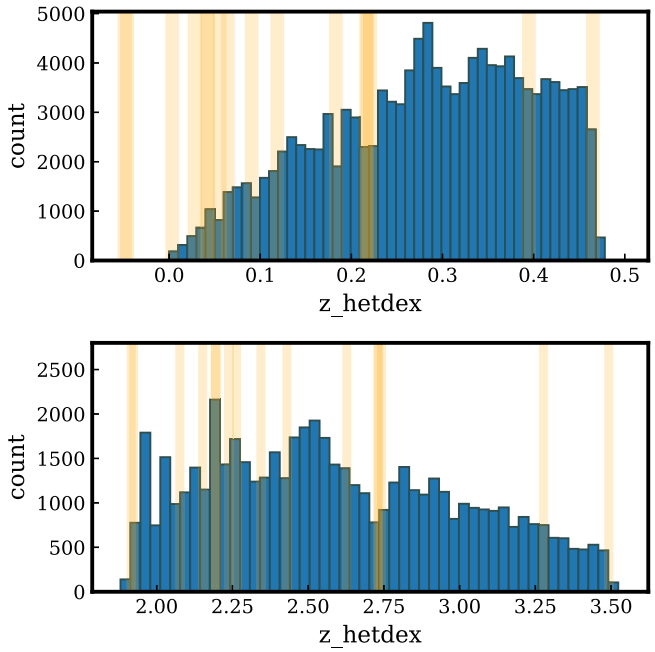

Figure 20. The redshift distribution of the low-z and high-z galaxy samples is shown in the top and bottom panels, respectively. The low-z sample is a combination of O II emitters and LZGs; the high-z data set is limited to LAEs with S/N > 5.5. The brightest sky lines are marked by light yellow bars, which cause a suppression in number counts in the high-z distribution.

variability in counts as a function of wavelength is due to the complex sensitivity variations caused by variable observing conditions, detector spectral response, and fiber-to-fiber (and amplifier-to-amplifier) variations. Details concerning HETDEX's complex selection will be described in D. Farrow et al. (2022, in preparation).

6.6. Overall Sample Validation

HETDEX is designed to search for faint, low-S/N emission lines in a large amount of data. HDR2 consists of 208 million fiber spectra, each with 1036 spectral resolution elements. This means examining over 210 billion resolution elements. Noise is ultimately the biggest contaminator in our catalog, and the vast majority of our spectra observe the blank sky. Attempts to quantify the HETDEX false-positive rate are ongoing but here we briefly summarize our efforts to confirm sources.

Multiple methods are used to measure the confirmation rate of HETDEX line emission sources. The method that provides the highest number of confirmed sources involves using the HETDEX data themselves. As described in Section 6.3.2, we assume that any emission-line source found in three independent observations is real. We create a validation sample by considering every OII and LAE in the catalog and checking to see if the location of the source has been targeted multiple times. This is done by cross-matching the catalog with the fiber database. If a source has fiber coverage from at least six observations, we put it in the validation sample. We note that just because a sky position has fiber coverage, that does not mean the observation is useful, as varying observing conditions may prevent a real source from being observed. Ultimately the confirmation rate provides only an upper limit on our false-positive rate. Consequently, we also use spectroscopic redshifts from the literature to validate HETDEX detections if the redshift matches the redshift in the literature to within $\Delta z < 0.02$

For the S/N > 5.5 validation sample, 91.0% of LAEs are confirmed; for the OII sample, the fraction is 99.3%. The combined fraction is 98.1% because of the high fraction of [O II]-emitting galaxies in the catalog (123,891) compared to the number of S/N > 5.5 LAEs (51,863).

7. Summary

HETDEX is a medium-wide area, IFU spectroscopic survey that covers the wavelength range 3500 Å $\lesssim \lambda \lesssim$ 5500 Å at a resolving power of $750 < R < 950$. The survey will ultimately cover $\sim$540 deg^2 with noncontiguous tiling leading to 94 deg^2 of complete sky coverage. With the clustering of over 1 million high-z Lyα-emitting galaxies, HETDEX aims to measure the Hubble expansion rate, $H(z)$, and the angular diameter distance, $D_A(z)$, to better than 1% accuracy.

This paper describes the first publicly released version of the HETDEX Source Catalog. The catalog is generated by combining raw HETDEX line emission and continuum emission detections, which are performed in a grid search under point-source assumptions. While there is some overlap between the two samples, the line emission search offers the unique capability of detecting very distant galaxies with relatively modest continuum emission and stellar mass through their bright Lyα-emission. The catalog contains 51,863 LAEs, 123,891 [O II]-emitting galaxies, 37,916 stars, 5274 low-z non-line-emitting galaxies, and 4976 AGNs. By utilizing a three-prong classification approach, we provide robust spectroscopic redshifts and classifications for the entire catalog. When compared to external catalog spectroscopic redshifts, 96.1% of the sources are within $\Delta z < 0.02$.

Using a sample of repeat source observations, we create a "confirmed" sample of confident sources. This allows us to validate line emitters that have not been observed in any other data set. In this "confirmed" sample, we find that we can confirm 91.0% of the LAE sample and 99.3% of the OII sample through evidence in repeat observations, suggesting an upper limit of 9% for the false-positive rate in the LAE sample.

Without any imaging preselection, HETDEX offers a blind search for LAEs. A search for imaging counterparts to the LAE sample in deep ancillary HSC r-band imaging shows that 45% of the LAE sample has no detected imaging counterpart down to a limiting magnitude of $r = 26.2$. A sample with imaging preselection at this sensitivity would miss half the HETDEX LAE sample presented in this paper.

Data access and details about the catalog can be found at http://hetdex.org. A copy of the HETDEX Public Source Catalog (Version 3.2) is available on Zenodo doi:10.5281/zenodo.7448504. This Zenodo deposit includes the *Source Observation Table* (columns described in Table 3), a FITS file with source spectra corresponding to each source observation, and the *Detection Info Table* (columns described in Table 6) in multiple formats as well as a Jupyter notebook with access examples.

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, The Pennsylvania State University, Institut für Astrophysik Göttingen, The University of Oxford, Max-Planck-Institut für Astrophysik (MPA), The University of Tokyo, and Missouri University of Science and Technology. In addition to Institutional support, HETDEX is funded by the National Science Foundation (grant AST-0926815), the State of Texas, the US Air Force (AFRL FA9451-04-2-0355), and generous support from private individuals and foundations.

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

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 authors are thankful to the Dark Energy Spectroscopic Instrument Survey team for providing invaluable early Survey Validation observations of a subset of the HETDEX emission-line sample.

The Institute for Gravitation and the Cosmos is supported by the Eberly College of Science and the Office of the Senior Vice President for Research at the Pennsylvania State University. The Kavli IPMU is supported by World Premier International Research Center Initiative (WPI), MEXT, Japan.

This work makes use of the Sloan Digital Sky Survey IV, with funding provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

This work makes use of the Pan-STARRS1 Surveys (PS1) and the PS1 public science archive, which have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society, and its participating institutes.

This work makes use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

This work makes use of observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555.

K.G. acknowledges support from NSF-2008793.

Facility: The Hobby-Eberly Telescope (McDonald Observatory).

Software: This research was made possible by the open-source projects astropy (Astropy Collaboration et al. 2018), python (Van Rossum & Drake 2009), numpy (Harris et al. 2020), Scipy (Virtanen et al. 2020), hetdex-api (https://github.com/HETDEX/hetdex_api), elixer (https://github.com/HETDEX/elixer; Davis et al. 2021), diagnose (https://github.com/grzeimann/Diagnose; Zeimann et al. 2023, in preparation), photutils (Bradley et al. 2021), dustmaps (Green 2018), extinction (https://github.com/kbarbary/extinction).

Appendix
Detection Info Table

This Appendix describes the *Detection Info Table* with column descriptions provided in Table 6, which contains information for every line and continuum detection from our object detection search method (as described in Section 3.1). As described in Section 3.3, an HETDEX source can be composed of a collection of line emission and continuum emission detections. *The Source Observation Table* (columns described in Table 3) provides a simplified version of Table 6 with one row per source observation, providing basic information about a source such as coordinates, redshift, $g_{\rm HETDEX}$ magnitude, and the [O II] and Lyα line flux and luminosity where applicable. Table 6 presented in this Appendix is expanded to provide additional information on every detection in a source. While many columns are the same as those in Table 3, such as `source_id`, `source_name`, `RA`, `DEC`, `z_hetdex`, additional information is provided regarding line fit parameter information. This includes the specific position of the detection (`RA_det`, `Dec_det`) and wavelength (`wave`) for the detection, the detection's line width, (σ: `sigma`), continuum-subtracted line flux, and the local continuum measurement. Each observed wavelength is checked to see if it is a rest-frame match to a common line species at `z_hetdex`. Specifically, we consider C III, C IV, Hβ, Hδ, Hγ, He II, Lyα, [O II], and [O III].[31] If a match is found, it is listed in `line_id`. Not all detections have a `line_id`, as some HETDEX line emission detections can result from jumps in a spectrum or calibration issues. We attempt to mitigate these by excluding high line width sources that are not selected as the main detection (i.e., `selected_det` = =True) of a source. Other information as described in the text is also provided. This includes detection group information from 3D and 2D FOF detection grouping and `ELiXer` imaging counterpart information. Also included are specific observation parameters such as the image quality of the observation, `fwhm`, and its observation ID information (e.g., `shotid`, `date`, `obsid`, and `field`) and specific information related to the highest weight fiber in the spectral extraction of the detection (such as `multiframe`, `fiber_id`, and `weight`). The detection whose spectrum is included in Table 3 that is the best representative of a source (typically the brightest magnitude detection) is identified by `selected_det==True`. The description of all of the parameters is provided in Table 6.

[31] http://classic.sdss.org/dr6/algorithms/linestable.html

Table 6
Detection Info Table Column Descriptions

Name	Description
source_id	HETDEX Source Identifier
source_name	HETDEX IAU designation
RA	source_id R.A. (ICRS deg)
DEC	source_id decl. (ICRS deg)
z_hetdex	HETDEX spectroscopic redshift
z_hetdex_src	HETDEX spectroscopic redshift source
z_hetdex_conf	0 to 1 confidence HETDEX spectroscopic redshift source
source_type	options are `star`, `lae`, `agn`, `lzg`, `oii`, and `none`
detectid	emission line or detection ID
selected_det	best detect ID for Lyα flux or [O II] line flux
det_type	detection type: "line" or "continuum"
line_id	line identification at observed wavelength (wave) assuming redshift of z_hetdex
RA_det	detect ID R.A. (ICRS deg)
DEC_det	detect ID decl. (ICRS deg)
src_separation	separation in degrees between the detect ID (RA_det, DEC_det) and the source_id center (R. A., decl.)
n_members	number of detections in the source group
gmag_err	MCMC uncertainty in gmag
gmag	SDSS g-magnitude measured in HETDEX spectrum
Av	applied dust correction in the V band
ebv	applied selective extinction
wave	central wavelength of line emission (Å)
wave_err	MCMC error in wave (Å)
flux	dust-corrected line flux 10^{-17} erg s^{-1} cm^{-2}
flux_err	MCMC error in dust-corrected line flux
flux_obs	observed line flux 10^{-17} erg s^{-1} cm^{-2}
flux_obs_err	MCMC error in observed line flux
flux_aper	dust-corrected, O II line flux measured in elliptical galaxy aperture in 10^{-17} erg s^{-1} cm^{-2}
flux_aper_err	error in flux_aper
flux_aper_obs	O II line flux measured in elliptical galaxy aperture in 10^{-17} erg s^{-1} cm^{-2}
flux_aper_obs_err	error in flag_aper_obs
flag_aper	1 = aperture line flux used for lum_oii, −1 = PSF-line flux used from "flux" column
sigma	sigma line width in Gaussian line fit (Å)
sigma_err	MCMC error in sigma line width (Å)
continuum	local fitted observed continuum in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}
continuum_err	MCMC error in continuum in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}
continuum_obs	local fitted observed continuum in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}
continuum_obs_err	MCMC error in continuum in 10^{-17} erg s^{-1} cm^{-2} Å^{-1}
sn	signal-to-noise for line emission
sn_err	MCMC error in signal-to-noise
chi2	reduced χ^2 quality of line fit
chi2_err	MCMC uncertainty in reduced χ^2
flux_noise_1sigma_obs	observed 1σ flux sensitivity in 10^{-17} erg s^{-1} cm^{-2}
flux_noise_1sigma	dust-corrected 1σ flux sensitivity in 10^{-17} erg s^{-1} cm^{-2}
apcor	aperture correction applied to spectrum at 4500 Å
counterpart_mag	selected closest counterpart mag from source extracting on image data
counterpart_mag_err	uncertainty in counterpart_mag
counterpart_dist	distance to closest counterpart
counterpart_catalog	image catalog source of counterpart
counterpart_filter	image filter of counterpart
plya_classification	`ELiXer` likelihood line is Lyα ranges 0 to 1 (1 = high-probability line is Lyα)
best_z	`ELiXer` best redshift
best_pz	confidence in best_z
z_diagnose	best-fit redshift from `Diagnose`
cls_diagnose	best classification from `Diagnose`. Options are "STAR," "GALAXY," "QSO," and "UNKNOWN"
stellartype	`Diagnose` spectral type classification for stars
agn_flag	-1 not an AGN, 0 broad-line source but not confirmed AGN, 1 confident AGN and z_hetdex
wave_group_id	ID for 3D FOF clustering at common R.A., decl., wave
wave_group_a	semimajor axis from 3D FOF clustering
wave_group_b	semiminor axis from 3D FOF clustering
wave_group_pa	positional angle from 3D FOF clustering
wave_group_ra	mean ra from 3D FOF clustering
wave_group_dec	mean dec from 3D FOF clustering
wave_group_wave	mean wavelength from 3D FOF clustering
fwhm	measured seeing of the observation in arcseconds
throughput	relative spectral response at 4540 assuming a 360 s nominal exposure
shotid	integer represent observation ID: int(date+obsid)
field	field ID: cosmos, goods-n, dex-fall, dex-spring
date	date
obsid	observation number
multiframe	string identifier for the ifuslot/specid/ifuid/amp combination
fiber_id	string identifier for the highest weight fiber
weight	flux weight of the highest weight fiber
x_raw	x-value on the CCD of the detection (ds9 x-value)
y_raw	y-value on the CCD of the detection (ds9 y-value)
x_ifu	x-position in the ifu in arcseconds
y_ifu	y-position in the ifu in arcseconds
ra_aper	R.A. of aperture center of imaging counterpart in degrees
dec_aper	decl. of aperture center of imaging counterpart in degrees
catalog_name_aper	imaging source for measuring O II resolved apertures
filter_name_aper	filter of imaging used for measuring O II resolved apertures
dist_aper	distance between aperture center and detect ID position in arcseconds
mag_aper	photometric magnitude in aperture in imaging source
mag_aper_err	photometric magnitude error in aperture in imaging source
major	major axis of aperture ellipse of resolved O II galaxy defined by imaging
minor	minor axis of aperture ellipse of resolved O II galaxy defined by imaging
theta	angle in aperture ellipse

ORCID iDs

Erin Mentuch Cooper https://orcid.org/0000-0002-2307-0146
Karl Gebhardt https://orcid.org/0000-0002-8433-8185
Dustin Davis https://orcid.org/0000-0002-8925-9769
Daniel J. Farrow https://orcid.org/0000-0003-2575-0652
Chenxu Liu https://orcid.org/0000-0001-5561-2010
Gregory Zeimann https://orcid.org/0000-0003-2307-0629

Robin Ciardullo https://orcid.org/0000-0002-1328-0211
John J. Feldmeier https://orcid.org/0000-0003-2908-2620
Niv Drory https://orcid.org/0000-0002-7339-3170
Donghui Jeong https://orcid.org/0000-0002-8434-979X
William P. Bowman https://orcid.org/0000-0003-4381-5245
Michael Boylan-Kolchin https://orcid.org/0000-0002-9604-343X
Óscar A. Chávez Ortiz https://orcid.org/0000-0003-2332-5505
Maya H. Debski https://orcid.org/0000-0002-1998-5677
Mona Dentler https://orcid.org/0000-0002-5149-2282
Steven L. Finkelstein https://orcid.org/0000-0001-8519-1130
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
Ulrich Hopp https://orcid.org/0000-0003-1008-225X
Lindsay R. House https://orcid.org/0000-0002-1496-6514
Steven Janowiecki https://orcid.org/0000-0001-9165-8905
Hasti Khoraminezhad https://orcid.org/0000-0001-5610-4405
Wolfram Kollatschny https://orcid.org/0000-0002-0417-1494
Eiichiro Komatsu https://orcid.org/0000-0002-0136-2404
Martin Landriau https://orcid.org/0000-0003-1838-8528
Maja Lujan Niemeyer https://orcid.org/0000-0002-6907-8370
Hanshin Lee https://orcid.org/0000-0002-3559-5310
Ken Mawatari https://orcid.org/0000-0003-4985-0201
Masami Ouchi https://orcid.org/0000-0002-1049-6658
Shun Saito https://orcid.org/0000-0002-6186-5476
Donald P. Schneider https://orcid.org/0000-0001-7240-7449
Jan Snigula https://orcid.org/0000-0003-4044-5357
Benjamin P. Thomas https://orcid.org/0000-0002-0977-1974
Sarah Tuttle https://orcid.org/0000-0002-7327-565X
Tanya Urrutia https://orcid.org/0000-0001-6746-9936
Laurel Weiss https://orcid.org/0000-0002-4974-1243
Yechi Zhang https://orcid.org/0000-0003-3817-8739

References

Abareshi, B., Aguilar, J., Ahlen, S., et al. 2022, AJ, 164, 207
Acquaviva, V., Gawiser, E., Leung, A. S., et al. 2014, in IAU Symp. 306, Low/High Redshift Classification of Emission Line Galaxies in the HETDEX Survey (Cambridge: Cambridge Univ. Press), 365
Adams, J. J., Blanc, G. A., Hill, G. J., et al. 2011, ApJS, 192, 5
Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3
Aihara, H., AlSayyad, Y., Ando, M., et al. 2022, PASJ, 74, 247
Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Bacon, R., Brinchmann, J., Richard, J., et al. 2015, A&A, 575, A75
Barbary, K. 2016, JOSS, 1, 58
Barger, A. J., Cowie, L. L., & Wang, W. H. 2008, ApJ, 689, 687
Benitez, N., Dupke, R., Moles, M., et al. 2014, arXiv:1403.5237
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bonoli, S., Marin-Franch, A., Varela, J., et al. 2021, A&A, 653, A31
Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S5
Bradley, L., Sipőcz, B., Robitaille, T., et al. 2021, astropy/photutils: 1.3.0, v1.3.0, Zenodo, doi:10.5281/zenodo.5796924
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Cenarro, A. J., Peletier, R. F., Sanchez-Blazquez, P., et al. 2007, MNRAS, 374, 664
Chiang, C.-T., Wullstein, P., Jeong, D., et al. 2013, JCAP, 2013, 030
Chonis, T. S., Blanc, G. A., Hill, G. J., et al. 2013, ApJ, 775, 99
Ciardullo, R., Gronwall, C., Wolf, C., et al. 2012, ApJ, 744, 110
Cooper, M. C., Aird, J. A., Coil, A. L., et al. 2011, ApJS, 193, 14
Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319
Davis, D., Gebhardt, K., Mentuch Cooper, E., et al. 2021, ApJ, 920, 122
Davis, D., Gebhardt, K., Mentuch Cooper, E., et al. 2023, arXiv:2301.01799
Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Jr., et al. 1991, Third Reference Catalogue of Bright Galaxies (New York: Springer)
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Eriksen, M., Alarcon, A., Gaztanaga, E., et al. 2019, MNRAS, 484, 4200
Falcon-Barroso, J., Sanchez-Blazquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95
Farrow, D. J., Sanchez, A. G., Ciardullo, R., et al. 2021, MNRAS, 507, 3187
Ferreras, I., Pasquali, A., Malhotra, S., et al. 2009, ApJ, 706, 158
Finkelstein, S. L., Ryan, R. E., Jr., Papovich, C., et al. 2015, ApJ, 810, 71
Fitzpatrick, E. L. 1999, PASP, 111, 63
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
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
Grasshorn Gebhardt, H. S., Jeong, D., Awan, H., et al. 2019, ApJ, 876, 32
Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79
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
Hasinger, G., Capak, P., Salvato, M., et al. 2018, ApJ, 858, 77
Hawkins, K., Zeimann, G., Sneden, C., et al. 2021, ApJ, 911, 108
Hill, G. J., Lee, H., MacQueen, P. J., et al. 2021, AJ, 162, 298
Hill, G. J., MacQueen, P. J., Smith, M. P., et al. 2008, Proc. SPIE, 7014, 701470
Horne, K. 1986, PASP, 98, 609
Hu, E. M., & Cowie, L. L. 2006, Natur, 440, 1145
Konno, A., Ouchi, M., Nakajima, K., et al. 2016, ApJ, 823, 20
Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15
Kriek, M., van Dokkum, P. G., Franx, M., et al. 2008, ApJ, 677, 219
Kulas, K. R., Shapley, A. E., Kollmeier, J. A., et al. 2012, ApJ, 745, 33
Laseter, I. H., Finkelstein, S. L., Bagley, M. J., et al. 2022, ApJ, 940, 9
Leung, A. S., Acquaviva, V., Gawiser, E., et al. 2017, ApJ, 843, 130
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218
Liu, C., Gebhardt, K., Mentuch Cooper, E., et al. 2022, ApJS, 261, 24
Green, G. 2018, JOSS, 3, 695
Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2017, ApJ, 842, 21
McCarron, A. P., Finkelstein, S. L., Chavez Ortiz, O. A., et al. 2022, ApJ, 936, 131
Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, PASJ, 70, S1
Moffat, A. F. J. 1969, A&A, 3, 455
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Morton, D. C. 1991, ApJS, 77, 119
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Nilson, P. 1973, Uppsala General Catalogue of Galaxies (Uppsala Astronomiska Observatoriums Annaler: Acta Universitatis Upsaliensis)
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Ono, Y., Itoh, R., Shibuya, T., et al. 2021, ApJ, 911, 78
Ouchi, M., Ono, Y., & Shibuya, T. 2020, ARA&A, 58, 617
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
Oyarzún, G. A., Blanc, G. A., González, V., Mateo, M., & Bailey, J. I. I. 2017, ApJ, 843, 133
Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, ApJ, 674, 1217
Pâris, I., Petitjean, P., Aubourg, É., et al. 2018, A&A, 613, A51
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A1
Ramsey, L. W., Adams, M. T., Barnes, T. G., et al. 1998, Proc. SPIE, 3352, 34
Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, ApJ, 653, 1004
Rhoads, J. E., Malhotra, S., Dey, A., et al. 2000, ApJL, 545, L85
Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, MNRAS, 498, 2354
Santos, S., Sobral, D., Matthee, J., et al. 2020, MNRAS, 493, 141
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shoji, M., Jeong, D., & Komatsu, E. 2009, ApJ, 693, 1404
Sobral, D., Santos, S., Matthee, J., et al. 2018, MNRAS, 476, 4725
Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, ApJ, 825, 5
Spinoso, D., Orsi, A., López-Sanjuan, C., et al. 2020, A&A, 643, A149

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Urrutia, T., Wisotzki, L., Kerutt, J., et al. 2019, A&A, 624, A141
van Breukelen, C., Jarvis, M. J., & Venemans, B. P. 2005, MNRAS, 359, 895
Van Rossum, G., & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace), https://dl.acm.org/doi/book/10.5555/1593511
Vinko, J., Thomas, B.P., Weeler, C., et al. 2022, arXiv:2212.08444
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wirth, G. D., Trump, J. R., Barro, G., et al. 2015, AJ, 150, 153
Wirth, G. D., Willmer, C. N. A., Amico, P., et al. 2004, AJ, 127, 3121
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
Zhang, Y., Ouchi, M., Gebhardt, K., et al. 2021, ApJ, 922, 167